Simone Bassis
Anna Esposito
Francesco Carlo Morabito   *Editors*

Advances in Neural Networks: Computational and Theoretical Issues

Springer

Smart Innovation, Systems and Technologies

Volume 37

About this Series

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

More information about this series at http://www.springer.com/series/8767

Simone Bassis · Anna Esposito
Francesco Carlo Morabito
Editors

Advances in Neural Networks: Computational and Theoretical Issues

Springer

Editors
Simone Bassis
Computer Science Department
University of Milano
Milano
Italy

Anna Esposito
Dipartimento di Psicologia, Seconda
    Universitá di Napoli, Caserta, Italy

and

International Institute for Advanced
    Scientific Studies (IIASS)
Vietri sul Mare (SA)
Italy

Francesco Carlo Morabito
Department of Civil, Environmental,
    Energy, and Material Engineering
University Mediterranea of
    Reggio Calabria
Reggio Calabria
Italy

Preface

This research book aims to provide the reader with a selection of high-quality papers devoted to current progress and recent advances in the now mature field of Artificial Neural Networks (ANN). Not only relatively novel models or modifications of current ones are presented, but many aspects of interest related to their architecture and design are proposed, which include the data selection and preparation step, the feature extraction phase, and the pattern recognition procedures.

This volume focuses on a number of advances topically subdivided in Chapters. In particular, in addition to a group of Chapters devoted to the aforementioned topics specialized in the field of intelligent behaving systems using paradigms that can imitate human brain, three Chapters of the book are devoted to the development of automatic systems capable to detect emotional expression and support users' psychological well-being, the realization of neural circuitry based on "memristors", and the development of ANN applications to interesting real-world scenarios.

This book easily fits in the related Series, like an edited volume, containing a collection of contributes from experts, and it is the result of a collective effort of authors jointly sharing the activities of SIREN Society, the Italian Society of Neural Networks.

May 2015

Anna Esposito
Simone Bassis
Francesco Carlo Morabito

Acknowledgments

The editors express their deep appreciation to the referees listed below for their valuable reviewing work.

Referees

Sponsoring Institutions

Contents

Part VI: Emotional Expressions and Daily Cognitive Functions

Part VII: Memristor and Complex Dynamics in Bio-inspired Networks

Part I
Introductory Chapter

Recent Advances of Neural Networks Models and Applications: An Introduction

Anna Esposito[1], Simone Bassis[2], and Francesco Carlo Morabito[3]

[1] Second University of Napoli, Department of Psychology and IIASS, Italy
[2] University of Milano, Department of Computer Science, Italy
[3] University "Mediterranea" of Reggio Calabria,
Department of Civil Engineering, Energy, Environment and Materials (DICEAM), Italy
`iiass.annaesp@tin.it, bassis@di.unimi.it, morabito@unirc.it`

Abstract. Recently, increasing attention has been paid to the development of approximate algorithms for equipping machines with an automaton level of intelligence. The aim is to permit the implementation of intelligent behaving systems able to perform tasks which are just a human prerogative. In this context, neural network models have been privileged, thanks to the claim that their intrinsic paradigm can imitate the functioning of the human brain. Nevertheless, there are three important issues that must be accounted for the implementation of a neural network based autonomous system performing an automaton human intelligent behavior. The first one is related to the collection of an appropriate database for training and evaluating the system performance. The second issue is the adoption of an appropriate machine representation of the data which implies the selection of suitable data features for the problem at hand. Finally, the choice of the classification scheme can impact on the achieved results. This introductive chapter summarizes the efforts that have been made in the field of neural network models along the abovementioned research directions through the contents of the chapters included in this book.

Keywords: Neural network models, behaving systems, feature selection, big data collection.

1   Introduction

Human-machine based applications turn out to be increasingly involved in our personal, professional and social life. In this context, human expectations and requirements become more and more highly structured, up to the desire to exploit them in most environments, in order to decrease human workloads and errors, as well as to be able to interact with them in a natural way. Along these directions, neural network models have been privileged because of their computational paradigm based on brain functioning and learning. However, it has soon become evident that, in order for machines to show autonomous behaviors, it would not suffice to exploit human learning and functioning paradigms. There are issues related to database collection, feature selection and classification schema that must be accounted for in order to

3

obtain computational effectiveness and optimal performance. These issues are briefly discussed in Sections 2 to 4. Section 5 summarizes the contents of this book by grouping the received contributions into 5 different sections devoted to the use of neural networks for applications, new or improved models, pattern recognition, signal processing and special topics such as emotional expressions and daily cognitive functions, as well as bio-inspired networks memristor-based.

2   The Data Issue

In training and assessing neural networks as a paradigm for complex systems to show autonomous behaviors, the first issue that arises is the appropriateness of the data exploited for it. It has become evident that system performances strongly depend on the database used and the related complexity of the task. If the database is poor in reproducing the features of the task at hand, inaccurate inferences can be drawn, and the trained neural system cannot perform accurately on other similar data. Therefore, it is necessary to assess the database in order to ascertain if it reproduces a genuine setting of the real world environment it aims to describe. The questions that must then be raised in order to define the suitability of the data are:

a) Have data been collected in a natural or artificial context? As an example, this can be necessary if the system must discriminate among genuine emotional speech or real world seismic signals, as opposed to acted emotional speech or synthetic signals [3,4,6];
b) Are data equally balanced among the categories the system must discriminate? In this case, consider as an instance a speech recognition task. If gender is not an issue, then the data must be equally balanced between male and female subjects;
c) Are data representative of the final application they are devoted to? This last question calls for the importance, in designing the database, of the actual task the system is designed for.

3   Feature Selection

This issue relates to the way the data are processed in order to extract from them suitable features efficiently describing the different categories among those the system must discriminate for the task at hand. The selection of features can be very hard and difficult depending on the task. An interesting example to describe this problem is to consider a speech emotional recognition task. In this case, the features selection task can be simple (as for a speaker dependent approach [17]) or very complex (if the task is speaker independent [3,4]) and even more in a noisy environment (as in the case of speech collected through phone calls [1,7]). The features selection procedure is strongly dependent on the data and the task, and its effectiveness relies on the knowledge the experimenter applies to understand data and identify features for them, as illustrated by Likforman-Sulem et al. in this volume and deeply explained in [14]. In addition, features from different sources can be combined

and fused, as it is tradition in the field of speech, where linguistic (such as language and word models [12]) and/or prosodic information (such as F0 contour [19]) and visual features (such as action units [13] are fused with acoustic features [8,20]. Automatic approach to feature selection can produce a huge amount of features [2] making hard the neural network training process. Of course, the relevance of this step is not limited to speech signal processing (see, for example, [21]).

4 Classification Schema

There are several classification schema proposed in literature for detection and classification tasks. The most exploited are Artificial Neural Networks (ANN) Gaussian Mixture Models (GMM), Hidden Markov Models (HMM), and Support Vector Machine (SVM) [9,10,18,22]. Advantage and drawbacks in their use have been reviewed recently in [11]. It is not the aim of this short chapter to go deep inside the problematics of the different classification schema. However, it is important to point out that they can be fused together in more complex models as reported in [15] or be complicated by sophisticated learning algorithms as those related to deep learning architectures, illustrated by Schuller in this volume and deeply explained in [5].

5 Contents of This Book

For over twenty years, Neural Networks and Machine Learning (NN/ML) have been an area of continued growth. The need for a Computational (bioinspired) Intelligence has increased dramatically for various reasons in a number of research areas and application fields, spanning from Economic and Finance, to Health and Bioengineering, up to the industrial and entrepreneurial world. Besides the practical interest in these approaches, the progress in NN/ML derives from its interdisciplinary nature.

This book is a follow-up of the scientific workshop on Neural Network held in Vietri sul Mare, Italy in May 15-16th 2014, as a continued tradition since its founder, Professor Eduardo Caianiello, thought to it as a way of exchanging information on worldwide activities on the field. The volume brings together the peer-reviewed contributions of the attendees: each paper is an extended version of the original submission (not elsewhere published) and the whole set of contributions has been collected as chapters of this book. It is worth emphasizing that the book provides a balance between the basics, evolution, and NN/ML applications.

To this end, the content of the book is organized in six parts: four general sections are devoted to Neural Network Models, Signal Processing, Pattern Recognition, and Neural Network Applications; two sections focused on more specialized topics, namely, "Emotional Expression and Daily Cognitive Functions" and "Memristors and Complex Dynamics in Bio-inspired Networks".

This organization aims indeed at reflecting the wide interdisciplinarity of the field, which on the one hand is capable of motivating novel paradigms and relevant improvement on known paradigms, while, on the other hand, is largely accepted in

many applicative fields as an efficient and effective way to solve classification, detection, identification and related tasks.

In Chapter 2 either novel ways to apply old learning paradigms or recent updates to new ones are proposed. To this aim the chapter includes six contributions respectively on Belief propagation in Normal Factor Graphs (proposed by Buonanno et al.), Genetic Embedding and NN regression (proposed by Panella et al.), Echo-State Networks and Pruning for Reservoir's Neurons (proposed by Scardapane et al.), Functional Link (proposed by Comminiello et al.), Continuous-Time Spiking Neural Networks (proposed by Cristini et al.) and Online Spectral Clustering (proposed by Rovetta & Masulli).

Chapter 3 presents interesting signal processing procedures and results obtained using either Neural Networks or Machine Learning techniques. In this context, section 1 (proposed by Labate et al.) describes an Empirical Mode Decomposition (EMD) to diagnose brain diseases. The following section reports on the effects of artifact rejection and the complexity of EEG (Labate et al., 2015b). Section 3 (proposed by D'Auria et al.) describes the ability of Self-Organizing Maps to de-noise real world as well as synthetic seismic signals, explaining how a self-learning algorithm would be preferable in this context. The following two sections in this chapter focus respectively on the integration of audio and video clues for source localization (by Parisi et al.) and an integrated system based on Spiking Neural Networks known as NeuCube (by Capecci et al.) to model EEGs in Alzheimer Disease data.

Chapter 3 main objective is to illustrate pattern recognition procedures defined through neural networks and machine learning algorithms. To this aim, Camastra et al. propose semantic graphs for document characterization, while Graph Neural Networks are used for web spam detection by Belahcen et al. Some complex network concepts, like hubs and communities, are proposed (by Mahmoud et al.) in financial applications. The last section of this chapter (proposed by Di Nardo et al.) presents a video-based access control by automatic license plate recognition.

Chapter 4 is devoted to various applications of ML/NN. They span different research fields such as behavioral analysis in maritime environment (by Castaldo et al.), forecasting of domestic water and natural gas demand (by Fagiani et al.), referenceless thermometry (by Agnello et al.), risk assessment (by Cardin and Giove), fingerprint classification (by Vitello et al.), FEEM sustainable composite indicator (by Farnia and Giove); autonomous physical rehabilitation at home (by Borghese et al.) and building automation systems (by De March et al.).

Chapter 5 is devoted to illustrate the contributions that were submitted to the workshop special session on emotional expressions and daily cognitive functions organized by Anna Esposito, Vincenzo Capuano and Gennaro Cordasco form the International Institute for Advanced Scientific Studies (IIASS) and the Second University of Napoli (Department of Psychology). The session intended to collect contributes on the current efforts of research for developing automatic systems capable to detect and support users' psychological wellbeing. To this aim the proposed contributions were on behavioral emotional analysis and perceptual experiments aimed to the identification of cues for detecting healthy and/or non-healthy psychological/physical states such as stress, anxiety, and emotional disturbances, as well as cognitive declines from a social and

psychological perspective. These aspects are covered by the contributions proposed by Esposito et al., as well as, Maldonato and Dell'Orco, Matarazzo and Baldassarre, Baldassarre et al., Hristova and Grinberg, Senese et al, Gnisci et al., included in this volume. In addition, the special session was also devoted to show possible applications and algorithms, biometric and ICT technologies to design innovative and adaptive systems able to detect such behavioral cues as a multiple, theoretical, and technological investment. These aspects are covered by the sections proposed by Schuller, as well as, Likforman et al., and Faundez-Zanuy et al.

Chapter 6 includes five papers on Memristive NN, a fast developing field for NN neurons and synapses implementation based on the original concept invented by Leon Chua, in 1971 [16]. They have been presented within the related session, organized by Fernando Corinto and Eros Pasero from the Polytechnic of Milano, Italy. Memristive systems are used for the synchronization of two Rossler oscillators (in Frasca et al.); for realizing an electrostatic loudspeaker (by Troiano et al.); for an analogic implementation of nonlinear networks in complex dynamic analysis (by Petrarca et al.); for high efficient learning with binary synapses circuitry (by Secco et al.); for quantum-inspired optimization techniques (by Fiaschè).

The nature of an edited volume like this, containing a collection of contributions from experts that have been first presented and discussed at the WIRN 2014 Workshop, and then developed in a full paper is quite different from a journal or a conference publication. Each work has been left the needed space to present the details of the proposed topic. The chapters of the volume have been organized in such a manner that the readers can easily seek for additional information from a vast number of cited references. It is our hope the book can contribute to the progress of NN/ML related methods and to their spread to many different fields, as it was in the original spirit of the SIREN (Italian Society of Neural Networks – Società Italiana REti Neuroniche) Society.

References

1. Atassi, H., Smékal, Z., Esposito, A.: Emotion recognition from spontaneous Slavic speech. In: Proceedings of 3rd IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2012), Kosice, Slovakia, December 2-5, pp. 389–394 (2012)
2. Atassi, H., Esposito, A., Smekal, Z.: Analysis of high-level features for vocal emotion recognition. In: Proceedings of 34th IEEE International Conference on Telecom. and Signal Processing (TSP), Budapest, Hungary, August 18-20, pp. 361–366 (2011)
3. Atassi, H., Riviello, M.T., Smékal, Z., Hussain, A., Esposito, A.: Emotional vocal expressions recognition using the COST 2102 Italian database of emotional speech. In: Esposito, A., Campbell, N., Vogel, C., Hussain, A., Nijholt, A. (eds.) Second COST 2102. LNCS, vol. 5967, pp. 255–267. Springer, Heidelberg (2010)
4. Atassi, H., Esposito, A.: Speaker independent approach to the classification of emotional vocal expressions. In: Proceedings of IEEE Conference on Tools with Artificial Intelligence (ICTAI 2008), Dayton, OH, USA, November 3-5, vol. 1, pp. 487–494 (2008)
5. Bengio, Y.: Learning Deep Architectures for AI. Foundations and Trends in Machine Learning 2(1), 1–127 (2009)

6. D'Auria, L., Esposito, A.M., Petrillo, Z., Siniscalchi, A.: Denoising magnetotelluric recordings using Self-Organizing Maps. In: Bassis, S., Esposito, A., Morabito, F.C. (eds.) Recent Advances of Neural Networks Models and Applications. SIST, vol. 37, pp. 139–149. Springer, Heidelberg (2015)

7. Galanis, D., Karabetsos, S., Koutsombogera, M., Papageorgiou, H., Esposito, A., Riviello, M.T.: Classification of emotional speech units in call centre interactions. In: Proceedings of 4th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2013), Budapest, Hungary, December 2-5, pp. 403–406 (2013)

8. Karunaratnea, S., Yanb, H.: Modelling and combining emotions, visual speech and gestures in virtual head models. Signal Processing: Image Comm. 21, 429–449 (2006)

9. Kwon, O., Chan, K., Hao, J., Lee, T.: Emotion recognition by speech signal. In: Proceedings of EUROSPEECH 2003, Geneva, Switzerland, September 1-4, pp. 125–128 (2003)

10. Labate, D., Palamara, I., Mammone, N., Morabito, G., Foresta, F.L., Morabito, F.C.: SVM classification of epileptic EEG recordings through multiscale permutation entropy. In: Proc. of Int. Joint Conf. on Neural Networks (IJCNN), Dallas, TX, USA, August 4-9 (2013)

11. Larochelle, H., Erhan, D., Courville, A., Bergstra, J., Bengio, Y.: An empirical evaluation of deep architectures on problems with many factors of variation. In: Proc. of 24th Int. Conf. on Machine Learning (ICML 2007), Corvallis, OR, USA, June 20-24, pp. 473–480 (2007)

12. Lee, C., Pieraccini, R.: Combining acoustic and language information for emotion recognition. In: Proceedings of the ICSLP 2002, pp. 873–876 (2002)

13. Lien, J., Kanade, T., Li, C.: Detection, tracking and classification of action units in facial expression. J. Robotics Autonomous Syst. 31(3), 131 (2002)

14. Lin, F., Liang, D., Yeh, C.-C., Huang, J.-C.: Novel feature selection methods to financial distress prediction. Expert Systems with Applications 41(5), 2472–2483 (2014)

15. Mohamed, A., Dahl, G.E., Hinton, G.: Acoustic Modeling Using Deep Belief Networks. IEEE Transactions on Audio, Speech, and Language Processing 20(1), 14–22 (2012)

16. Morabito, F.C., Andreou, A.G., Chicca, E.: Neuromorphic engineering: from neural systems to brain-like engineered systems. Neural Networks 45, 1–3 (2013)

17. Navas, E., Luengo, H.I.: An objective and subjective study of the role of semantics and prosodic features in building corpora for emotional TTS. IEEE Transactions on Audio, Speech, and Language Processing 14, 1117–1127 (2006)

18. Ou, G., Murphey, Y.L.: Multi-class pattern classification using neural networks. Pattern Recognition 40, 4–18 (2007)

19. Ishi, C.T., Ishiguro, H., Hagita, N.: Automatic extraction of paralinguistic information using prosodic features related to F0, duration and voice quality. Speech Communication 50(6), 531–543 (2008)

20. Schuller, B., Rigoll, G., Lang, M.: Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief-network architecture. In: Proceedings of the ICASSP 2004, vol. 1, pp. 577–580 (2004)

21. Simone, G., Morabito, F.C., Polikar, R., Ramuhalli, P., Udpa, L., Udpa, S.: Feature extraction techniques for ultrasonic signal classification. International Journal of Applied Electromagnetics and Mechanics 15(1-4), 291–294 (2001)

22. Vlassis, N., Likas, A.: A greedy EM algorithm for Gaussian mixture learning. Neural Process. Lett. 15, 77–87 (2002)

Part II
Models

Simulink Implementation of Belief Propagation in Normal Factor Graphs

Amedeo Buonanno and Francesco A.N. Palmieri

Seconda Università di Napoli (SUN)
Dipartimento di Ingegneria Industriale e dell'Informazione,
via Roma 29, 81031 Aversa (CE) - Italy
`{amedeo.buonanno,francesco.palmieri}@unina2.it`

Abstract. A Simulink Library for rapid prototyping of belief network architectures using Forney-style Factor Graph is presented. Our approach allows to draw complex architectures in a fairly easy way giving to the user the high flexibility of Matlab-Simulink environment. In this framework the user can perform rapid prototyping because belief propagation is carried in a bi-directional data flow in the Simulink architecture. Results on learning a latent model for artificial characters recognition are presented.

Keywords: Belief Propagation, Factor Graph, Pattern Recognition, Machine Learning.

1 Introduction

Graphical models are a "marriage between probability theory and graph theory" [1] as they compactly encode complex distributions over a high-dimensional space. When a problem can be formulated in the form of a graph, it is very appealing to study the variables involved as part of an interconnected system where the reached equilibrium point is the solution. The similarities with the working of the nervous system makes this paradigm even more fascinating [2]. Bayesian inference on graphs, pioneered by Pearl [3], has become a very popular paradigm for approaching many problems in different fields such as communication, signal processing and artificial intelligence [4]. The Factor Graph is a particular type of Graphical model and represents an interesting way to model the interaction between stochastic variables. Following the formulation of Forney-style Factor Graphs (FFG) [5] (or normal graphs), Bayesian graphs can be drawn as block diagrams and probability distribution easily transformed and propagated. In this paper we report the results of our work in which we have designed and implemented a Simulink Library for quick prototyping of several network architectures using the FFG paradigm.

In Section 2 we briefly review the Factor Graph paradigm introducing the building blocks of our proposed Simulink Library. In Section 3 the two operating modes are introduced. In Section 4 we present the application of this tool to an artificial character recognition task.

2   Simulink Factor Graph Library

Factor Graphs model the interaction among stochastic variables. In the FFG approach there are blocks, variables and directed edges [5]. Even if edges have a defined direction, probability flows in both directions (foward and backward) [4]. To associate to each stochastic variable two messages, we have used the built-in Two-Way Connection block that in Simulink allows bidirectional signal flow. In our Simulink implementation all the architectures can be built with just three main functional blocks: Variable, Factor and Diverter (Figure 1) that will be described in the folllowing. In our notation, we avoid the upper arrows [4] and use explicit letters: b for *backward* and f for *forward*.
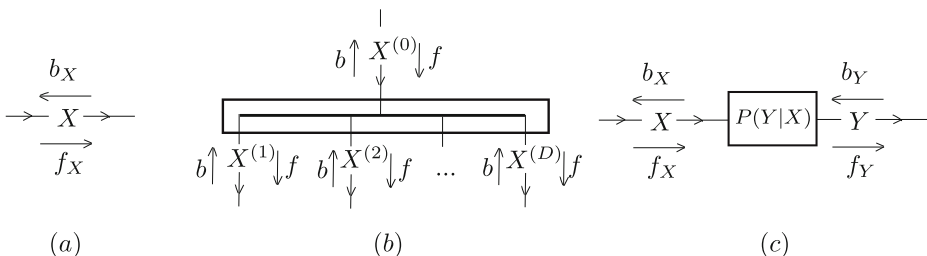
Fig. 1. Functional Blocks: (a) Variable, (b) Diverter, (c) Factor

2.1   Variable

For a *variable* X (Figure 1(a)) that takes values in the discrete alphabet $\mathcal{X} = \{x^1, x^2, ..., x^{M_X}\}$, forward and backward messages are in function form:

$$b_X(x^i), f_X(x^i), \quad i = 1 : M_X$$

and in vector form

$$\mathbf{b}_X = (b_X(x^1), b_X(x^2), ..., b_X(x^{M_X}))^T$$
$$\mathbf{f}_X = (f_X(x^1), f_X(x^2), ..., f_X(x^{M_X}))^T$$

All messages are proportional ($\propto$) to discrete distributions and may be normalized to sum to one. Comprehensive knowledge about X is contained in the distribution p_X obtained through the product rule (in function form):

$$p_X(x^i) \propto f_X(x^i) b_X(x^i), \quad i = 1 : M_X$$

or $\mathbf{p}_X \propto \mathbf{f}_X \odot \mathbf{b}_X$, in vector form, where $\odot$ denotes the element-by-element product.

Each message b, f or p in the data flow is an $nT \times M$ matrix with nT the number of realizations and M the variable cardinality. Two-way connection blocks

allow the construction of a bi-directional data flow. The implementation for an Internal Variable block is shown in Figure 2 where the forward message on the port up (f_b_up) is transmitted on the port down (f_b_down) and conversely the backward message on the port down is transmitted on the port up. All distribution flow can be saved to workspace.
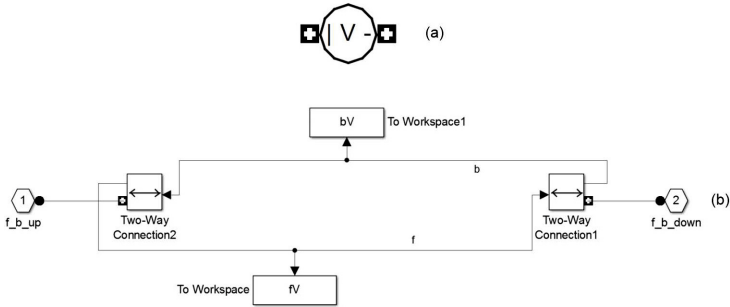
Fig. 2. The implementation of the Internal Variable block. The icon in the library (a) and its detailed scheme (b)

Similarly Figure 3 shows the detailed schemes of Source and Sink Variable blocks.
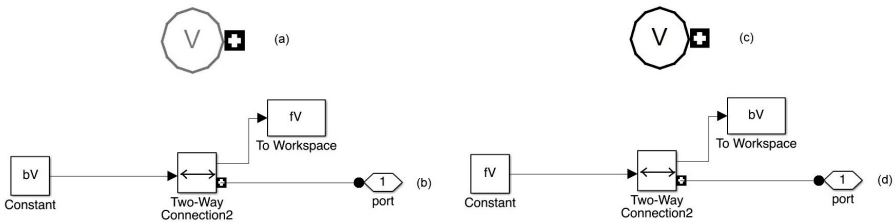
Fig. 3. The implementation of the Source Variable block and of the Sink Variable block. The icon in the library (a,c) and its detailed scheme (b,d) respectively for the Source and for the Sink

2.2 Diverter Block

The *diverter* block (Figure 1(b)) in the Bayesian model represents the equality constraint with the variable X replicated $D + 1$ times. Messages for incoming and outgoing branches carry different forward and backward information.

Messages that leave the block are obtained as the product of the incoming ones (in function form):

$$b_{X^{(0)}}(x^i) \propto \prod_{j=1}^{D} b_{X^{(j)}}(x^i)$$

$$f_{X^{(m)}} \propto f_{X^{(0)}}(x^i) \prod_{j=1, j\neq m}^{D} b_{X^{(j)}}(x^i), \quad m = 1 : D, i = 1 : M_X$$

In vector form:

$$\mathbf{b}_{X^{(0)}} \propto \odot_{j=1}^{D} \mathbf{b}_{X^{(j)}},$$

$$\mathbf{f}_{X^{(m)}} \propto \mathbf{f}_{X^{(0)}} \odot_{j=1, j\neq m}^{D} \mathbf{b}_{X^{(j)}}, \quad m = 1 : D$$

Figure 4 shows the detailed scheme of our implementation of the Diverter Block. Each port is connected to a variable in the network. After element-wise product among variables each variable is returned after normalization to one (each message is normalized to be a valid distribution).
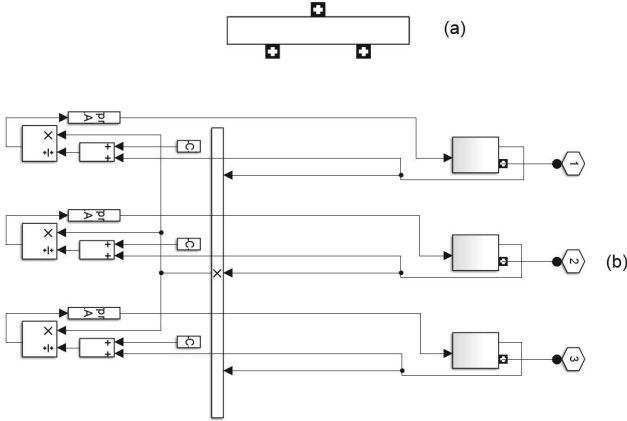
Fig. 4. Simulink implementation of a Diverter Block with three ports. The icon in the library (a) and its detailed scheme (b)

2.3   Factor Block

The *factor* block (Figure 1(c)) is the main block that represents the conditional probability matrix of Y given X. More specifically if X takes values in the discrete alphabet $\mathcal{X} = \{x^1, x^2, ..., x^{M_X}\}$ and Y in $\mathcal{Y} = \{y^1, y^2, ..., y^{M_Y}\}$, $P(Y|X)$ is the $M_X \times M_Y$ row-stochastic matrix:

$$P(Y|X) = [Pr\{Y = y^j | X = x^i\}]_{i=1:M_X}^{j=1:M_Y} = [\theta_{ij}]_{i=1:M_X}^{j=1:M_Y} = \theta$$

Outgoing messages are (in function form):

$$f_Y(y^j) \propto \sum_{i=1}^{M_X} \theta_{ij} f_X(x^i), \quad b_X(x^i) \propto \sum_{j=1}^{M_Y} \theta_{ij} b_Y(y^j)$$

In vector form:
$$\mathbf{f}_Y \propto P(Y|X)^T \mathbf{f}_X, \quad \mathbf{b}_X \propto P(Y|X)\mathbf{b}_Y$$

The above rules are rigorous translation of Bayes' theorem and marginalization (a complete review and proofs can be found in classical papers [4], [6]).

Figure 5 shows our implementation of the Factor Block with a Level2-MATLAB S-Function that wraps the Maximum Likelihood (ML) algorithm described in [7]. The system learns locally using nT realizations of the forward message of variable X, the nT realizations of backward message of variable Y and an initial value of matrix P. During learning, a new value of P is produced on each epoch and nT realizations of backward message for variable X and forward message for Y are sent to the adjacent blocks.

If the number of iteration is set to 0, the Block simply computes the nT realizations of backward of variable X and the nT realizations of forward message of variable Y (using the results in [8]).
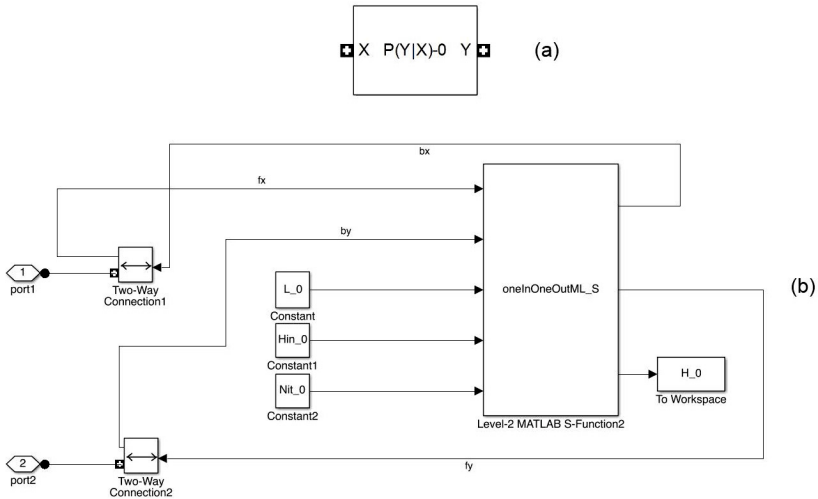
Fig. 5. Simulink implementation of the Factor Block. The icon in the library (a) and its detailed scheme (b) - During learning phase, given the initial value of Conditional Probability Matrix (Hin), the bacward messages for variable Y, the forward messages for variable X and the learning mask (L), a new value of H is computed applying Nit iterations of ML algorithm. If the Nit is set to 0, the block works in inference mode.

Using the implemented library, simply by dragging and connecting, the user can define a wide range of architectures that otherwise would have required the
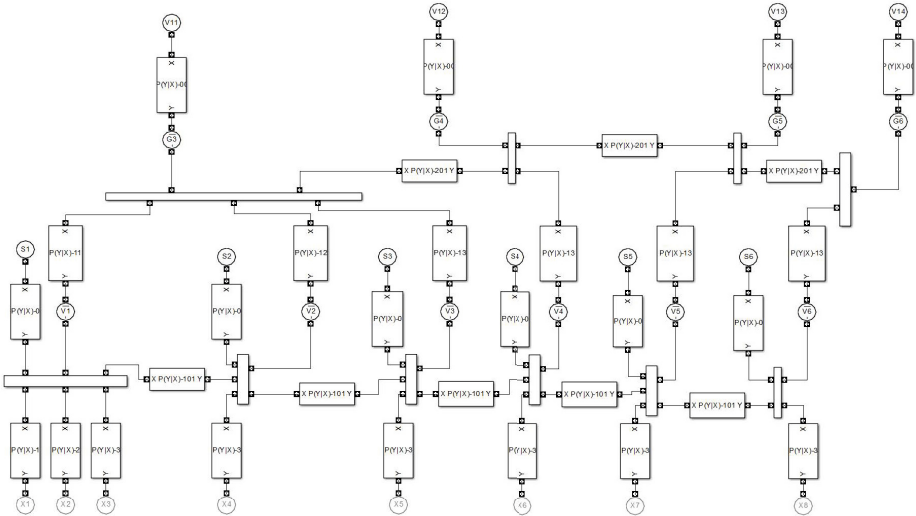
Fig. 6. A complex architecture designed using the proposed library

writing of a custom algorithm of belief propagation. Figure 6 shows a complex network drawn using the building blocks previously introduced.

3   Flow Control

During the simulation, each block uses messages coming from connected blocks and evolves producing new messages. The distributions exchanged among blocks are bi-directional and simultaneous, but the network flow is controlled from the top by a MATLAB script that sets parameters, triggers execution and collects results. The network can work in *Inference Mode*, when the block parameters are fixed, and in *Learning Mode*, when the block parameters are learned. In the Learning Phase (Figure 7(a)), based on epochs, after the Network Initialization (set to uniform all the variables, set the dimension of the messages), the model simulation is started defining purposely the Simulation Time and Model Parameters (values of Factors). At the end of simulation the new Model Parameters are used as initialization values for next epoch. This is done until the Maximum Number of Epochs is reached. In the Evolution Phase (Figure 7(b)), in the Parameter Initialization, the user has to adopt the correct values of parameters learned during Learning Phase.

The Model Simulation step is performed in the Simulink environment that has to be purposely configured using Fixed-Step Solver Type and with a Fixed Size Time Step. During the updating phase of simulation, Simulink determines the order in which the block methods must be triggered. The user cannot explicitly change this order, but he can assign priorities to non virtual blocks to indicate to Simulink their execution order relative to other blocks. Simulink tries to honor
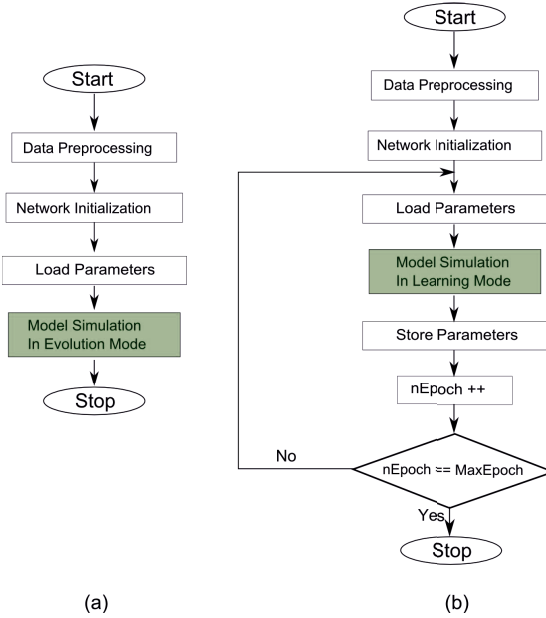
Fig. 7. Scheme for model simulation in the Inference mode (a) and in the Learning mode (b)

block priority settings, unless there is a conflict with data dependencies [9]. We have verified that Simulink automatically assigns the correct execution order, evaluating the From Workspace block (in the source blocks) and then the other blocks. To avoid wrongly assigned variables, each variable in each block is initialized with an uniform distribution. Each block automatically determines the dimension of the variable to which it is connected. During the simulation, each block uses the inputs coming from other blocks and evolves producing output to connecting blocks using the rules outlined in [8].

4   Characters Recognition Example

We have used the proposed Library in several applications. In this work we present the result obtained with a simple Latent Model applied to a recognition task on the Artificial Characters Dataset [10]. This dataset is formed by thousands of 12x8 black and white images representing the characters {'A', 'C', 'D', 'E', 'F', 'G', 'H', 'L', 'P', 'R'}. The network we have implemented is composed of 96 factors (a factor for each pixel) and only one hidden variable.

An image is a matrix of pixels, where each pixel can be considered as a stochastic variable that can assume value in a finite alphabet (2 symbols for black and white images). We have a set of random variables $\{X_1, X_2, ..., X_n\}$ that belong to
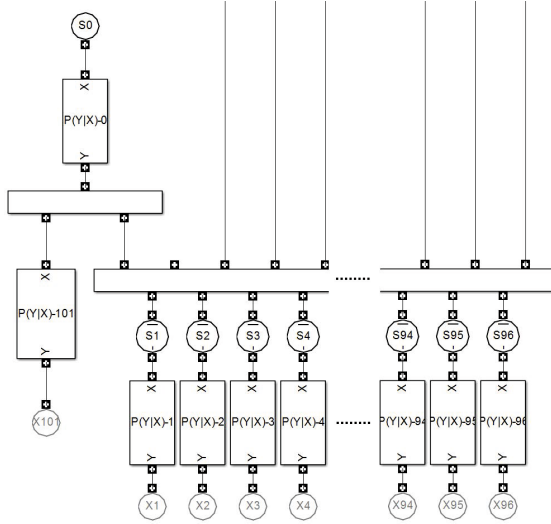
Fig. 8. The designed network for Artificial Characters recognition task using the implemented Library

a same finite alphabet $\mathcal{X}$. This set of variables is fully characterized by its joint probability mass function $p(X_1, X_2, ..., X_n)$. All the mutual interactions among the variables is contained in the structure of p. A variable can be: 1) known (instantiated): the backward message is the delta distribution; 2) completely unknown (erased): the backward message is a uniform distribution; 3) known softly: the backward message is a density. In all cases after message propagation the system responds with a forward message that is related to information stored in the system during the learning phase [11]. We use a simple Latent Model where each variable X_i (pixel) is connected to a Latent Variable (Figure 8) and there is also a Variable that contains the information of the presented character (X_{101}). In the Learning Phase the instantiated variables of training examples are injected in the network and using the ML algorithm in [7] the matrices $P(Y|X) - i$ are learned.

4.1   A Simulation

Using the Artificial Characters Dataset [10] we have trained our network with 800 training images of 12x8 black and white images representing the characters: {'A', 'C', 'D', 'E', 'F', 'G', 'H', 'L', 'P', 'R'} (Figure 9). The dimension of the embedding space is set to 150. The number of epochs for learning phase is set to 20 and each epoch is formed by 10 evolution steps.

To store all configurations the embedding space should have been set to 2^{96}, but the real configurations are much less. We limited the embedding space to 150 because computational issues. Even if we have used a small dimension of the embedding space, the system stores relevant structures of the presented images
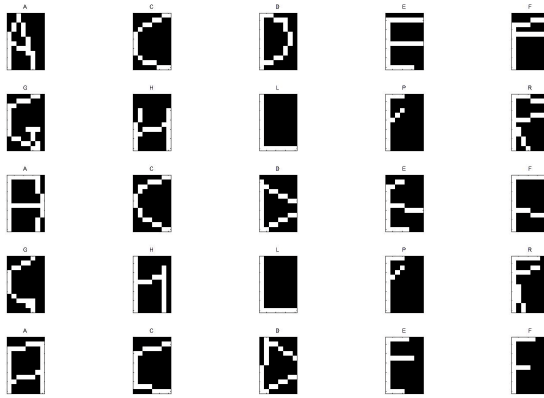
Fig. 9. 25 samples from the Training Set
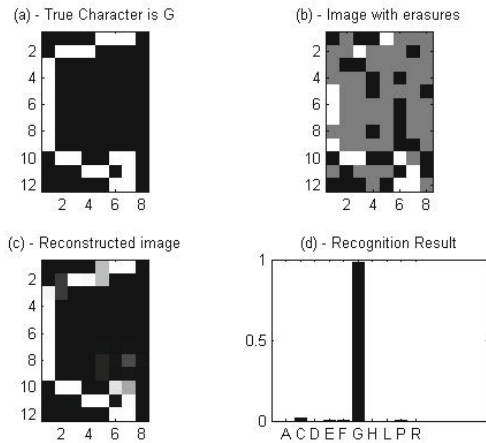
Fig. 10. Network answer - An image is retrieved from the Test Set (a), a big percentage of pixels are erased (gray pixels in (b)) and this information is injected in the network as backward messages. The network, after evolution, returns the Reconstructed image (c) and a probability distribution on the character set (d))

and presenting 800 test images, the system recognize the characters presented with an accuracy of 76%.

In Figure 10 the results of the recognition and completion task are presented. An image is retrieved from Test Set (Figure 10 (a)), a big percentage of pixels are erased (gray pixels in (Figure 10 (b))) and this information is injected in the network as backward messages of Source variables. The information about the presented character is set to uniform. The network, after the evolution (Inference Mode) returns the forward messages of Source variables that, combined

with the provided backward messages, give us the Reconstructed image (Figure 10 (c)). The network provides also the probability distribution on whole vocabulary (Figure 10 (d))

5   Conclusion

We have implemented a Library of Simulink blocks that permits to rapidly design a wide range of architectures using the Factor Graph paradigm. This approach allows to experiment on different architectures using Simulink bi-directional connections as probability pipelines. Current efforts are devoted to use this paradigm for various applications and to find more efficient implementations when the architectures grow in size and complexity.

References

1. Jordan, M. (ed.): Learning in Graphical Models. MIT Press (1998)
2. Hawkins, J.: On Intelligence (with Sandra Blakeslee). Times Books (2004)
3. Pearl, J.: Probabilistic reasoning in intelligent systems - networks of plausible inference. Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann (1989)
4. Loeliger, H.A.: An introduction to factor graphs. IEEE Signal Processing Magazine 21(1), 28–41 (2004)
5. Forney, G.D.: Codes on graphs: Normal realizations. IEEE Transactions on Information Theory 47(2), 520–548 (2001)
6. Kschischang, F., Member, S., Frey, B.J., Loeliger, H.-A.: Factor graphs and the sum-product algorithm. IEEE Transactions on Information Theory 47, 498–519 (2001)
7. Palmieri, F.A.N.: A Comparison of Algorithms for Learning Hidden Variables in Normal Graphs. ArXiv e-prints (2013)
8. Palmieri, F.: Notes on factor graphs. In: Apolloni, B., Bassis, S., Marinaro, M. (eds.) WIRN. Frontiers in Artificial Intelligence and Applications, vol. 193, pp. 154–162. IOS Press (2008)
9. MATLAB Documentation Center - R2014A, ch. Control and Display the Sorted Order
10. Guvenir, H.A., Acar, B., Muderrisoglu, H.: Artificial characters data set. In: Bache, K., Lichman, M. (eds.) UCI Machine Learning Repository (2013), https://archive.ics.uci.edu/ml/datasets/Artificial+Characters
11. Palmieri, F., Ciuonzo, D., Mattera, D., Romano, G., Rossi, P.S.: From examples to bayesian inference. In: Apolloni, B., Bassis, S., Esposito, A., Morabito, F.C. (eds.) WIRN. Frontiers in Artificial Intelligence and Applications, vol. 234, pp. 97–104. IOS Press (2011)

Time Series Analysis by Genetic Embedding and Neural Network Regression

Massimo Panella, Luca Liparulo, and Andrea Proietti

DIET Department, University of Rome "La Sapienza"
via Eudossiana 18, 00184 Rome, Italy
massimo.panella@uniroma1.it
http://massimopanella.site.uniroma1.it

Abstract. In this paper, the time series forecasting problem is approached by using a specific procedure to select the past samples of the sequence to be predicted, which will feed a suited function approximation model represented by a neural network. When the time series to be analysed is characterized by a chaotic behaviour, it is possible to demonstrate that such an approach can avoid an ill-posed data driven modelling problem. In fact, classical algorithms fail in the estimation of embedding parameters, especially when they are applied to real-world sequences. To this end we will adopt a genetic algorithm, by which each individual represents a possible embedding solution. We will show that the proposed technique is particularly suited when dealing with the prediction of environmental data sequences, which are often characterized by a chaotic behaviour.

Keywords: time series prediction, embedding technique, genetic algorithm, environmental data.

1 Introduction

Environmental data sequences often exhibit a chaotic behaviour that is typical for almost all real-world observed systems. In this regard, the performance of a predictor depends on how accurate it models the unknown context delivering the sequence to be predicted. Due to the actual importance of forecasting, the technical literature is full of proposed methods for implementing a predictor, especially in the field of neural and fuzzy neural networks [3], [8], [9], [11], [12].

The general approach to solve a prediction problem is based on the solution of a suitable function approximation problem, that is by synthesizing the function that links the actual sample to be predicted to a suitable set of past ones. The embedding technique is the way to determine the input vector based on past samples of a sequence $S(n)$, which can be considered as the output of an unknown autonomous system that is observable only through $S(n)$. Consequently, the sequence $S(n)$ should be embedded in order to reconstruct the state-space evolution of this system that, in actual applications, is inherently both non-linear and non-stationary. In this regard, the relationship between the reconstructed

state and its corresponding output must be a non-linear function [1]. It follows that the implementation of a predictor will coincide with the estimation of a non-linear model by using any data driven function approximation technique.

As a case study, in this paper we consider the observation of some pollution agents in Rome (Italy), whose prediction is very important in terms of health monitoring and risk prevention of daily activities. In this regard, we suggest to use a neural network approach because of its efficacy and flexibility in solving such problems. Classical neural networks (such as MultiLayer Perceptron - MLP, Radial Basis Function - RBF, Mixture of Gaussian - MoG, etc.) are function approximation models that can easily fail in the case of environmental data sequences. In fact, the complexity of the function to be approximated, caused by the chaotic behaviour, is further enhanced by the contamination of spurious noise. This inconvenience is evidently due to a lack of an accurate and complete description of data, which can be provided by means of a full conditional density $p(y|x)$ [2], [7].

In the case of the problem introduced above, the process to be estimated is often represented by a training set of P input-output pairs $x_i, y_i, i = 1 \ldots P$. Several approaches, based on a suitable clustering procedure of the training set, can be found for the synthesis of $p(y|x)$. In fact, in [10] different types of clustering approaches are proposed; one of the described approaches estimates the joint density $p(x, y)$ with no distinction between input and output variables. The joint density is successively conditioned, so that the resulting $p(y|x)$ can be used for obtaining the mapping to be approximated, i.e.:

$$p(y|x) = \frac{p(x, y)}{p(x)} = \frac{p(x, y)}{\int_{y \in \Re} p(x, y) dy} .$$

(1)

In this paper, we will refer to this approach since it ensures the largest robustness with respect to the approximation of non-convex multi-valued mappings [4]. The most popular way for obtaining $p(y|x)$ is therefore based on the prior determination of $p(x, y)$; a useful approach to the modelling of $p(x, y)$ is commonly based on a mixture of Gaussian components [13]. The determination of the said mixture yields directly the architecture of neural networks such as RBF or MoG, which are involved in this paper.

In Sect. 2 the significance of a chaotic system will be introduced. Unfortunately, the classical embedding approach, which will be briefly summarized in Sect. 3, may lead to an unsatisfactory prediction accuracy even when advanced neural network learning paradigms are used. In fact, trying to synthesize directly the unknown mapping between the current sample to be predicted and the past ones can be a difficult task that often corresponds to an ill-posed function approximation problem [6]. For these reasons, we will propose in Sect. 4 a different approach, which is based on a genetic algorithm as an advanced embedding technique. In this way, each individual in a generation represents a possible solution for the vector of past samples of $S(n)$ to be used in the approximation task. The use of a genetic algorithm allows the automatic determination of past samples without using the classical techniques for estimating the embedding parameters,

which are often characterized by a critical accuracy when applied to real-world data sequences. Moreover, the choice of the optimal parameters depends upon the use of a specific approximation model (i.e., a neural network), since the fitness of each individual is evaluated through that model fitted on the basis of the given individual (i.e., the embedded past samples).

We will consider in this work some environmental time series relevant to air pollution, whose forecasting is very important in terms of pollution control and resource management. In Sect. 5 we will discuss the chaotic nature of these sequences and we will demonstrate the suitability of the proposed technique for their prediction, as the performances in terms of accuracy are better than other well-known prediction models. The performances are evaluated by using a custom implementation of a 'Master-Slave' distributed genetic algorithm in a cluster of computers connected through the intranet of our laboratories.

2   Time Series Forecasting: Embedding for State Space Reconstruction

As previously said, a chaotic sequence $S(n)$ can be considered as the output of a chaotic system that is observable only through $S(n)$, which should be embedded in order to reconstruct the state-space evolution of this system. The general embedding technique is based on the determination of the following parameters [1]:

- embedding dimension D of the reconstructed state-space attractor, obtained by using the False Nearest Neighbors (FNN) method [14];
- time lag T between the embedded past samples of $S(n)$, obtained by using the Average Mutual Information (AMI) method; i.e.:

$$x_n = \begin{bmatrix} S(n) & S(n-T) & \dots & S(n-(D-1)T) \end{bmatrix}, \qquad (2)$$

where x_n is a row vector representing the reconstructed state at time n.

The solution of the embedding problem is useful for time series prediction. In a chaotic sequence, the prediction of $S(n)$ can be obtained by using the relationship between the (reconstructed) state and the system output. In fact, the embedding of $S(n)$ is intended to obtain an 'unfolded' version of the actual system attractor, so that the difficulty of the prediction task can be reduced. Therefore, the prediction of a chaotic sequence $S(n)$ can be considered as the determination of the function $f : \Re^D \to \Re$ that approximates the link between the reconstructed state x_n and the output sample $S(n + m)$ at the prediction distance m, being $m > 0$. Another technique can be based on the determination of the function $F : \Re^D \to \Re^D$ that approximates the link between the reconstructed state x_n and the reconstructed state x_{n+m} at the prediction distance m. Both these methods will be described in detail in the next Sect. 3.

3 Time Series Forecasting: Function Approximation Method

A chaotic system is intrinsically characterized by non-linear and non-stationary properties; consequently, its dynamic evolution should be modelled by non-linear functions determined by using data driven techniques only, especially in the case of time series prediction. In other words, the system identification and the prediction of $S(n)$ can be solved through the solution of the same function approximation problem and following two possible different approaches:

- a first approach aims at determining $F(\cdot)$, by which $\widetilde{x}_{n+m} = F(x_n)$ and the prediction is achieved by extracting the predicted sample $\widetilde{S}(n+m)$ from the estimated state $\widetilde{x}_{n+m}$. The determination of $F(\cdot)$ realizes a regularized prediction of $S(n+m)$, since the synthesis of the model $F(\cdot)$ is constrained by the simultaneous approximation of $S(n+m)$ and of the other samples embedded in x_{n+m}, i.e. $S(n+m-T)$, $S(n+m-2T)$, and so on. However, we must determine in this case a vector function $F(\cdot)$ instead of a scalar one $f(\cdot)$; by the way, this implies a greater computational cost of the learning procedure;
- a second approach will determine $f(\cdot)$, by which $\widetilde{S}(n+m) = f(x_n)$ and the identification is achieved by embedding the predicted sample $\widetilde{S}(n+m)$ in order to estimate the state $\widetilde{x}_{n+m}$. In this way, the implementation of a predictor will coincide with the determination of a non-linear data driven function approximation model. However, this approach can lead to the solution of an ill-posed problem since, even when an optimal embedding of $S(n)$ is ensured, the function approximated by $f(\cdot)$ might violate the condition of uniqueness and/or continuity [6]. The solution to this problem, suggested by several authors in the technical literature, is to adopt regularized neural network learning paradigms, as the well-known Tikhonov regularization theory [15].

To determine the sequence of reconstructed states through the approximation $F(\cdot)$ will coincide with an identification task that is necessary either when a limited number of samples of $S(n)$ are known or when the availability of these samples is delayed. However, in the following of the paper we will adopt the approach based on the estimation of $f(\cdot)$; to this end, we suggest the use of the MoG model trained by the Splitting Hierarchical Expectation Maximization (SHEM) algorithm, which will be denoted in the following as 'MoG Predictor'. In fact, it is particularly suited to the solution of multi-valued non convex function approximation problems [13].

4 Genetic Embedding for Prediction

Usually, in order to solve the embedding problem, we assume implicitly that all the past samples of $S(n)$, $n > 0$, are relevant to its solution. However, often we have no a priori information about the existence of a relationship among the past samples and the one to be predicted. In this case, a basic problem consists in

1	0	1	0	0	0	1	0	1	1	0	1	0	1	1

1		3			....						12		14	15

| $S(n\text{-}1)$ | | $S(n\text{-}3)$ | | | .... | | | | | | $S(n\text{-}12)$ | | $S(n\text{-}14)$ | $S(n\text{-}15)$ |

$$\tilde{S}(n) = f\left(\left[S(n\text{-}1)\ S(n\text{-}3)\ \dots\ S(n\text{-}12)\ S(n\text{-}14)\ S(n\text{-}15)\right]\right)$$
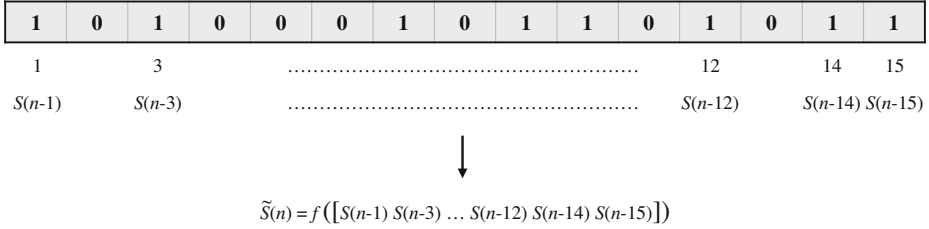
Fig. 1. Genetic encoding for fitmode1

determining how much a subset of past samples is relevant to the prediction task. The technique proposed in this paper is based on a genetic algorithm for selecting the optimal subset of past samples for the assigned prediction task. Consequently, this reduces the input space dimension and improves the prediction accuracy.

Genetic algorithms belong to the particular class of biologically inspired optimization techniques [5]; they are based on some concepts of natural selection, such as inheritance, mutation and crossover. Genetic algorithms are designed in order to manage a population of individuals, i.e. a set of potential solutions for the optimization problem at hand. Each individual is unequivocally represented by a genetic code, which is typically a string of binary digits.

The fitness of a particular individual coincides with the corresponding value assumed by the objective function to be optimized. In our application, the adopted fitness function is the prediction accuracy measured using a chosen approximation model (i.e., linear, RBF, MoG, etc.) and the subset of past samples related to the genetic code whose fitness is evaluated. In fact, once the embedding is determined, the prediction problem must be completed by the solution of a function approximation problem, that is by the determination of the function $f(\cdot)$. As aforementioned, it will be a non-linear function determined by using in general a data driven technique and, in particular, the full conditional density approach previously introduced.

For the prediction problem we have implemented two different alternatives for the genetic code:

- the genetic code is a binary string representing a subset of past samples, where the ith digit is equal to 1 if the corresponding sample is embedded in the reconstructed state and hence it feeds the approximation model, otherwise it is equal to 0. An illustrative example of this genetic code is illustrated in Fig. 1. This method will be denoted in the following as 'fitmode1';
- the genetic code is a binary string representing three subsets of bits. Each subset is the binary coding of the prediction step m and of the two embedding parameters T and D, respectively. An illustrative example of this genetic code is illustrated in Fig. 2; this method will be denoted in the following as 'fitmode2'.

A genetic algorithm produces a succession of sets of individuals (generations), aiming at increase the fitness of the best individual. The evolution starts from a

0	0	0	1	1	1	0	0	1	1

$m = 0 + 1 = 1$ $T = 3 + 1 = 4$ $D = 19 + 1 = 20$

$$\tilde{S}(n) = f\left(\left[S(n\text{-}1)\ S(n\text{-}5)\ S(n\text{-}9)\ ...\ S(n\text{-}77)\right]\right)$$
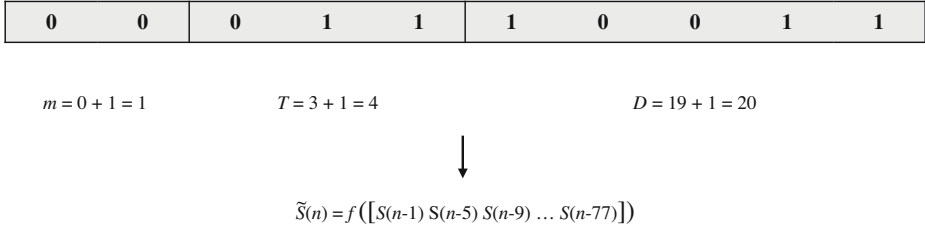
Fig. 2. Genetic encoding for fitmode2

population of completely random individuals. Starting from the kth generation G_k, the next generation G_{k+1} is determined by applying selection, mutation and crossover operators. In other words, in each generation the fitness of each individual is evaluated, multiple individuals are randomly selected from the current population (based on their fitness) and they are modified (mutated or recombined) to form the new generation.

The particular algorithm employed for our task can be summarized as follows:

1. Initialization: a population G_0 with P individuals is created and set as the current generation.
2. The individuals of G_0 are sorted by descending values of the fitness function.
3. The next generation is created by means of standard cloning, mutation and crossover operators from the current one.
4. The next generation becomes the current one.

Steps 2, 3 and 4 are iterated for a predefined fixed number M_{gen} of generations.

The behaviour of the whole algorithm depends on P and M_{gen} values, as well as on the mutation rate M_R and on the crossover rate C_R, which are two probability thresholds that control the mutation and the crossover operators. The next generation G_{k+1} is produced from the current one G_k as follows:

1. The last two individuals of G_k are deleted.
2. The best individual of G_k is cloned and put in G_{k+1} (elitism). This assures a non-decreasing behaviour of the best fitness value from a generation to the successive one.
3. The second individual of G_k is mutated with probability equal to M_R and put in G_{k+1}.
4. A pair of parents are randomly selected, with a selection probability proportional to their fitness. With a probability equal to C_R, the two parents are crossed-over. Each of the two resulting individuals is mutated with probability equal to M_R. The two resulting individuals are placed in G_{k+1}.

Step 4 is repeated until the next generation contains exactly P individuals.

5   Illustrative Tests

The forecasting performances of the proposed predictor have been carefully investigated by several simulation tests we carried out in this regard. We will

Table 1. Prediction results for the Benzene sequence (SNR in dB)

Predictor	AMI-FNN	fitmode1	fitmode2
LSE training	9.681	10.337	10.408
LSE test	9.364	10.289	10.209
RBF training	11.551	12.444	12.667
RBF test	9.194	10.749	10.488
MoG training	11.870	12.142	12.088
MoG test	8.801	10.514	10.294

Table 2. Prediction results for the PM10 sequence (SNR in dB)

Predictor	AMI-FNN	fitmode1	fitmode2
LSE training	22.184	27.043	22.509
LSE test	27.468	27.934	27.935
RBF training	22.235	22.676	22.348
RBF test	28.482	28.599	28.504
MoG training	22.234	22.465	23.649
MoG test	28.420	28.936	28.756

Table 3. Prediction results for the NO sequence (SNR in dB)

Predictor	AMI-FNN	fitmode1	fitmode2
LSE training	9.770	10.054	9.770
LSE test	9.524	9.607	9.617
RBF training	11.706	12.146	11.991
RBF test	8.786	9.450	9.216
MoG training	12.279	11.745	10.416
MoG test	8.146	9.514	9.696

illustrate in the following the results concerning actual environmental data sequences. They consist on the observation of some pollution agents in Rome (Italy): Benzene, Particulate (PM10) and Nitrogen Oxide (NO).

In order to validate the proposed prediction technique based on the genetic synthesis of the embedding vector, the prediction accuracy of the two variants fitmode1 and fitmode2 are compared with respect to a standard embedding technique, where the embedding dimension D and the time lag T are evaluated by FNN and AMI methods, respectively.

Several data driven modelling techniques have been taken into consideration: a linear predictor determined by the well-known least-squares (LSE) technique; an RBF neural network; an MoG neural network. All the predictors are trained on the first 2000 samples of $S(n)$. The same set of samples is used to compute the embedding dimension D and the time lag T by the AMI and FNN methods in the classical embedding technique. The performance of the resulting predictors, in terms of prediction accuracy, is tested on the successive 1000 samples of the sequence. It is measured by the signal-to-noise ratio (SNR), which is a commonly adopted normalised measure of the prediction accuracy where the energy of the original sequence is normalised with respect to the mean squared prediction error. Thus, the higher is the SNR the better is the prediction accuracy.

The genetic algorithm has been implemented in a Master-Slave configuration, using a client for driving the genetic evolution and a cluster of multi-core workstations. The parameters of the genetic process are $P = 100$, $M_{gen} = 30$, $M_R = 0.3$, $C_R = 1$, Roulette Wheel selection algorithm and two-point crossover.

We illustrate in Tables 1-3 the results obtained using the considered prediction models. For each row, we report the performance on both training and test sets. Considering the results of the test set, we obtain that the proposed genetic methods always outperform the classic embedding technique. Nevertheless, the fitmode1 method is better than fitmode2, since it relaxes the constraints due to Takens' theorem for what concerning the embedding parameters T and D. In fact, in the case of fitmode1 the choice of past samples does not consider a fixed time lag between them; past samples are picked up according to the genetic code associated with the best individual at the end of the genetic optimization routine.

6   Conclusions

In this paper, we considered the forecasting of three different time series related to the problem of pollution control. It is well-known that these sequences exhibit a chaotic behaviour, which is also contaminated by noise. For this reason neural networks are particularly suited to solve the forecasting problem, due to the possible robustness of their learning algorithms. This is confirmed by the performances obtained by the MoG predictor, which overcomes other prediction systems well-known in the technical literature as, for instance, the RBF neural network.

The proposed prediction approach relies on the selection of past samples to be used for prediction on the basis of a genetic algorithm optimization as an alternative approach with respect to standard embedding techniques. As evidenced by the results illustrated in this paper, the performances assured by the proposed genetic selection show an increase of prediction accuracy with respect to the commonly adopted method based on the AMI and FNN techniques.

References

1. Abarbanel, H.: Analysis of Observed Chaotic Data. Springer, New York (1996)
2. Bishop, C.: Neural Networks for Pattern Recognition. Oxford Univ. Press Inc., N.Y (1995)
3. Chen, C.H., Hong, T.P., Tseng, V.S.: Fuzzy data mining for time-series data. Applied Soft Computing 12(1), 536–542 (2012)
4. Ghahramani, Z.: Solving inverse problems using an em approach to density estimation. In: Proceedings of the 1993 Connectionist Models Summer School. Erlbaum Ass., Hillsdale (1994)
5. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning, 1st edn. Addison-Wesley Longman Publishing Co., Inc., Boston (1989)
6. Haykin, S., Principe, J.: Making sense of a complex world. IEEE Signal Processing Magazine, 66–81 (1998)
7. Huang, C.F.: A hybrid stock selection model using genetic algorithms and support vector regression. Applied Soft Computing 12(2), 807–818 (2012)
8. Khashei, M., Bijari, M.: A new class of hybrid models for time series forecasting. Expert Systems with Applications 39(4), 4344–4357 (2012)
9. Masulli, F., Studer, L.: Time series forecasting and neural networks. In: Invited tutorial in Proc. of IJCNN 1999, Washington D.C., U.S.A. (1999)
10. Panella, M.: Advances in biological time series prediction by neural networks. Biomedical Signal Processing and Control 6(2), 112–120 (2011)
11. Panella, M., Barcellona, F., D'Ecclesia, R.: Forecasting energy commodity prices using neural networks. Advances in Decision Sciences 2012, 1–26 (2012)
12. Panella, M., Liparulo, L., Barcellona, F., D'Ecclesia, R.: A study on crude oil prices modeled by neurofuzzy networks. In: Proceedings of FUZZ-IEEE 2013, Hyderabad, India (2013)
13. Panella, M., Rizzi, A., Martinelli, G.: Refining accuracy of environmental data prediction by MoG neural networks. Neurocomputing 55(3-4), 521–549 (2003)
14. Rhodes, C., Morari, M.: The false nearest neighbors algorithm: An overview. Computers & Chemical Engineering 21(suppl.), S1149 – S1154 (1997),
http://www.sciencedirect.com/science/article/pii/S0098135497876570,
supplement to Computers and Chemical Engineering 6th International Symposium on Process Systems Engineering and 30th European Symposium on Computer Aided Process Engineering
15. Tikhonov, A., Arsenin, V.: Solutions of Ill-posed Problems. W.H. Winston Ed. (1977)

Significance-Based Pruning for Reservoir's Neurons in Echo State Networks

Simone Scardapane, Danilo Comminiello, Michele Scarpiniti, and Aurelio Uncini

Department of Information Engineering, Electronics and Telecommunications (DIET),
"Sapienza" University of Rome, via Eudossiana 18, 00184, Rome
{simone.scardapane,danilo.comminiello,
michele.scarpiniti}@uniroma1.it,aurel@ieee.org

Abstract. Echo State Networks (ESNs) are a family of Recurrent Neural Networks (RNNs), that can be trained efficiently and robustly. Their main characteristic is the partitioning of the recurrent part of the network, the *reservoir*, from the non-recurrent part, the latter being the only component which is explicitly trained. To ensure good generalization capabilities, the reservoir is generally built from a large number of neurons, whose connectivity should be designed in a sparse pattern. Recently, we proposed an unsupervised online criterion for performing this sparsification process, based on the idea of *significance* of a synapse, i.e., an approximate measure of its importance in the network. In this paper, we extend our criterion to the direct pruning of neurons inside the reservoir, by defining the significance of a neuron in terms of the significance of its neighboring synapses. Our experimental validation shows that, by combining pruning of neurons and synapses, we are able to obtain an optimally sparse ESN in an efficient way. In addition, we briefly investigate the resulting reservoir's topologies deriving from the application of our procedure.

Keywords: Echo State Networks, Recurrent Neural Networks, Pruning, Least-Square.

1 Introduction

In the machine learning community, *Recurrent Neural Networks* (RNNS) have always attracted a large interest, due to their dynamic behavior [2]. In fact, a RNN implemented in a digital computer can be shown to be as least as powerful as a Turing machine [6]. Hence, in principle it can perform any computation the digital computer can be programmed to. However, the same dynamic behavior has always made RNN training difficult and subject to a large number of theoretical and numerical drawbacks [2].

Over the last two decades, different researchers independently proposed three similar models that later converged in the field of *Reservoir Computing* (RC) [3]. An RC model is a RNN architecture whose processing is partitioned in two components. First, a recurrent network, called *reservoir*, is used to process the input and extract a large number of dynamic features. Then, a static network, called *readout*, is trained on top of these features. In this way, the overall training problem is itself partitioned in two easier subproblems. In particular, in *Echo State Networks* (ESNs), the reservoir is generally

built with random connections starting from a set of classical analog neurons, while the readout is trained using linear regression techniques [3]. In this way, the original non-linear optimization problem is transformed into a simpler least-square problem, whose solution can be computed efficiently using any linear algebra package.

According to ESN theory, a reservoir has to fulfill three main properties. First, it must be *stable*, in the sense that the effect of any input should vanish after a suitable time. More formally, the reservoir must possess the so-called *echo state property*, which is generally expressed in terms of the spectral radius of its weight matrix [3]. Secondly, the reservoir should be large enough so as to ensure sufficient generalization capabilities. Finally, the connections inside the reservoir (the *synapses*) should be constructed in a sparse fashion, to ensure that the resulting features are suitably heterogeneous. A large amount of research has gone into investigating the echo state property and the optimal sizing of the reservoir [3], while the problem of *sparsification* of the synapses is less explored. Practically, the only criterion in widespread use is to randomly generate only a predefined fraction $d \in [0, 1]$ of connections during the initialization of the reservoir. However, the difficulty of choosing an optimal value for d, together with the complete stochasticity of the process, does not lead in general to a significant improvement, which probably explains the large body of works considering fully-connected reservoirs, e.g. [1].

To improve over this, in [5] we introduced an online criterion for generating sparse reservoirs in an unsupervised fashion. The main idea, which is highly inspired to the classical concepts of *Hebbian learning*, is that each synapse has a relative importance in the learning process, which can be approximated well enough by computing an estimate of the linear correlation between its input and output neuron's states. We call this quantity the *significance* of the synapse. Updating the significance at every iteration for all the synapses requires a single outer product between two vectors, hence it does not increase the computational complexity of updating the whole ESN. At fixed intervals, this quantity is used to compute a probability that each synapse is pruned, using a strategy reminiscent of the simulated annealing optimization algorithm [5]. The experimental validation in [5] shows that this procedure is robust to a change of parameters, hence it does not require a complex fine-tuning. Moreover, it provides a significant increase in performance in some situations, which is robust to an increase in the level of memory and non-linearity requested by the task.

One of the questions that remained unanswered in [5] was whether the procedure can be extended directly to the pruning of neurons. This would provide similar advantages with respect to the pruning of synapses, although with one additional benefit, namely, that the reservoir's size itself would adapt during the learning process. Hence, it can potentially free the ESN's designer from choosing an optimal reservoir's size beforehand. In this paper, we answer this question by providing an extension of the concept of significance to the neurons themselves. In particular, we define the significance of a neuron in terms of a weighted average of its neighboring incoming and outgoing connections. Then, a neuron's probability of being deleted is computed in a similar way with what has been said before. We validate our approach by employing the extended polynomial introduced in [1]. Our experiments show that, by combining pruning of neurons
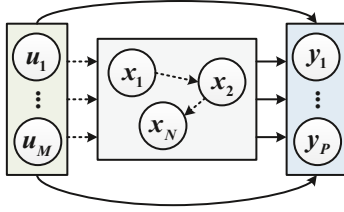
Fig. 1. General schema of an ESN, with no back-connections from the output layer. Fixed and trainable connections are represented with dashed and solid lines respectively.

and synapses, we are able to obtain an optimally sparse reservoir, with a concomitant increase in performance.

The rest of the paper is organized as follows. In Section 2 we briefly introduce the basics of ESN theory. Then, we detail our significance-based pruning for synapses in Section 3. The main novelty of this paper, its extension to the neurons of the reservoir, is in Section 4. Finally, we validate our approach in Section 5. To further investigate our procedure, we analyze the resulting reservoir's topologies in Section 5.3. We conclude with some final remarks in Section 6.

2   Echo State Networks

The ESN used in this work is represented schematically in Fig. 1. It is composed of an input layer of size M, a reservoir of size N, and an output layer of size P. For simplicity, in the following we consider $P = 1$, although everything we say extends naturally to the multi-output case. The connections going from the input layer to the reservoir, and the connections inside the reservoir, are randomly generated at the beginning of the training process. In particular, they are extracted from a normal distribution with unitary variance, in the form of an $N \times M$ matrix $\mathbf{W}_i^r$, and an $N \times N$ matrix $\mathbf{W}_r^r$, respectively. To ensure stability, the latter is then rescaled so as to achieve a predefined spectral radius ρ^* [3]. This is obtained as follows: denoting by ρ the spectral radius of $\mathbf{W}_r^r$, we rescale the original matrix by a factor ρ^*/ρ.

We suppose the network is fed with an input sequence of length S, given by $\{\mathbf{u}(1), \ldots, \mathbf{u}(S)\}$. For example, the i-th input $\mathbf{u}(i)$ may represent a sample of an audio signal, or an element of a spatial sequence. Denoting as $\mathbf{x}(n)$ the N-dimensional vector containing the states of the reservoir's neurons, we update it at every time instant as:

$$\mathbf{x}(n) = f(\mathbf{W}_r^r \mathbf{x}(n-1) + \mathbf{W}_i^r \mathbf{u}(n)) \tag{1}$$

where $f(\cdot)$ is the activation function of the neurons inside the reservoir. In this work, we use $f(\cdot) = \tanh(\cdot)$. The output of the network is computed similarly:

$$y(n) = \mathbf{w}_r^o \mathbf{x}(n) + \mathbf{w}_i^o \mathbf{u}(n) \tag{2}$$

where $\mathbf{w}_r^o$ is the N-dimensional vector linking the reservoir to the output and $\mathbf{w}_i^o$ the M-dimensional vector connecting the input to the output layer. In this work, we consider

the case of linear output, so there is no activation function in Eq. (2). In Fig. 1 we represented fixed connection with solid lines, whilst the trainable connections are shown with dashed lines. In particular, we are interested in learning the vectors $\mathbf{w}_r^o$ and $\mathbf{w}_i^o$. Denote by $\mathbf{d} = [d(1),\ldots,d(S)]$ the concatenation of all the desired outputs, by $\mathbf{s}(n)$ the "extended" state $\mathbf{s}(n) = \left[\mathbf{u}(n)^T \mathbf{x}(n)^T\right]^T$ and by $\mathbf{A}$ the concatenation of all such states $\mathbf{A} = [\mathbf{s}(1),\ldots,\mathbf{s}(S)]$. The optimal weights are given by solving the following regularized least-square optimization problem:

$$\min_{\mathbf{w}\in\mathbb{R}^{P+N\times 1}} \|\mathbf{d} - \mathbf{w}\mathbf{A}\|^2 + \lambda\|\mathbf{w}\|^2 \tag{3}$$

where $\lambda \in \mathbb{R}^+$ is a positive scalar balancing the two terms, and $\|\cdot\|$ denotes the L_2-norm of a vector. Provided we have enough samples, the solution to (3) is given by:

$$\mathbf{w} = \left(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I}\right)^{-1}\mathbf{A}^T\mathbf{d} \tag{4}$$

where $\mathbf{I}$ is the identity matrix. Practically, the initial values produced by the network are discarded due to their transient state, and are denoted as *dropout* elements. Moreover, multiple sequences in input (e.g., multiple audio signals) are handled by concatenating the resulting matrices.

3 Significance-Based Pruning for the Reservoir's Connections

In this section, we describe our pruning strategy for the reservoir's connections that we introduced in [5]. The strategy acts during the computation of the network states, and it does not require the evaluation of the error gradient. Loosely speaking, it can be seen as an hard thresholded version of the Hebbian rule, and it is inspired to some biologically existing processes in the brain [5]. The main idea is to consider the importance of a synapse (or the *significance*, as we denote it) in terms of the correlation between its input and output neurons. Practically, we define the significance of a synapse at time instant n as:

$$s_{ij}(n) = \frac{1}{T}\sum_{z=n-T}^{n} \frac{(x_i(z-1) - \hat{\mu}_x)(x_j(z) - \hat{\mu}_x)}{\hat{\sigma}_x^2} \tag{5}$$

where T is a time-interval chosen *a priori*, and $\hat{\mu}_x$ and $\hat{\sigma}_x$ are the empirical estimations of the mean and standard deviation of the neuron states. For simplicity, we suppose these are equivalent for all the neurons. The quantities $s_{ij}(n)$ are used to define the probability that a synapse is removed as:

$$p_{ij}(n) = \exp\left\{-\frac{|s_{ij}(n)|}{t(n)}\right\} \tag{6}$$

where $t(\cdot)$ is a positive, monotonically decreasing function of n. This is used to ensure that the probability of removing a synapse is maximal in the beginning of learning and goes to 0 afterwards. This is inspired from the Simulated Annealing optimization algorithm [5], and for this reason we adopt the corresponding terminology and call $t(n)$ the *temperature* of the system. Successively, every Q time instants, we prune each

---

Algorithm 1. Pseudo-code for a single update step of the pruned ESN, with direct pruning of synapses and neurons.

Data: Input signal $\mathbf{x}(n)$, desired output $d(n)$.

1   $\mathbf{x}(n) = f(\mathbf{W}_r^r \mathbf{x}(n-1) + \mathbf{W}_i^r \mathbf{u}(n))$

2   Update sums in Eqs. (5) and (8)

3   **if**   mod $(n, Q) = 0$ **then**

4      $p_{ij}(n) = \exp\left\{-\frac{|s_{ij}(n)|}{t(n)}\right\}$

5      $p_i(n) = \exp\left\{-\frac{|s_i(n)|}{\hat{t}(n)}\right\}$

6   Prune each synapse with probability $p_{ij}(n)$

7   Prune each neuron with probability $p_i(n)$

8   **end**

---

synapse with a probability given exactly by Eq. (6). We use an exponential profile for the temperature $t(n)$:

$$t(n) = \alpha^{(n/Q)-1} t_0 \tag{7}$$

where t_0 is the initial temperature, given *a priori*, and α is called the *scaling factor*. It can be seen from Eq. (7) that the temperature is scaled by a factor α at every "pruning step", given by $(n/Q) - 1$.

4   Extending Pruning to the Reservoir's Neurons

The pruning strategy introduced in the last section can be used to delete unnecessary connections inside the reservoir, hence promoting sparsity. In this section, we show how it can be extended to the direct pruning of neurons. In particular, denote as $\mathscr{I}_j(n)$ the set of *incoming* synapses of the j-th neuron at time-instant n, and by $\mathscr{O}_j(n)$ the set of *outgoing* synapses. The significance of the neuron is defined as:

$$s_j(n) = \frac{1}{2|\mathscr{I}_j(n)|} \sum_{z \in \mathscr{I}_j(n)} s_{jz}(n) + \frac{1}{2|\mathscr{O}_j(n)|} \sum_{z \in \mathscr{O}_j(n)} s_{zj}(n) \tag{8}$$

where $|\cdot|$ denotes the cardinality of the set. Thus, the significance of the neuron is defined as a weighted average of the significance of its neighboring synapses. In this way, neurons belonging to less "significant" clusters, i.e. whose connections are not significant in the sense of Eq. (5), will be denoted by a small value of Eq. (8). We can use this quantity to prune neurons in a similar way with respect to the last section. In particular, every Q time instants we define the probability of removing a given node as:

$$p_i(n) = \exp\left\{-\frac{|s_i(n)|}{\hat{t}(n)}\right\} \tag{9}$$

The new temperature $\hat{t}(n)$ must respect the same properties depicted in the last section. Practically, in all our experiments we use the exponential profile defined by Eq. (7). The overall algorithm, inclusive of pruning of the neurons and of the synapses, is summarized in Algorithm 1.

5   Experimental Validation

5.1   Experimental Setup

To test the efficacy of our strategy, we consider the extended polynomial detailed in [1], which we already adopted previously in [5]. The input to the system consists in one random number extracted from a uniform distribution over $[-1, +1]$. The output is given by:

$$y(n) = \sum_{i=0}^{p} \sum_{j=0}^{p-i} c_{ij} u^i(n) u^j(n-d) \tag{10}$$

where the coefficients c_{ij} are randomly distributed over the same distribution as the input data, and the two parameters p and d control the requirements of the task in term of memory and non-linearity. In particular, increasing p increases the power of the polynomial, while increasing d extends its delay. We consider a reservoir with $N = 250$ neurons in the reservoir. The input-to-reservoir matrix is initialized with full connectivity and weights extracted uniformly from the set $\{-0.1, 0.1\}$. The reservoir weights are extracted from a normal Gaussian distribution, and then $\mathbf{W}_r^r$ is rescaled to have a spectral radius of 0.9, following the results of [1]. We generate 20 sequences of 1000 elements each according to Eq. (10), with the addition of Gaussian noise with variance 0.01. To compute performance, we perform a 10-fold cross validation over the sequences, i.e., at every fold we use 18 sequences for training and the 2 remaining sequences for testing. We prune the network every 100 time instants, and we set $T = Q = 100$. The optimal regularization factor λ in Eq. (4) is found to be around 0.001.

Regarding our strategy, we compare the performance (i) without pruning, (ii) with pruning of the synapses, (iii) with pruning of the neurons, and (iv) with pruning of both simultaneously. In all our experiments, after some trials we set the initial temperature to 0.3. By performing an inner 3-fold cross validation for the scaling factor, we found that the optimal values are around 0.9 for the pruning of the synapses, and 0.5 for the pruning of the neurons.

5.2   Generalization Performance

We perform experiments when increasing simultaneously the power and the delay of the polynomial from 1 to 9. The *Mean-Squared Error* (MSE) averaged over the 10 folds is shown in Fig. 2. We see that the sparse versions of the reservoir have a significant decrease in testing error, which becomes more pronounced for high levels of memory and non-linearity. Moreover, the same testing error is achieved when pruning only synapses, only neurons, or both. As an example, for $p = d = 9$, we have that the sparse ESNs obtain an MSE of 0.53, compared to the original MSE of 0.87, with a 30% decrease approximately.

The original ESN is built from 250 neurons in the reservoir, and full connectivity, for a total of 62500 connections. The ESN with synapse pruning has instead an average number of 7700 connections, i.e., slightly more than 12% of the original one. Similarly, the ESN with pruning of the neuron has a final number of neurons around
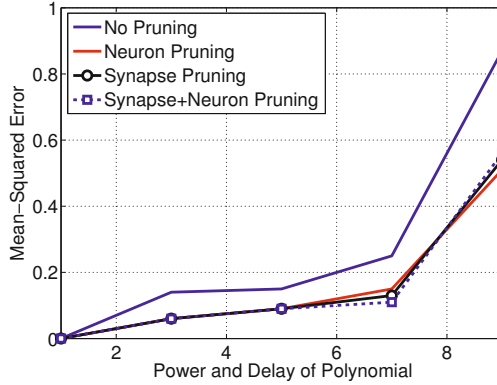
Fig. 2. Average MSE in the four cases under consideration, when increasing simultaneously the delay and power of the polynomial from 1 to 9

Table 1. Resulting number of neurons and synapses when using the different pruning strategies

Case	Neurons	Synapses
Original	250	62500
Synapse pruning	250	7700
Neuron pruning	110	12000
Full pruning	50	1700

110, with approximately 12000 connections. This means that, together with an increase in performance, the ESN is also faster to train, and easier to eventually implement on an hardware platform. The largest advantage, however, is obtained when both pruning strategies are applied simultaneously. In this case, the final ESN has an approximate number of 50 neurons, with 1700 connections left. This is summarized in Table 1.

5.3   Analysis of the Reservoirs Topology

Before concluding, we briefly investigate an interesting aspect of our strategies, namely, the resulting topologies of the reservoir after pruning. In Fig. 3 we plotted the histograms of the average number of outgoing connections, the so-called *outdegree* [4], inside the reservoirs, after applying the three criteria. We see from Fig. 3-(a) that, when pruning only the synapses, most of the resulting neurons have a relatively small outdegree (ranging in 10-30), and the outdegree decays linearly. The distribution when pruning the neurons is slightly more complex, and is depicted in Fig. 3-(b). We can see that the outdegree has a set of three peaks which are evenly distributed, then decays linearly in both verses. The most interesting aspect, however, is relative to the distribution when pruning both neurons and synapses simultaneously, depicted in Fig. 3-(c). We see that most neurons have a very small number of outgoing connections $(10 - 20)$, and a

(a) Synapse pruning   (b) Neuron pruning   (c) Simultaneous pruning of neurons and synapses
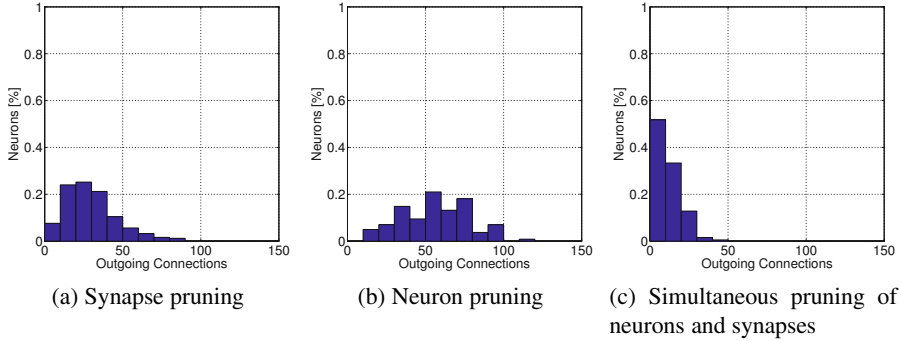
Fig. 3. Histogram of the average number of outgoing connections in the ESN, after applying the three pruning strategies

smaller fraction an outdegree of 30. Hence, it seems that the overall strategy is indeed creating compact "clusters" of neurons. This is an interesting behavior that we are eager to investigate more deeply in a future work.

6   Conclusions

Echo State Networks allows for an efficient training of Recurrent Neural Networks in real-world applications. Due to their nature, however, they generally require large networks, which may be non-applicable in realistic contexts. In this paper, we have extended an algorithm that we proposed for the pruning of the synapses, to the direct pruning of neurons. Our results show that, when the two strategies are applied simultaneously, we are able to obtain optimally sparse reservoir without increasing the computational complexity of the training process.

References

1. Butcher, J., Verstraeten, D., Schrauwen, B., Day, C., Haycock, P.: Reservoir computing and extreme learning machines for non-linear time-series data analysis. Neural Networks 38, 76–89 (2013)
2. Jaeger, H.: The "echo state" approach to analysing and training recurrent neural networks - with an erratum note. Tech. rep. (2001)
3. Lukoševičius, M., Jaeger, H.: Reservoir computing approaches to recurrent neural network training. Computer Science Review 3(3), 127–149 (2009)
4. Newman, M.: Networks: an introduction. Oxford University Press (2010)
5. Scardapane, S., Nocco, G., Comminiello, D., Scarpiniti, M., Uncini, A.: An effective criterion for pruning reservoir's connections in echo state networks. In: 2014 International Joint Conference in Neural Networks, pp. 1205–1212 (2014)
6. Siegelmann, H.T.: Neural and super-turing computing. Minds and Machines 13(1), 103–114 (2003)

Online Selection of Functional Links
for Nonlinear System Identification

Danilo Comminiello, Simone Scardapane, Michele Scarpiniti,
Raffaele Parisi, and Aurelio Uncini

Department of Information Engineering, Electronics and Telecommunications (DIET)
"Sapienza" University of Rome
via Eudossiana 18, 00184 Rome, Italy
{danilo.comminiello,simone.scardapane,michele.scarpiniti,
raffaele.parisi}@uniroma1.it,aurel@ieee.org

Abstract. This paper introduces a new method for improving nonlinear modeling performance in online learning by using functional link-based models. The proposed algorithm is capable of selecting the useful nonlinear elements resulting from the functional expansion, while setting to zero the ones that does not bring any improvement of the modeling performance. This allows to reduce any gradient noise due to a possible overestimate of the solution, thus preventing any overfitting phenomena. The proposed model is assessed in several nonlinear identification problems, including different levels of nonlinearity, showing significant improvements.

Keywords: Nonlinear Modeling, Functional Links, Nonlinear Transformation, Nonlinear System Identification, Sparse Systems.

1 Introduction

The nonlinearity degree in a signal may depend on several factors related to the signal itself, such as its nonstationary or time-varying nature. Therefore, in nonlinear system identification problems, it becomes very difficult to design *a priori* a nonlinear model to be used without incurring in any overfitting issue. All along the years, in offline learning problems, pruning methods have been widely applied to batch and sequential models, due to the modeling performance improvement that they produce [8,9,12,11,16]. However, in online learning problems, these methods may not be appropriate, sometimes due to expensive computational load, or even due to batch processing. In this case, other methods can be adopted that do not actually prune unnecessary elements, but just perform an online selection of the useful elements.

In this paper, we propose a new method for improving nonlinear modeling performance in online learning, which performs an online selection of the nonlinear elements, thus reducing any gradient noise that may be generated by a possible overestimate of the solution. We focus on a class of linear-in-the-nonlinear adaptive models [18], which are based on a nonlinear transformation of the input signal that projects it in a higher dimensional space. Then, the transformed

signal can be processed by a linear model. In particular, we take into account the nonlinear *functional link adaptive filters* (FLAFs) [3], in which the nonlinear expansion is carried out by the so-called *functional links* [14,15,19,1], and the subsequent linear model is an adaptive filter.

One of the main advantages of the FLAF model lies in its flexibility, since the setting of several parameters is allowed in order to fit the model to a specific application. In this regard, an important choice in the FLAF design concerns the amount of functional links to be employed for the modeling. This choice is strictly related to the nonlinearity degree introduced by the unknown system. However, when the number of functional links is too high with respect to the nonlinearity to be modeled, it may lead to overfitting phenomena that may cause a decrease of the modeling performance [4]. This is due to the fact that only a portion ,i.e. a sparse representation, of functional links actively contributes to the filtering.

In order to address this problem, we exploit a sparse representation of the functional links [4] to perform an online selection and thus improve the performance. This approach is based on the proportionate adaptive algorithms [7,10,13]. Proportionate algorithms were developed to improve the convergence performance in linear systems when the impulse response to be estimated shows a sparse nature, i.e., many of its coefficients are zero or very close to zero. However, sparsity is not only related to linear systems, but it can also occur in the estimation of nonlinearities, as not all the elements of a nonlinear model may be useful for a correct modeling. Unlike [4], where a *split FLAF* architecture was proposed for nonlinear acoustic echo cancellation, in this paper, we focus on a *μ-law* rule [6] to exploit a sparse functional link representation for nonlinear FLAF. The resulting *μ-law proportionate FLAF* algorithm gives a greater importance to the coefficients that contribute actively to the nonlinear modeling. At the same time, any overfitting phenomenon caused by the unnecessary coefficients is avoided. The effectiveness of the the proposed method is assessed in the nonlinear system identification problems, which requires online processing.

The paper is organized as follows: the nonlinear FLAF model is introduced in Section 2 and the proposed algorithm is detailed in Section 3. Results are discussed in Section 4 and, finally, in Section 5 our conclusions are drawn.

2   A Brief Review on the Nonlinear FLAF

The FLAF model is based on the representation of the input signal in a higher-dimensional space [14], in which an enhanced nonlinear modeling is allowed. Such approach derives from the machine learning theory, more precisely from the Cover's Theorem on the separability of patterns (see for example [8]).

The purely nonlinear FLAF is composed of two main parts: a nonlinear *functional expansion block* (FEB) and a subsequent linear adaptive filter, as depicted in Fig. 1. The FEB consists of a series of functions, which might be a subset of a complete set of orthonormal basis functions satisfying universal approximation constraints. The term "functional links" actually refers to the functions
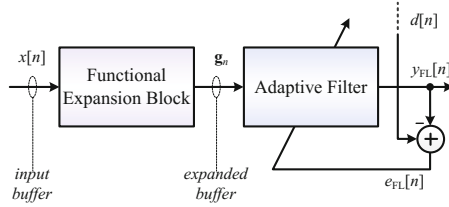
Fig. 1. The nonlinear functional link adaptive filter

contained in the chosen set $\Phi = \{\varphi_0(\cdot), \varphi_1(\cdot), \ldots, \varphi_{Q-1}(\cdot)\}$, where Q is the number of functional links. At the n-th time instant, the FEB receives the input sample $x[n]$, which is stored in an input buffer $\mathbf{x}_{\mathrm{N},n} \in \mathbb{R}^{M_{\mathrm{i}}} = [x[n] \, x[n-1] \ldots x[n-M_{\mathrm{i}}+1]]^T$, where M_{i} is defined as the input buffer length. Each element of $\mathbf{x}_{\mathrm{N},n}$ is passed as argument to the chosen set of functions Φ, thus yielding a subvector $\overline{\mathbf{g}}_{i,n} \in \mathbb{R}^Q$:

$$\overline{\mathbf{g}}_{i,n} = \left[\varphi_0(x[n-i]) \, \varphi_1(x[n-i]) \ldots \varphi_{Q-1}(x[n-i]) \right]. \tag{1}$$

The concatenation of all the subvectors, for $i = 0, \ldots, M_{\mathrm{i}} - 1$, engenders an *expanded buffer* $\mathbf{g}_n \in \mathbb{R}^{M_{\mathrm{e}}}$:

$$\mathbf{g}_n = \left[\overline{\mathbf{g}}_{0,n}^T \, \overline{\mathbf{g}}_{1,n}^T \cdots \overline{\mathbf{g}}_{M_{\mathrm{i}}-1,n}^T \right]^T = \left[g_0[n] \, g_1[n] \ldots g_{M_{\mathrm{e}}-1}[n] \right]^T \tag{2}$$

where $M_{\mathrm{e}} \geq M_{\mathrm{i}}$ represents the length of the expanded buffer. Note that $M_{\mathrm{e}} = M_{\mathrm{i}}$ only when $Q = 1$. As functional expansion, we choose a nonlinear trigonometric series expansion:

$$\varphi_j(x[n-i]) = \begin{cases} \sin(p\pi x[n-i]), \, j = 2p-2 \\ \cos(p\pi x[n-i]), \, j = 2p-1 \end{cases} \tag{3}$$

where $p = 1, \ldots, P$ is the expansion index, being P the *expansion order*, and $j = 0, \ldots, Q-1$ is the functional link index. The trigonometric expansion (3) implies a functional link set Φ composed of $Q = 2P$ functional links. Some convergence properties of the nonlinear FLAF of Fig. 1 can be found in [5]. It is worth noting that (3) actually refers to a *memoryless* expansion, since it does not involve cross-products, but it can be easily extended to a memory expansion (see [3] for a detailed explanation). A way of considering the memory of a nonlinearity is that of taking into account the outer products of the i-th input sample with the functional links of the previous input samples. The FLAF with memory is characterized by a *memory order K* [3].

The expanded buffer $\mathbf{g}_n$ is then fed into a linear adaptive filter $\mathbf{w}_{\mathrm{FL},n} \in \mathbb{R}^{M_{\mathrm{e}}} = \left[w_{\mathrm{FL},0}[n] \, w_{\mathrm{FL},1}[n] \ldots w_{\mathrm{FL},M_{\mathrm{e}}-1}[n] \right]^T$, thus providing the nonlinear output:

$$y_{\mathrm{FL}}[n] = \mathbf{g}_n^T \mathbf{w}_{\mathrm{FL},n-1}. \tag{4}$$

Thereby, the nonlinear error signal is:

$$e_{\text{FL}}[n] = d[n] - y_{\text{FL}}[n] \tag{5}$$

which is used for the adaptation of $\mathbf{w}_{\text{FL},n}$. In (5), $d[n]$ represents the desired signal for the nonlinear model. Being $\mathbf{w}_{\text{FL},n}$ a conventional linear filter, it can be adapted by any adaptive algorithm based on the minimization of the mean square error [20]. The use of an adaptive filter after the expansion allows to apply the FLAF model to several online learning applications, such as active noise reduction, acoustic echo cancellation [19,2,17,3].

3   The μ-law Proportionate FLAF

Very often, nonlinearities affecting an input signal may vary in time and frequency. This behavior is further stressed when the input signals has a nonstationary nature. This is the reason why not all the nonlinear elements of an expanded buffer may be useful in the same way to model a dynamic nonlinear channel. A possible solution to this drawback is that of using a weighted mask for the nonlinear filter in an attempt to give more prominence to those nonlinear elements of the expanded buffer that have an active role in the modeling of nonlinearities. To this end, we introduce a weighted adaptive algorithm for the nonlinear FLAF, that provides a sparse representation of functional links, thus performing an online selection of them. In order to exploit the sparsity in the expanded buffer, proportionate adaptive algorithms may be adopted [7,10,13]. Among such algorithms, we take into account the μ-*law proportionate normalized least mean square* (MPNLMS) algorithm, which is based on an approximation of the optimal proportionate step size [6]. According to this, the update equation of the MPNLMS-FLAF can be expressed as:

$$\mathbf{w}_{\text{FL},n} = \mathbf{w}_{\text{FL},n-1} + \eta \frac{\mathbf{Q}_n \mathbf{g}_n}{\mathbf{g}_n^T \mathbf{Q}_n \mathbf{g}_n + \delta_{\text{P}}} e_{\text{FL}}[n] \tag{6}$$

where η is the step-size parameter and δ_{rP} is a regularization factor. $\mathbf{Q}_n$ is a diagonal weighting matrix that contains the proportionate coefficients $q_m[n]$, with $m = 0, \ldots, M_e - 1$, whose values are computed according to [6]:

$$q_m[n] = \frac{\gamma_m[n]}{\frac{1}{M_e} \sum_{i=0}^{M_e - 1} \gamma_i[n]}, \qquad m = 0, \ldots, M_e - 1. \tag{7}$$

The coefficients $\gamma_m[n]$ can be computed by introducing a function of the estimate of the optimal filter coefficient:

$$\theta_m[n] = \frac{\ln\left(1 + \mu \left|w_{\text{FL},m}[n-1]\right|\right)}{\ln\left(1 + \mu\right)}, \qquad m = 0, \ldots, M_e - 1 \tag{8}$$

where the step size can be represented as $\mu = 1/\varepsilon$. The parameter ε is a very small positive number and its value can be chosen according to the measurement

noise. As a default choice, we set $\varepsilon = 0.001$ (i.e., $\mu = 1000$) that means that the noise below -60 dB is negligible. In (8), the constant 1 inside the logarithm has been introduced in order to avoid a singular point when $|w_{\text{FL},m}[n]| = 0$. Moreover, the denominator normalizes the function to be in the range $[0, 1]$. It is worth noting that the function $\theta_m[n]$ is nothing but the μ-law used in nonuniform compression in telecommunication applications [6].

Based on (8), it is possible to define, first, a lower bound for the coefficients $\gamma_m[n]$:

$$\gamma_{\min}[n] = \rho \max \{\xi, \theta_0[n], \ldots, \theta_{M_e-1}[n]\} \tag{9}$$

where ρ is a scaling factor and ξ is a threshold value, usually chosen respectively as $\rho = 0.01$ and $\xi = 0.01$. Then, using (8) and (9), it is possible to define the coefficients $\gamma_m[n]$:

$$\gamma_m[n] = \max \{\gamma_{\min}[n], \theta_m[n]\} \tag{10}$$

that can be finally used to derive the proportionate coefficients in (7) and, thus, to achieve the update equation (6) for the MPNLMS-FLAF.

4   Experimental Results

In this section, we evaluate the nonlinear modeling performance of the proposed MPNLMS-FLAF. We assess the effectiveness of the MPNLMS-FLAF over three different system identification scenarios, which are distinguished according to the nonlinearity degree introduced by an unknown system. In all the scenarios, the plant to be identified is composed of a nonlinear system followed by a linear system, in a Hammerstein configuration as depicted in Fig. 2. The input signal is generated by a first-order autoregressive model, whose transfer function is $\sqrt{1 - \theta^2}/(1 - \theta z^{-1})$, with $\theta = 0.8$, fed with an independent identically distributed (i.i.d.) Gaussian random process. The length of the input signal is $L = 20000$ samples. In each scenario, the linear system is formed by $M = 7$ independent random values between -1 and 1. An additive i.i.d. noise signal $v[n]$ is added at the output of the whole plant, in order to provide 30 dB of signal-to-noise ratio (SNR). Performance is evaluated in terms of the *excess mean square error* (EMSE), in dB:

$$\text{EMSE}[n] = 10 \log_{10} \left(\text{E} \left\{ (e[n] - v[n])^2 \right\} \right) \tag{11}$$

which is averaged over 1000 runs with respect to input and noise. Moreover, in order to facilitate the visualization, curves are smoothed by a moving-average
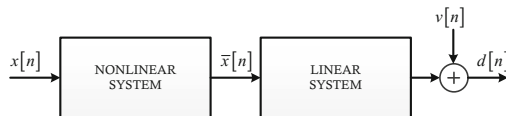
Fig. 2. General scheme for nonlinear system identification scenarios

(a)   (b)
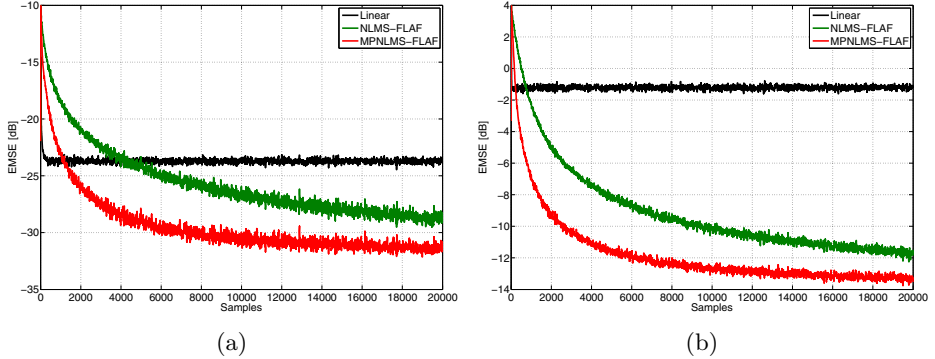
Fig. 3. Performance behavior in terms of ERLE in the presence of: (a) mild nonlinearity, and (b) strong nonlinearity

filter. In all the scenarios, the proposed MPNLMS-FLAF is compared, in terms of the EMSE, with the standard NLMS-FLAF, in which the FLAF is simply adapted by an NLMS algorithm [3], and with a simple NLMS algorithm, in order to have a linear reference.

In the first scenario, we assume that the nonlinear system applies a symmetrical soft-clipping distortion to the input signal, described by the following equation [4]:

$$\overline{x}\,[n] = \begin{cases} \frac{2}{3\zeta} x\,[n] & \text{for } 0 \leq |x\,[n]| \leq \zeta \\ \text{sign}\,(x\,[n]) \frac{3-(2-x[n]/\zeta)^2}{3} & \text{for } \zeta \leq |x\,[n]| \leq 2\zeta \\ \text{sign}\,(x\,[n]) & \text{for } 2\zeta \leq |x\,[n]| \leq 1 \end{cases} \tag{12}$$

where ζ is a threshold chosen in the range $(0\,,0.5]$. For this experiment this threshold is chosen as $\zeta = 0.15$, thus providing a medium/mild degree of nonlinearity. We normalize the input signal to limit its amplitude in the range $[-1, 1]$. The resulting signal $\overline{x}\,[n]$ is convolved with the linear impulse response, as in Fig. 2. The parameter setting of FLAF-based models for this experiment is: $\mu_{\text{FL}} = 0.2$, $\delta_{\text{FL}} = 10^{-3}$, $M_{\text{i}} = M$, $P = 15$. Memoryless FLAFs are considered for this experiments (i.e., $K = 0$). Results are shown in Fig. 3(a), in which it can be seen that the proposed MPNLMS-FLAF outperforms the NLMS-FLAF.

In the second scenario, as regards the nonlinear system, we consider the same symmetrical soft-clipping distortion of (12), but we choose a threshold $\zeta = 0.05$, which yields a high degree of nonlinearity. Differently from the previous parameter setting, in this case we use a higher expansion order, i.e., $P = 30$. Results are shown in Fig. 3(b), where it is possible to notice that the gap between the linear model and the FLAF-based ones is larger with respect to the previous scenario, since the nonlinearity introduced is stronger than before. Moreover, the proposed MPNLMS-FLAF keeps its performance gain over the NLMS-FLAF.

In the last scenario, we increase the difficulty of the problem by considering a dynamic nonlinearity, whose function involves a temporal dependence.

In particular, the nonlinear system in Fig. 2 is composed of two subsequent blocks. The first block receives the input signal $x[n]$ and applies the following dynamic nonlinearity:

$$
\begin{aligned}
\overline{u}[n] = {} & \frac{3}{2}\cos\left(\frac{\pi}{2}x[n]\right) - \frac{3}{10}\cos\left(\frac{\pi}{2}x[n]\right)^2 + \frac{9}{5}\cos\left(\frac{\pi}{2}x[n]\right)\sin\left(\frac{\pi}{4}x[n-1]\right) \\
& + \frac{1}{2}\sin\left(\frac{\pi}{2}x[n]\right)\cos\left(\frac{\pi}{8}x[n-2]\right) - \frac{2}{5}\sin\left(\frac{\pi}{2}x[n]\right)\cos\left(\frac{\pi}{16}x[n-3]\right) \\
& - \frac{3}{2}\sin\left(\frac{\pi}{4}x[n-1]\right)\cos\left(\frac{\pi}{8}x[n-2]\right) \\
& + \frac{9}{10}\sin\left(\frac{\pi}{2}x[n]\right)\cos\left(\frac{\pi}{4}x[n-1]\right)\sin\left(\frac{\pi}{16}x[n-3]\right).
\end{aligned} \tag{13}
$$

The resulting signal $\overline{u}[n]$ is then processed by a 3-rd order Chebyshev filter with transfer function:

$$
H(z) = \frac{0.6055 + 0.7785z^{-1} + 0.778z^{-2} + 0.6055z^{-3}}{1 + 0.6416z^{-1} + 0.7692z^{-2} + 0.3574z^{-3}}, \tag{14}
$$

thus yielding the nonlinear signal $\overline{u}[n]$ that is then fed into the linear system according to the sceme in Fig. 2. For this experiment, we choose the same expansion order of the previous experiment, $P = 30$, but we take into account also some memory in the functional links in order to better model the time-dependent nonlinearity. In particular, we choose a memory order $K = 5$, which represents a good compromise between performance and computational cost [3]. The other parameters are setting as in the previous experiments. Results are shown in Fig. 4, where it can be seen that, while the linear model is not able to provide any improvement, the FLAF-based models achieve good results, and, in particular, the MPNLMS-FLAF takes advantage of its parameter selection to provide a larger improvement with respect to the NLMS-FLAF. It should be noted that, although the nonlinearity adopted in the third scenario is very strong, it is composed of sine and cosines, thus facilitating the modeling by trigonometric functional links.
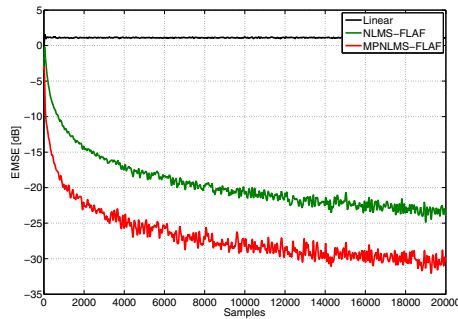
Fig. 4. Performance behavior in terms of the EMSE in the presence of strong dynamic nonlinearity

By considering the three experiments above, it is worth noting that the proposed MPNLMS-FLAF achieves an improvement as large as higher the nonlinearity degree. This is due to the fact that a strong nonlinearity needs a larger expansion buffer, whose effectiveness of the nonlinear elements is not uniform but sparse. Moreover, it can be noticed that, unlike the MPNLMS-FLAF, the NLMS-FLAF considers also the useless functional links, which generates overfitting. Therefore, we can conclude that the performance gap of the NLMS-FLAF from the proposed MPNLMS-FLAF is essentially due to the overfitting.

5   Conclusion

In this paper, a new algorithm has been proposed to perform an online selection of functional links in a FLAF. The selection of functional links serves to prevent the occurrence of overfitting phenomena when the system to be identified is unknown. The proposed MPNLMS-FLAF is based on the μ-law proportionate adaptive algorithm and exploits the sparse representation of functional links. This model has been assessed in nonlinear system identification problems. In particular, three scenarios with different nonlinearity degrees have been considered. Results have proved the effectiveness of the proposed method for all the scenarios. In future works, the adopted algorithm may be extended also to more sophisticated FLAF-based architectures for problems like the nonlinear acoustic echo cancellation. Moreover, other kinds of proportionate algorithms can be investigated to online select functional links. The method can be also extended to other classes of linear-in-the-parameters nonlinear models.

References

1. Alhamdoosh, M., Wang, D.: Fast decorrelated neural network ensembles with random weights. Inform. Sciences 264, 104–117 (2014)
2. Comminiello, D., Azpicueta-Ruiz, L.A., Scarpiniti, M., Uncini, A., Arenas-Garcia, J.: Functional link based architectures for nonlinear acoustic echo cancellation. In: Proc. IEEE J. Works. Hands-free Speech Commun. Microph. Arrays (HSCMA 2011), Edinburgh, UK, pp. 180–184 (May 2011)
3. Comminiello, D., Scarpiniti, M., Azpicueta-Ruiz, L.A., Arenas-García, J., Uncini, A.: Functional link adaptive filters for nonlinear acoustic echo cancellation. IEEE Trans. Audio, Speech, Lang. Process. 21(7), 1502–1512 (2013)
4. Comminiello, D., Scarpiniti, M., Azpicueta-Ruiz, L.A., Arenas-García, J., Uncini, A.: Nonlinear acoustic echo cancellation based on sparse functional link representations. To appear in IEEE Trans. Audio, Speech, Lang. Process. (2014), doi:10.1109/TASLP.2014.2324175
5. Comminiello, D., Scarpiniti, M., Parisi, R., Uncini, A.: Convergence properties of nonlinear functional link adaptive filters. IET Electron. Lett. 49(14), 873–875 (2013)
6. Doroslovački, M., Deng, H.: Proportionate adaptive algorithms for network echo cancellation. IEEE Trans. Signal Process. 54(5), 1794–1803 (2006)
7. Duttweiler, D.L.: Proportionate normalized least-mean-squares adaptation in echo cancelers. IEEE Trans. Speech Audio Process. 8(5), 508–518 (2000)

8. Haykin, S.: Neural Networks and Learning Machines, 3rd edn. Prentice Hall, Upper Saddle River (2008)
9. Karnin, E.: A simple procedure for pruning back-propagation trained neural networks. IEEE Trans. Neural Netw. 1(2), 239–242 (1990)
10. Naylor, P., Cui, J., Brookes, M.: Adaptive algorithms for sparse echo cancellation. Signal Process. 86(6), 1182–1192 (2006)
11. Omlin, C., Giles, C.: Pruning recurrent neural networks for improved generalization performance. IEEE Trans. Neural Netw. 5(5), 848–851 (1994)
12. Orlandi, G., Piazza, F., Uncini, A., Ascone, A.: A biological approach to plasticity in artificial neural netowrks. In: Proc. IEEE Int. J. Conf. Neural Netw (IJCNN 1991), Seattle,WA, July 8-12, vol. 2, pp. 583–586 (1991)
13. Paleologu, C., Benesty, J., Ciochină, S.: Sparse Adaptive Filters for Echo Cancellation. Morgan & Claypool Publishers (2010)
14. Pao, Y.H.: Adaptive Pattern Recognition and Neural Networks. Addison-Wesley, Reading (1989)
15. Patra, J.C., Pal, R.N., Chatterji, B.N., Panda, G.: Identification of nonlinear dynamic systems using functional link artificial neural networks. IEEE Trans. Syst., Man, Cybern., B: Cybern. 29(2), 254–262 (1999)
16. Scardapane, S., Nocco, G., Comminiello, D., Scarpiniti, M., Uncini, A.: An effective criterion for pruning reservoir's connections in echo state networks. In: Proc. IEEE WCCI - Int. J. Conf. Neural Netw. (IJCNN 2014), Beijing, China, July 6-11 (2014)
17. Scarpiniti, M., Comminiello, D., Parisi, R., Uncini, A.: A collaborative approach to adaptive noise cancellation. In: Proc. XXII Italian Works. Neural Netw. (WIRN), Vietri Sul Mare, Italy, May 17-19 (2012)
18. Sicuranza, G., Carini, A.: On a class of nonlinear filters. In: Tabus, M., Egiazarian, K. (eds.) Festschrift in Honor of Jaakko Astola on the Occasion of his 60th Birthday. TICPS, vol. 47, pp. 115–144 (2009)
19. Sicuranza, G.L., Carini, A.: A generalized FLANN filter for nonlinear active noise control. IEEE Trans. Audio, Speech, Lang. Process. 19(8), 2412–2417 (2011)
20. Uncini, A.: Fundamentals of Adaptive Signal Processing. Springer (2014) ISBN 978-3-319-02806-4

A Continuous-Time Spiking
Neural Network Paradigm

Alessandro Cristini*, Mario Salerno, and Gianluca Susi

Departement of Electronic Engineering, University of Rome "Tor Vergata",
via del Politecnico 1, 00133 Rome, Italy
alessandro.cristini84@gmail.com,
{salerno,gianluca.susi}@uniroma2.it

Abstract. In this work, a novel continuous-time spiking neural network paradigm is presented. Indeed, because of a neuron can fire at any given time, this kind of approach is necessary. For the purpose of developing a simulation tool having such a property, an ad-hoc event-driven method is implemented. A simplified neuron model is introduced with characteristics similar to the classic Leaky Integrate-and-Fire model, but including the spike latency effect. The latency takes into account that the firing of a given neuron is not instantaneous, but occurs after a continuous-time delay. Both excitatory and inhibitory neurons are considered, and simple synaptic plasticity rules are modeled. Nevetheless the chance to customize the network topology, an example with Cellular Neural Network (CNN)-like connections is presented, and some interesting global effects emerging from the simulations are reported.

Keywords: Neuron Model, Spike Latency, Spiking Neural Network, Synaptic Plasticity, Continuous-Time Paradigm, Event-Driven Simulation.

1 Introduction

In recent decades there has been a significant increase in the development, implementation and general purpose use of spiking neural networks [1,2,3]. The attractiveness of this kind of neural networks lies in the bio-inspired neuron models and the related peculiar characteristics, such as: subthreshold decay of the membrane potential, spatio-temporal integration of the incoming synaptic inputs, excitatory and inhibitory effects, threshold phenomena, spike latency, synaptic plasticity, etc. Several neuron models have been proposed in the literature, from the simplest Integrate-and-Fire [4] to the most bio-realistic Hodgkin-Huxley models [5]. However, these are typically described by means of ODEs (Ordinary Differential Equations). Usually, the more ODEs are complex, the more the neuronal membrane potential is faithfully followed. A comparative

---

* Corresponding author.

review of the main neuron models is listed in [6]. With the aim of making possible simulations of very large networks on a PC (Personal Computer), often the choice fall on the simplest model. Indeed, the latter allows the investigation of global effects arising only in the case of large networks of spiking neurons, such as: Polychronization [7,8] or the formation of spontaneous neuronal groups that can be affected by the Neuronal Group Selection [9,10]. Moreover, in some cases I&F model can be analytically studied (e.g., [11,12]).

There are two main methods to simulate such networks: clock-driven (or "synchronous") and event-driven (or "asynchronous") strategies. In the first one, all neurons are updated simultaneously at every tick of a clock, whereas in the second one, all neurons are updated only when they receive or emit a spike. The latter kind of strategies is developed for exact simulations, thus allowing a high precision computation. The obvious drawback of clock-driven methods is that spike timings are aligned to a grid (ticks of the clock), thus the simulation is approximated even when the differential equations are computed exactly. Other specific errors come from the fact that threshold conditions are checked only at the ticks of the clock, implying that some spikes might be missed. Whereas, event-driven methods implicitly assume that we can calculate the state of a neuron at any given time, i.e., we have an explicit solution of the differential equations [13]. This condition cannot be always satisfied; for instance, the Hodgkin-Huxley equations have no explicit solution.

However, in the present work we apply an event-driven method in order to implement a novel continuous-time spiking neural network paradigm. The neuron model is not described using ODEs, but it explicitly makes an algebraic sum of any incoming inputs to a target neuron. Indeed, both excitatory and inhibitory effects are considered. Each input is weighted by a synaptic strength, that may be affected by the synaptic plasticity (see [14] for an overview about the biological mechanisms), thus we have implemented simple rules (described in *synaptic plasticity rules* section) to take into account this phenomenon. Note that, in this model when a threshold is crossed, the neuron will fire with a continuous-time delay called latency [15]. Of course, the threshold crossing is often prevented by the subthreshold decay. Therefore, a spatio-temporal integration is implemented. Finally, after the spike generation, the neuron will be reset to its resting potential, becoming not excitable for a time equal to the absolute refractory period. Also, we have considered a Cellular Neural Network (CNN)-*like* topology [16], in order to arrange the connections among neurons; then, each neuron fires to a number of target neurons belonging to a fixed neighborhood. In *simulation results* section, we will show simulation results about global effects arising thanks to this kind of model and neuronal paradigm.

2   Neuron Model

We propose here a simplified neuron model in which the variables are updated step-by-step (i.e., in correspondence to incoming events). Note that, only normalized real quantities are considered. In addition, we call "firing neuron" an emitting neuron, and "burning neuron" a receiving one (see Fig. 1). Furthermore, we define "passive mode" the operating mode of the neuron when its inner state is less than a threshold, "active mode" otherwise.
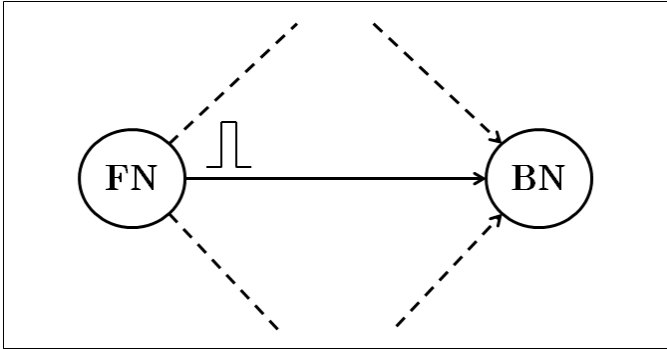
Fig. 1. A Firing Neuron (FN) emitting a pulse to a Burning Neuron (BN). The dashed lines indicate other connections linking the FN to other BNs or incoming connections from other FNs to the depicted BN. Of course, each neuron can be both FN and BN, depending on the direction of the activity.

The state of each burning neuron is evaluated through the following equations:

$$S = S_p + P_r P_w - T_l \ , \ \text{for } S < S_{th} \tag{1}$$

$$S = S_p + P_r P_w + T_r \ , \ \text{for } S \geq S_{th} \tag{2}$$

In (1)–(2), S denotes the inner state of the neuron; when $S = 0$ the neuron is in its resting state. S_p is the previous state, whereas, S_{th} represents the spiking threshold. The latter is conventionally chosen equal to $1+d$, where d indicates a threshold constant. The quantity P_r, "presynaptic weight", represents the signal emitted by a firing neuron to a number of burning neurons; this quantity is conventionally equal to the inverse of the fan-out of the firing neuron, but other choices can be taken into account. Of course, this is a simplification by which we only consider inputs with the same amplitude. For the purpose of considering the inhibitory effect, P_r is chosen negative for inhibitory neurons. The quantity P_w, "postsynaptic weight", represents the connection strength between a couple of neurons. If P_w is equal to 0, the related connection is not present.

Finally, with T_l (*leakage term*) we take into account the subthreshold decay for the passive mode ($S < S_{th}$). In particular, we have chosen a linear decay behavior (this kind of decay is used in [17]), then $T_l = Ld\Delta t$; in which Ld is

the linear parameter, whereas, Δt represents the temporal distance between a couple of consecutive input spikes.

In active mode ($S \geq S_{th}$), the neuron is ready to fire: its firing is not instantaneous, but occurs after a continuous-time delay called *time-to-fire*. This quantity can be affected by inputs, making the neuron sensitive to possible changes in the network for a time window bounded by the time-to-fire itself. The inner state and the time-to-fire are related through the following bijective relationship, called *firing equation*.

$$t_f = \frac{1}{(S-1)} \, .$$ (3)

Equation (3) represent an approximation of the curve that we have obtained through the simulation of a membrane patch stimulated by brief current pulses (0.01 ms of duration), solving the Hodgkin-Huxley equations [5] making use of NEURON simulator [18]. Similar behaviors have been investigated by other authors; for example, in Wang et al. [19], using DC inputs (see Fig. 1 in [19]).

In Fig. 2, a qualitatively comparison between the simulated behavior of the latency and the firing equation is shown.
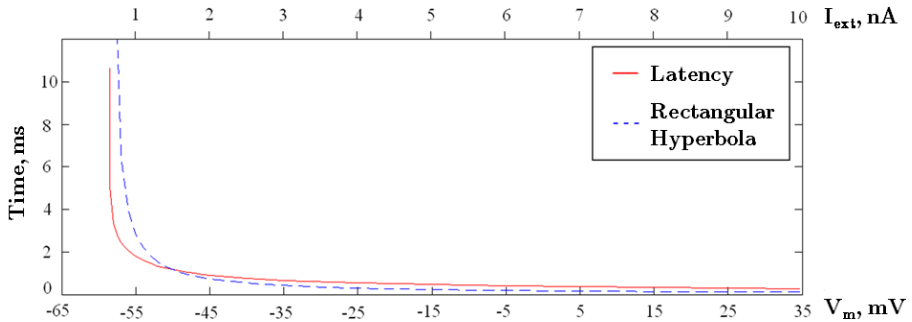
Fig. 2. The red line indicates the latency as a function of the membrane potential V_m (red marked scale), or else of the current amplitude I_{ext}, equivalently, obtained by means of simulations in NEURON environment. The dashed blue line indicates the rectangular hyperbola (i.e., the firing equation properly shifted for the comparison). In addition, the normalized scale S is reported (blue marked scale). Note that, below the S_{th} value no spike can be generated (fading blue area).

Note that, using (3) under proper considerations, it is possible to obtain T_r (*rise term*) in active mode, as follows:

$$T_r = \frac{(S_p - 1)^2 \Delta t}{1 - (S_p - 1)\Delta t} \, .$$ (4)

in which S_p represents the previous state, whereas $\Delta t = t_c - t_p$ is the temporal distance between two consecutive incoming spikes; where t_c and t_p represent the times related to current and previous states, respectively.

Note that, the denominator of (4) must be positive, thus $\Delta t < 1/(S_p - 1)$ (i.e., $\Delta t < t_{fp}$, in which t_{fp} is the previous time-to-fire value).

Equation (4) allows us to determinate the inner state of a burning neuron at the moment when it receives further inputs during the t_f time window.

Finally, for $S = S_{th} = 1 + d$, the time-to-fire is equal to $t_{f,max} = 1/d$, representing the upper bound of the time-to-fire. The latter consideration is crucial in order to have a finite maximum latency [15].

In order to make more clear the behavior of the model, we have depicted the quantities introduced by means of (1)–(2)–(3) in Fig. 3. The effect of (4) is shown in Figs. 4a-b.
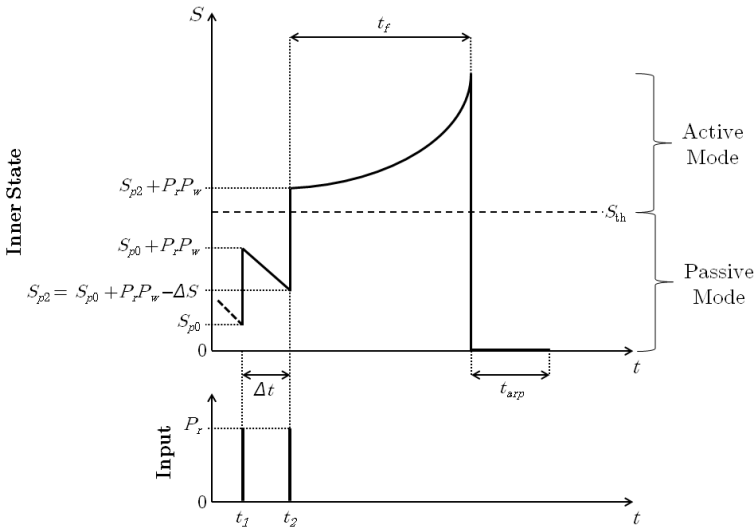
Fig. 3. In this figure, an example of the qualitatively inner state behavior of a neuron in passive and active modes is illustrated. An incoming excitatory input at t_1 causes an instantaneous increase of the state from S_{p0} to $S_{p0} + P_r P_{w1}$. At t_2 a second excitatory input is applied, then the state increase his value from S_{p2} to $S_{p2} + P_r P_{w2}$ (in this example, we have chosen $P_{w2} \neq P_{w1}$). Note that, $S_{p2} < (S_{p0} + P_r P_{w1})$; indeed, under the spiking threshold (S_{th}) the neuron is affected by a linear decay. Moreover, due to the latency effect, the firing is not instantaneous but occurs after t_f. Finally, after the firing, the neuron is reset to its resting potential (i.e., $S = 0$) for a time equal to t_{arp} (i.e., absolute refractory period).

Note that, excitatory (inhibitory) inputs increase (decrease) the inner state of a burning neuron. Therefore, when this neuron is in active mode, excitatory (inhibitory) inputs decrease (increase) the related time-to-fire (Fig. 4a and 4b, respectively).

Moreover, if the inhibitory effect is so strong to pull the burning neuron state under the spiking threshold, the time-to-fire will be canceled and the state will come back to the passive mode.
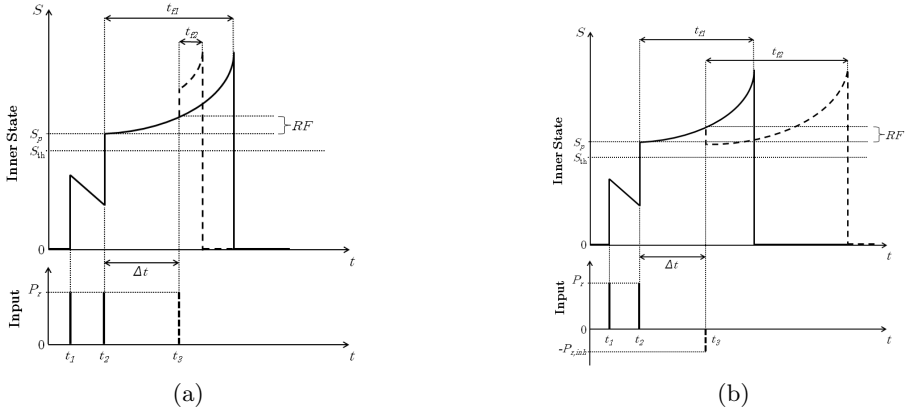
Fig. 4. (a) A hypothetical third excitatory input at t_3 (dashed pulse) would cause a reduction of the spike latency; then, the neuron would fire at $t = t_3 + t_{f2}$ (dashed line). (b) A hypothetical inhibitory input at t_3 (dashed pulse) would cause an increase of the spike latency $t = t_3 + t_{f2}$ (dashed line). In general, the amplitude of the presynaptic inhibitory input ($P_{r,inh}$) is different from the excitatory ones. In this case, a smaller value of $P_{r,inh}$ has been chosen. However, since in (a) and (b) we have assumed that each presynaptic input came from various synapses (implying different P_w values), then the $P_r P_w$ product is different. Also, T_r effect is shown.

3   Network Topology and Plasticity Rules

3.1   CNN-*like* Topology

In order to show behaviors emerging from simulations conducted through the paradigm here proposed, in this section we present a model characterized by a CNN-*like* architecture topology. Therefore, a firing neuron emits its spikes to a number of burning neurons belonging to a neighborhood, "v". It is possible to change the size of the neighborhood by setting the parameter "v".

In Fig. 5 the grids for both excitatory (Fig. 5a) and inhibitory (Fig. 5b) neurons are illustrated.

Note that, in order to maintain the balance between excitation and inhibition, the number of inhibitory neurons is less than the excitatory ones, about 15%-25% of the global neuron population [20]. Typically, this represents a condition necessary but not sufficient to guarantee the network stability, as pointed out recently [21]. Synaptic plasticity provides a further contribution in order to maintain stable the network activity.

In the model, we have considered the following synapse classes: *excitatory-to-excitatory* (s_{ee}), *excitatory-to-inhibitory* (s_{ei}), and *inhibitory-to-excitatory* (s_{ie}). Note that, self-connections are not contemplated.

Since for simplicity we assumed that also inhibitory neurons are integrators, *inhibitory-to-inhibitory* synapses are not implemented. In fact, if this class were present, this would entail a reduction of the number of inhibitory neurons dur-
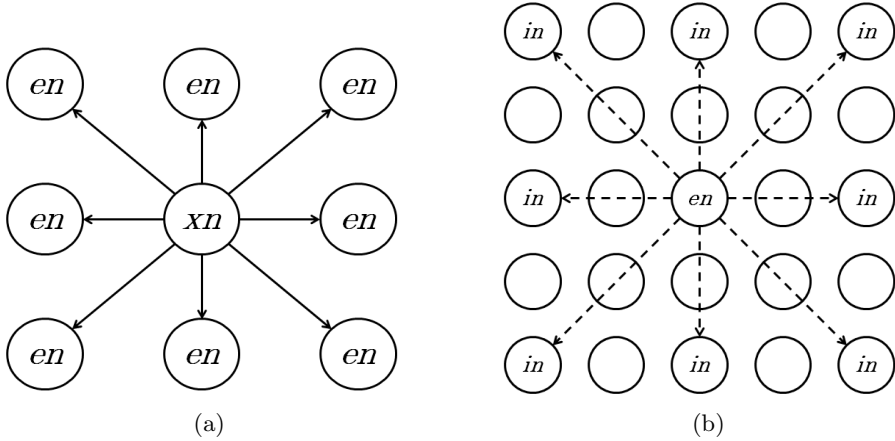
Fig. 5. (a) Synapses grid for excitatory firing neurons (*en*); *xn* indicates an excitatory or inhibitory (*in*) firing neuron. (b) Synapses grid for inhibitory neurons.

ing the simulation, pushing the network activity toward instability due to an uncontrolled excitation.

In general, the spiking activity strictly depends on the synaptic circuitry [22], but in this work we have considered a regular topological structure rather than a biological one, which is too complex to reproduce. However, this choice does not affect the basic paradigm.

Finally, in further works we will simulate different topologies such as modular (e.g., [23]), hierarchical (e.g., [24]), small-world (e.g., [25,26]), etc.

3.2 Synaptic Plasticity Rules

For the purpose of taking into account the synaptic plasticity phenomenon, we propose the following simple rules. They represent a functional simplification of the *Postsynaptic Rule*: in its most general form, this rule states that it is the positional pattern and timing of heterosynaptic inputs with respect to the homosynaptic inputs to a given synapse that governs the change in postsynaptic efficacy induced by a modifying substance at that synapse [27].

- *Exponential decay.* All postsynaptic weights grow down to the minimum value in an exponential way with a proper time constant (τ).

$$P_w = P_{w,min} + (P_w - P_{w,min})\mathrm{e}^{-\frac{\Delta t}{\tau}} \ . \tag{5}$$

- *Homosynaptic enhancement.* When a spiking event occurs from a certain synapse, the postsynaptic weight grows up, in function of previous spikes on the same neuron and from the same synapse, occurred in a specified time window (homosynaptic window).

− *Heterosynaptic enhancement.* When a spiking event occurs from a certain synapse, the postsynaptic weight grows up, in function of previous spikes on the same neuron, in a specified time window (heterosynaptic window) from other synapses.

The growing rates related to homo- and heterosynaptic rules are properly applied using the following equation:

$$\Delta P_w = \eta(P_{w,max} - P_w) \ . \tag{6}$$

with $0 < \eta \leq 1$, representing the *learning rate*. $P_{w,max}$ represents a cut-off value, but thanks to the exponential decay the saturation of the weights is avoided.

As show in subsection 4.2, these simple rules, together with the CNN-*like* topology, seem to be suitable to realize a *confinement-competition-selection model* [9].

In future works, we would also implement alternative strategies such as STDP (e.g., [28]) or Synaptic Scaling (e.g., [29]).

4 Event-Driven Simulations

4.1 Event-Driven Approach for the Network Simulation

As we have already stressed in the introduction, for the purpose of emulating a continuous-time behavior an event-driven approach is required [17], [30,31]. Therefore, a simple MATLAB code has been implemented, by means of which the simulation proceeds searching for the active neuron with the minimum time-to-fire, in order to determinate the next firing event to be scheduled in the spike timing array list. Then, the evaluation of firing event effects on all the directly burning neurons is made. A simulation procedure summary is illustrated below.

1. Pseudo-random inner state assignment for all neurons. Then, some neurons could be active when the simulation is running.
2. If all neuron states are less than the spiking threshold S_{th}, no active neuron is present. In this case, there is no activity.
3. The inputs are applied and the inner states S are computed for each burning neuron.
4. If a subset of neurons become over threshold, the time-to-fire for all active neurons will be evaluated, i.e. the following cyclic simulation procedure is applied:
   (a) Find the neuron N_0 with the minimum time-to-fire, t_{f0}. Apply to the global simulation time an increase equal to t_{f0}. According to this choice, update the time-to-fire and the inner states for each active neuron.
   (b) Firing of the neuron N_0. According to this event, make N_0 passive and update the states of all directly connected burning neurons. Update the postsynaptic weights according to the synaptic rules. Finally, the quantity t_{f0} is subtracted to the time-to-fire of all active neurons.
   (c) Update the set of the active neurons. If no active neuron is present the simulation is terminated. Otherwise, repeat the procedure from step (a).

4.2   Simulation Results

The whole distribution of the synapses is dynamically stored in a $N \times M$ matrix, called P_w (i.e., postsynaptic weights). Each entry of this matrix represents the weight of the particular synapse linking the firing neuron (j-th column of the matrix) to the burning one (i.e., i-th row of the matrix). Because of some synapses are not present (i.e., the network is not fully connected), some entries are equal to zero.

Defining nen as the number of excitatory neurons, nin as the number of inhibitory ones, nef as the number of external sources, and nt as the total number of neurons (including the external sources) the P_w matrix can be divided into submatrices:

$$P_w = \begin{pmatrix} P_{w_{11}} & P_{w_{12}} & P_{w_{13}} \\ P_{w_{21}} & P_{w_{22}} & P_{w_{23}} \end{pmatrix}$$

in which:

1. $P_{w_{11}}(pw_{i,j})$ represents the submatrix for the excitatory-to-excitatory synapses (s_{ee}), with $i = j = 1, ..., nen$; since no self-connection is considered, each entry of the main diagonal is equal to zero.
2. $P_{w_{12}}(pw_{i,j})$ represents the submatrix for the inhibitory-to-excitatory synapses (s_{ie}), with $i = 1, ..., nen$ and $j = nen + 1, ..., nen + nin$.
3. $P_{w_{13}}(pw_{i,j})$ represents the submatrix for the connections among external sources and excitatory neurons (s_{ese}), with $i = 1, ..., nen$ and $j = nen + nin + 1, ..., nt$.
4. $P_{w_{21}}(pw_{i,j})$ represents the submatrix for the excitatory-to-inhibitory synapses (s_{ei}), with $i = nen + 1, ..., nen + nin$ and $j = 1, ..., nen$.
5. $P_{w_{22}}(pw_{i,j})$ represents the submatrix for the inhibitory-to-inhibitory (s_{ii}), with $i = nen + 1, ..., nen + nin$ and $j = nen + 1, ..., nen + nin$, and it is a zero matrix as this class of connections is not considered here.
6. $P_{w_{23}}(pw_{i,j})$ represents the submatrix for the connections among external sources and inhibitory neurons (s_{esi}), with $i = nen + 1, ..., nen + nin$ and $j = nen + nin + 1, ..., nt$.

Note that, since the input signal cannot back-propagate, then the submatrices related to excitatory- and inhibitory-to-external sources, and external source self-connections are not present. Of course, nen, nin and nef can be chosen arbitrarily large.

Moreover, we have defined a $nt \times 5$ matrix, called S, in which are dynamically stored all the parameters related to all neurons of the network: the *state* (i.e., $S(:, 1)$), the *time-to-fire* (i.e., $S(:, 2)$), the *tlastfire* (i.e., $S(:, 3)$, representing the time from the last spike emitted), the *tlastburning* (i.e., $S(:, 4)$, representing the time from the last spike received) and the *presynaptic weight* P_r (i.e., $S(:, 5)$), for each neuron. This list includes the external sources, but they are not affected by the same rules of the neurons (i.e., Eqs. (1)–(2)–(3)–(4)–(5)–(6)). Indeed, external sources are thought as access nodes by which we provide spike sequences to the network.

In the following figures, we show in a 2D map the spiking activity before and after stimulation. In this case, we have applied random spike sequences. Each point of the map represents the state level in gray scale: darker points imply a high activity. Note that, in order to avoid boundary effects, the 2D neuron map (obtained by the combination of the two grids shown in Figs. 5a–b) has been folded as a taurus.

(a)   (b)
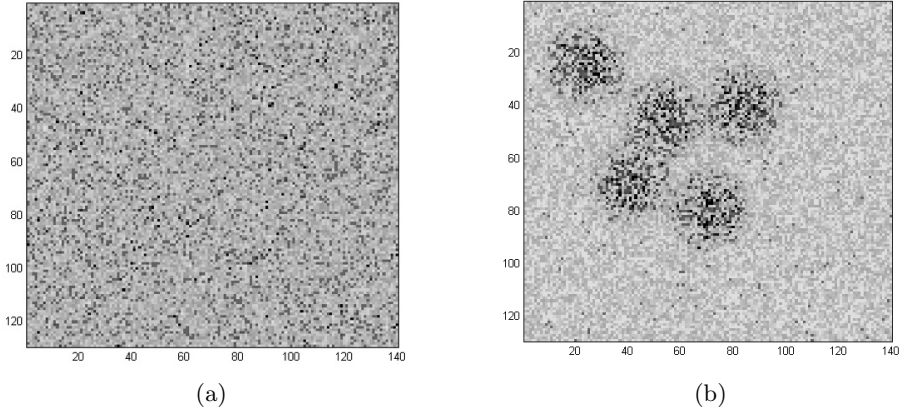
Fig. 6. (a) Spontaneous spiking activity. (b) Formation of neuronal groups after stimulation.

For this preliminary study, the network was composed by 18060 excitatory and 2021 inhibitory neurons; moreover, we have applied 25 external sources in a pseudo-random fashion when the simulation was already run. The size of the 2D map was 140×129 (i.e., the size of the excitatory neuron grid). Finally, we have chosen a neighborhood $v = 4$, then each firing neuron could fire to 80 burning neurons (i.e., $[(2v+1)^D - 1]$, with $D = 2$).

As regards the neuron model parameters, we have assumed a threshold constant $d = 0.04$, implying a spiking threshold $S_{th} = 1 + d = 1.04$, and a maximum time-to-fire $t_{f,max} = 1/d = 25$; a linear decay $Ld = 0.001$. In addition, we have considered the postsynaptic weights P_w in the range $[0.1, 3]$.

In Fig. 6a a spontaneous activity of the network is depicted, which imply that some neurons were active when the simulation was run. We have also reported the spiking activity as a consequence of a stimulation consisting of pseudo-random spike sequences (Fig. 6b). Notice the emerging of 5 neuronal groups, in which the activity is higher than the rest of the network. Moreover, when we removed the external input these groups maintained their activity stable, preserving the shapes depicted in Fig. 6b. We believe that this confinement/selection behavior is due to both the architecture topology and synaptic plasticity rules implemented.

Of course, by means of proper techniques, further studies on the "memory" implications are needed (e.g., using statistical tools).

5   Conclusions

In this paper, we have introduced a simple paradigm in order to realize a continuous-time spiking neural network simulator. Since the simulations are conducted on a digital PC, we have implemented an *ad-hoc* event-driven method based on an array list. In this array, spike times of active neurons are stored and the algorithm proceeds searching for the minimum spike time in the list. Thus, a scheduling of the events is performed. Such a method allows us to implement a "continuous-time" behavior and to reduce the computational cost as well. Note that, event-driven simulations seem to be more suitable for the purpose of emulating the realistic dynamics of biological systems; indeed, as pointed out in the introduction, clock-driven simulations cause some relevant errors in the computation, which can mask the real network behavior.

We have taken into account some important neuronal characteristics such as subthreshold decay, spike latency, synaptic integration, excitatory and inhibitory effects, and synaptic plasticity. Even though we addressed the network implementation more from a functional point of view rather than a biological one, some interesting preliminary results have been obtained. In particular, formation and maintenance of neuronal groups after stimulation have been observed.

Further works will be focused on the statistical implications about the activity of these neuronal groups and the chance of storing information, realizing then *analog memories.*

References

1. Maass, W.: Networks of spiking neurons: The third generation of neural network models. Neural Netw. 10(9), 1659–1671 (1997)
2. Belatreche, A., Maguire, L.P., McGinnity, M.: Advances in design and application of spiking neural networks. Soft Computing - A Fusion of Foundations, Methodologies and Applications 11(3), 239–248 (2006)
3. Ponulak, F., Kasiński, A.: Introduction to spiking neural networks: Information processing, learning and applications. Acta Neurobiol. Exp. 71(4), 409–433 (2011)
4. Brunel, N., van Rossum, M.C.W.: Lapicque's 1907 paper: from frogs to integrate-and-fire. Biol. Cybern. 97(5-6), 337–339 (2007)
5. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and application to conduction and excitation in nerve. J. Physiol. 117(4), 500–544 (1952)
6. Izhikevich, E.M.: Which Model to Use for Cortical Spiking Neurons? IEEE Trans. on Neural Networks 15(5), 1063–1070 (2004)
7. Izhikevich, E.M.: Polychronization: Computation with spikes. Neural Comput. 18(2), 245–282 (2006)
8. Chrol-Cannon, J., Gruning, A., Yaochu, J.: The emergence of polychronous groups under varying input patterns, plasticity rules and network connectivities. In: The 2012 International Joint Conference on Neural Networks (IJCNN), pp. 1–6. IEEE (2012)
9. Edelman, G.M.: Neural Darwinism: The Theory of Neuronal Group Selection. Basic Book, Inc., New York (1987)
10. Izhikevich, E.M., Gally, J.A., Edelman, G.M.: Spike-timing Dynamics of Neuronal Groups. Cerebral Cortex 14(8), 933–944 (2004)

11. Burkitt, A.N.: A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. Biol. Cybern. 95(1), 1–19 (2006)
12. Burkitt, A.N.: A review of the integrate-and-fire neuron model: II. Inhomogeneous synaptic input and network properties. Biol. Cybern. 95(2), 97–112 (2006)
13. Brette, R., Rudolph, M., Carnevale, T., Hines, H., Beeman, D., Bower, J.M., Diesmann, M., Morrison, A., Goodman, P.H., Harris Jr., F.C., Zirpe, M., Natschläger, T., Pecevski, D., Ermentrout, B., Djurfeldt, M., Lansner, A., Rochel, O., Vieville, T., Muller, E., Davison, A.P., El Boustani, S., Destexhe, A.: Simulation of networks of spiking neurons: A review of tools and strategies. J. Comput. Neurosci. 23(3), 349–398 (2007)
14. Citri, A., Malenka, R.C.: Synaptic plasticity: multiple forms, functions, and mechanisms. Neuropsychopharmacology 33(1), 18–41 (2008)
15. FitzHugh, R.: Mathematical models of threshold phenomena in the nerve membrane. Bull. Math. Biophys. 17(4), 257–278 (1955)
16. Chua, L., Yang, L.: Cellular Neural Networks: Theory. IEEE Trans. on Circuits and Systems 35(10), 1257–1272 (1988)
17. Mattia, M., Del Giudice, P.: Efficient event-driven simulation of large networks of spiking neurons and dynamical synapses. Neural Comput. 12(10), 2305–2329 (2000)
18. NEURON simulator, http://www.neuron.yale.edu/neuron/
19. Wang, H., Chen, Y., Chen, Y.: First-spike latency in Hodgkin's three classes of neurons. J. of Theoretical Biology 328, 19–25 (2013)
20. Okun, M., Lampl, I.: Balance of excitation and inhibition. Scholarpedia 4(8), 7467 (2009), http://www.scholarpedia.org/article/Balance_of_excitation_and_inhibition
21. Pernice, V., Staude, B., Cardanobile, S., Rotter, S.: Recurrent interactions in spiking networks with arbitrary topology. Physical Review E 85, 031916 (2012)
22. Buzsáki, G.: Rhythm of the brain. Oxford University Press, Inc. 198 Madison Avenue, New York (2006)
23. Parasuraman, K., Elshorbagy, A., Carey, S.: Spiking modular neural networks: a neural network modeling approach for hydrological processes. Water Resources Research 42(5), 1–14 (2006)
24. Wu, Q.X., McGinnity, M., Maguire, L., Cai, R., Chen, M.: Simulation of Visual Attention Using Hierarchical Spiking Neural Networks. In: Huang, D.-S., Gan, Y., Premaratne, P., Han, K. (eds.) ICIC 2011. LNCS, vol. 6840, pp. 26–31. Springer, Heidelberg (2012)
25. Watts, D.J., Strogatz, S.H.: Collective dynamics of "small-world" networks. Nature 393(1), 440–442 (1998)
26. Newman, M.E.J.: The structure and function of complex networks. SIAM Review 45(2), 167–256 (2003)
27. Finkel, L.H., Edelman, G.M.: Interaction of synaptic modification rules within populations of neurons. Proc. Natl. Acad. Sci. USA, 1291–1295 (1985)
28. Song, S., Miller, K.D., Abbott, L.F.: Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature 3(9), 919–926 (2000)
29. Sullivan, T.J., de Sa, V.R.: Homeostatic synaptic scaling in self-organizing maps. Neural Networks 19, 734–743 (2006)
30. Ros, E., Carrillo, R., Ortigosa, E.M., Barbour, B., Agís, R.: Event-Driven Simulation Scheme for Spiking Neural Networks Using Lookup tables to Characterize Neuronal Dynamics. Neural Comput 18(12), 2959–2993 (2006)
31. D'Haene, M., Schrauwen, B., Van Campenhout, J., Stroobandt, D.: Accelerating Event-Driven Simulation of Spiking Neurons with Multiple Synaptic Time Constants. Neural Comput. 21(4), 1068–1099 (2009)

Online Spectral Clustering and the Neural Mechanisms of Concept Formation

Stefano Rovetta[1] and Francesco Masulli[1,2]

[1] DIBRIS – University of Genova, Via Dodecaneso 35, 16146 Genova, Italy
[2] Temple University, Philadelphia PA, USA
{stefano.rovetta,francesco.masulli}@unige.it

Abstract. Spectral clustering can provide surprising performances. As all kernel methods, is uses a similarity matrix, whose size grows with n^2, and it requires to solve a possibly large eigenproblem. In this paper we focus on a method for spectral embedding of stream data, modeled as an unbounded quantity of input observation. A second purpose of this work is to analyze the proposed method and compare it with traditional neural network implementations: current knowledge about computations in neurons and the brain does not contrast with the computing primitives required for a local implementation of the proposed technique. A hypothesis stemming from this work could be that concept formation and discrimination in neurons and the brain could be explained by a spectral embedding framework.

Keywords: Spectral clustering, Online learning, Concept formation, Unsupervised learning, Neural networks.

1 Introduction

Spectral clustering is a family of unsupervised machine learning techniques capable of providing surprising performances, such as the detection of clusters of more or less arbitrary shape [26]. From the cognitive modeling standpoint, two weaknesses of spectral clustering are that solutions in non-clusterable data sets tend to be less meaningful than with other, more traditional techniques such as k means; and that, especially with data embedded in $\mathbb{R}^d$ where the heat kernel

$$K(\mathbf{x}, \mathbf{y}) = e^{-\|\mathbf{x}-\mathbf{y}\|/\sigma^2} \tag{1}$$

is commonly used, in the presence of clusters of different densities the choice of the parameter σ is critical and it may not be possible to find a unique optimal value [28,15]. Generalization, or "out-of-sample extension", is not directly provided by these methods, but several techniques can be used to this purpose [4,10,8]

The present work, however, is concerned with one specific computational limitation of the method. Spectral clustering, as all kernel methods [9], is based on the use of a Gram (or similarity) matrix, whose size grows with n^2 (where n is the data set cardinality), and therefore computations usually scale as n^3. The methods require solving an eigenproblem, with related computational complexity.

In this paper we focus on the problem of providing a spectral embedding solution to the problem of clustering stream data, which can be modeled as an unbounded quantity of input observation ($n \to \infty$). This is motivated by the growth of available raw stream data. For instance, some applications currently receiving a lot of attention are wearable sensors for health monitoring, data from mobile devices in crowd and traffic management in "smart city" projects, and sensors for ambient-assisted living. Clearly, solving this problem requires some approximations to the method, which will be introduced in Section 4. We provide a technique that, while not directly tested here on stream data, is nevertheless optimized by online training, making it suitable in this framework.

A second purpose of this work is to analyze the proposed method and compare it with traditional neural network implementations. We will see that current knowledge about computations in neurons and about the structure of several brain areas, for instance those related to early vision, does not contrast with the computing primitives required for a local implementation of the proposed technique.

A hypothesis stemming from this work could be that, contrary to what is suggested by many "canonical" models of computation in the brain based on simple maximum similarity matching (e.g., the "prototype effect" [12] in perception and recognition), concept formation and discrimination could be explained by a more flexible and powerful spectral embedding framework.

2   Spectral Clustering

Spectral clustering is a clustering criterion that can be justified as arising from spectral graph partitioning [6] or from several other principles, such as random walks, diffusion phenomena and the heat equation, or Laplacian-of-Gaussian filters for edge detection. In the case of data embedded in $\mathbb{R}^d$, it uses a similarity matrix W to construct a neighborhood graph, and then it analyzes the spectral properties of this graph by studying the eigenvalues and eigenvectors of the graph Laplacian $L = D - W$ or one of its normalized versions, $L_{\mathrm{rw}} = I - D^{-1}W$ or $L_{\mathrm{sym}} = I - D^{-1/2}WD^{-1/2}$, where D is the diagonal degree matrix whose values are the row (or column) sums of W.

The eigenvectors and eigenvalues of graph Laplacians provide information about the number of connected components of the graph, although in different forms depending on the normalization. In particular, L and L_{rw} have piece-wise constant eigenvectors, all corresponding to eigenvalue 0, and whose multiplicity equals the number of connected components of the graph. These eigenvectors are indicator vectors of the connected components, being 0/1 valued. If clusters are defined in a more general and realistic way, i.e., as sub-graphs which are not connected components, but have stronger within-group connectivity than between-group connectivity, then the eigenvectors are still
approximate indicator vectors.

The algorithm by Ng, Jordan and Weiss [16] is slightly different in that it uses a complementary but equivalent definition of L_{sym}, which is $L'_{\mathrm{sym}} = I - L_{\mathrm{sym}}$ and has the same eigenvectors; moreover if λ_i is an eigenvalue for L_{sym} then $1 - \lambda_i$ is an eigenvalue for L'_{sym}, so that this approach studies the largest-valued (ideally 1), as opposed to smallest-valued (ideally 0), eigenvalues of the Laplacian spectrum.

After computing the eigendecomposition of the Laplacian, the top eigenvectors are arranged as the columns of a matrix; then a spectral embedding is performed, where the l-th data point is represented by the l-th row of the matrix. Finally these representations, after having been row-normalized, are clustered with a simple method (often k means).

A very good introduction to the different flavors of spectral clustering is provided in [26]. In ref. [9] an overall survey of the properties of spectral, as well as kernel-based clustering is provided.

3 Online and Incremental Versions of Spectral Clustering

Due to the mentioned computational limitations, spectral clustering has been the subject of several modifications. These fall into two typical broad categories. The first one is that of exact algorithms which exploit the sparsity of graph data, i.e., the fact that the number of edges is less than $n(n-1)/2$. The second category is approximated algorithms, where the approximation may apply to the data (not all data are kept), to the similarity matrix, or to other aspects. Many of these modifications are iterative and can be used for online training, although not all are suitable for the clustering of stream data.

An instance of the first category is the approach by H. Ning [17] which directly updates the eigenvectors and eigenvalues by decomposing the graph and identifying only the individual elements that need updating. It applies only to cases of graph with limited connectivity; the approach is suitable for instance for studying the Internet graph, although it relies on some hypotheses on the extent of modifications required at each updating step, which should be limited.

The second category is represented by the "Nyström method", i.e., the use of the Nyström formula to obtain a reduced-rank approximation of a Gram (similarity) matrix, which has been proposed for use in spectral clustering in [10]. The "fast approximate" method from [27] approximates the data rather than the eigensystem by using a pre-clustering step, to which spectral clustering is then applied.

4 An Approximated, Online Spectral Clustering Method

Rather than approximating the data or the eigensystem, in this work we propose to approximate the matrix W and therefore the normalized Laplacian. The technique includes two steps: approximation and eigendecomposition.

For the first step, approximation, we assume that the input data $\mathbf{x}$ are satisfactorily described as realizations of a stationary, discrete-time, stochastic vector process

$$\mathbf{x} = \mathbf{c}_j + v, \qquad j \in \{1 \dots m\}, \tag{2}$$

where j is a random integer between 1 and m and v is a random (noise) term for which we assume a reasonable probability distribution, i.e., unimodal, symmetric, and zero-centered. Therefore we approximate the data with a set $\{\mathbf{c}_1, \dots, \mathbf{c}_c\}$ of reference or landmark [7] points which are optimized to minimize a mean squared distortion criterion

$$J = \int (\mathbf{x} - \mathbf{c}(\mathbf{x}))^2 p(\mathbf{x}) d\mathbf{x}, \tag{3}$$

where $\mathbf{c}(\mathbf{x})$ is the nearest landmark to data point $\mathbf{x}$. This is a vector quantization problem. We require that its solution approximately reflect the data distribution $p(\mathbf{x})$, but it does not necessarily have to pinpoint any structure (clusters) within it. A vector quantization problem is usually solved by stochastic approximation methods.

The landmark points are then used for approximating the asymmetric normalized Laplacian

$$L_{\text{rw}} = D^{-1}W \tag{4}$$

by replacing the computation of the similarity $K(\mathbf{x}, \mathbf{x}')$ of any given data point to the remaining points $\mathbf{x}'$, a set of possibly infinite cardinality in our hypotheses, with the similarity of the same point $\mathbf{x}$ to each landmark $\mathbf{c}_j$, a set of finite cardinality m. The chosen similarity function $K(\cdot, \cdot)$ can be for instance the heat kernel (1). The task is therefore that of identifying the similarity matrix $W_{jk} = K(\mathbf{c}_j, \mathbf{c}_k)$ from a possibly unbounded sequence of observed samples $\mathbf{x}$ generated according to model (2). Note that, according to the asymmetric normalization chosen, if the current sample is $\mathbf{x} = \mathbf{c}_j + v$ we have

$$L_{jk} = \frac{W_{jk}}{\sum_h W_{jh}} \approx \frac{K(\mathbf{x}, \mathbf{c}_k)}{\sum_h K(\mathbf{x}, \mathbf{c}_h)}, \tag{5}$$

which depends only on $\mathbf{x}$, not on other samples as it would in the case of the symmetrical normalization $D^{-1/2}WD^{-1/2}$.

In the second step the structure of the data distribution is analyzed by means of the eigendecomposition of the normalized Laplacian. Since in this setting the Laplacian is noisy (random) and given by a sequence of row vectors, each of the form

$$\left[\frac{K(\mathbf{x}, \mathbf{c}_1)}{\sum_h K(\mathbf{x}, \mathbf{c}_h)}, \cdots, \frac{K(\mathbf{x}, \mathbf{c}_m)}{\sum_h K(\mathbf{x}, \mathbf{c}_h)} \right],$$

then we may use Oja's subspace rule [19,18] which gives the eigendecomposition of a matrix $S^T S$, known through a sequence of noisy samples of S; the eigenvectors of $S^T S$ are the same as the right eigenvectors of S (while the eigenvalues are squared w.r.t. those of S). Note that, since this algorithm requires centered data, the mean input vector is also learned as a set of bias terms, and then subtracted.

To complete the spectral clustering process, the embedded data should be clustered, usually by k means. However, due to the properties of the spectral embedding, this last step is usually almost trivial. In this work we will mainly focus on embedding.

To sum up, the following is an outline of the proposed online algorithm:

1. Input one pattern $\mathbf{x}$
2. Compute similarities from landmarks: $K(\mathbf{x}, \mathbf{c}_j)$
3. Compute the corresponding row of the normalized Laplacian: $\lambda_j = \frac{K(\mathbf{x}, \mathbf{c}_j)}{\sum_h K(\mathbf{x}, \mathbf{c}_h)}$
4. Update landmarks
5. Compute one subspace step using Laplacian row as input
6. Update subspace projection
7. Go to step 1
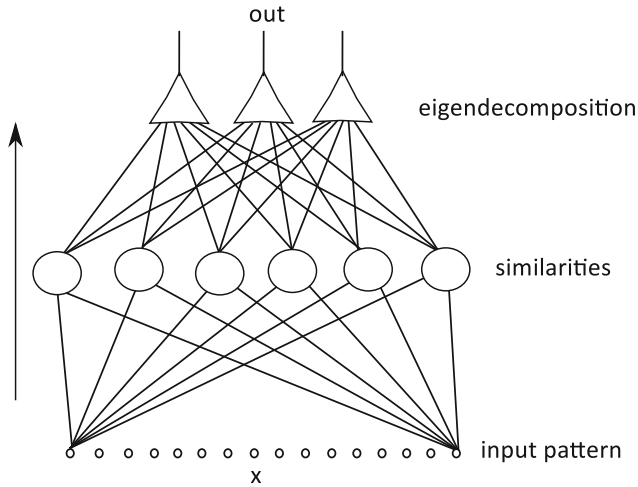
Fig. 1. Graphical representation of the method. Small circles are inputs; large circles compute similarities, each unit storing one prototype $\mathbf{c}_j$; triangles compute eigendecomposition of the approximated Laplacian, each unit computing the projection on one dimension of the embedding space. For clarity not all connections between inputs and similarity units are shown.
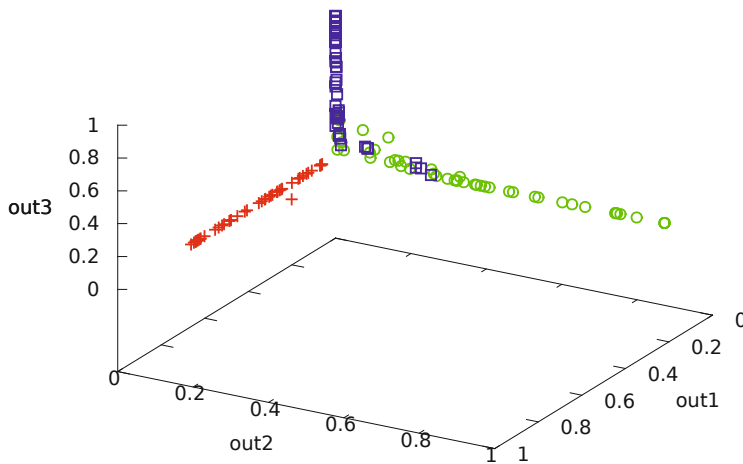
Fig. 2. Iris data: The spectral embedding obtained. Crosses is *Setosa*, circles is *Virginica*, and squares is *Versicolor*; "out1", "out2" and "out3" indicate three output components.

5   Experimental Results

Some experiments have been performed to check the consistency of the proposed method with standard approaches, rather than proving its quality in absolute terms. The data sets used are Anderson's Iris data [2], a data set composed of two concentric circles ("circles", see left of Fig. 3), and a data set of random samples from the letters W I R N ("WIRN", see top of Fig. 4).

Anderson's Iris data needs little presentation. It is a three-class dataset with 4 inputs (petal width, petal length, sepal width, sepal length, all expressed in centimeters) and 50 instances per class, for a total of 150 patterns. The data set was downloaded from the UCI Machine Learning repository [3].

The circles and WIRN data were generated by randomly sampling points from 2-dimensional geometrical structures within the (dimensionless) square $[0, 1] \times [0, 1]$, as shown in the respective figures, so they are both 2-input data sets. The circles data has two clusters and 200 instances per cluster (total cardinality: 400), whereas the WIRN data has four clusters and instances distributed as follows: 499 in cluster "W", 226 in "I", 469 in "R" and 376 in "N" (total cardinality: 1570). The different cardinalities are due to sampling different shapes and sizes with uniform random density.

The latter two experiments can be directly compared with the results presented in [16] on similar data. Note, however, that both in traditional approaches and in the online version proposed here the results are strongly dependent on the choice of σ, therefore comparisons, even those contained in [16], are not necessarily fair.

The method was implemented in C++. Vector quantization was performed with a centroid optimization heuristic similar to online k means, with some degree of inter-action between the best-matching vector and the remaining ones, to reduce the risk of false minima; the interaction degree was annealed during training with an exponential
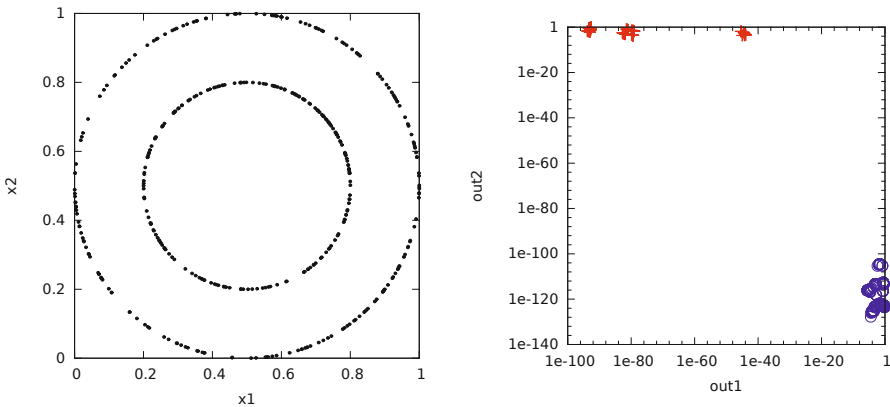
Fig. 3. The "Circles" data set. Left, data; "x1" and "x2" indicate two input components. Right: Spectral embedding (learned representation), log-log scale; "out1" and "out2" indicate two output components.

rule. Eigendecomposition and mapping was performed with Oja's generalized Hebb rules plus an orthogonalization step (a method roughly equivalent to Sanger's GHA rule [24]). Both steps are purely online, with no memory required in addition to the already described quantities: landmarks, eigenvectors, bias terms.

The figures show the data and the results of the spectral embeddings obtained. These are graphs with axes representing the two or three components of the embeddings themselves (actually the plotted values are the outputs of the network corresponding to each input pattern presented). From the figures it is clear that the subsequent, actual clustering is trivial with circles and WIRN, while for Iris, which only contains two separable clusters, some misattributed data remain, as it is to be expected.
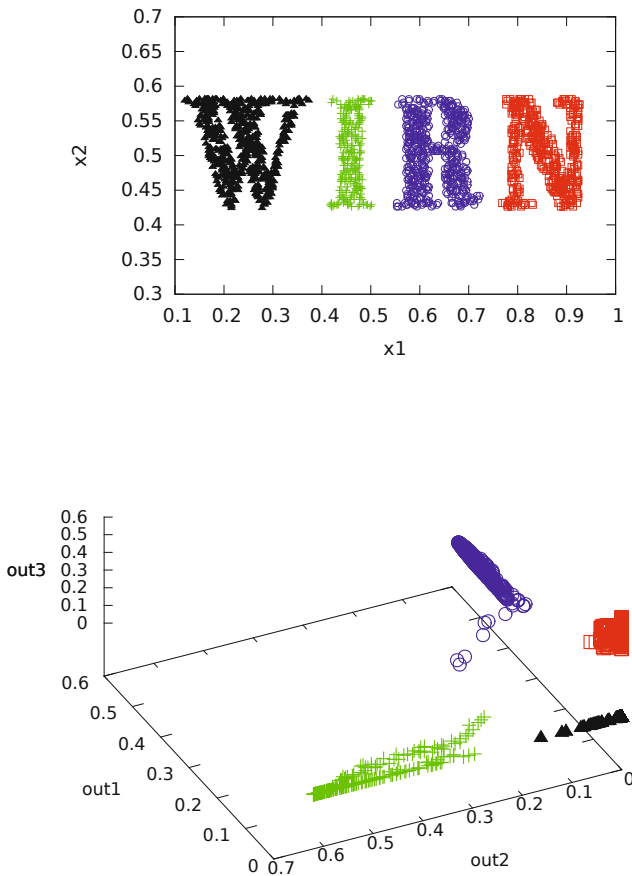
Fig. 4. The "WIRN" data set. Top, data labeled after the clustering result; "x1" and "x2" indicate two input components. Bottom, spectral embedding (learned representation); "out1", "out2" and "out3" indicate three output components.

For the WIRN data also the result of the final clustering step, using k means on the normalized output patterns, is presented.

In general, the embeddings have a dimensionality that equals the number of clusters sought, so for Iris all three components are shown. For WIRN we show three of the four components.

6   A Neural Implementation

An intriguing property of the method presented is that it is completely local and, as already noted, it requires constant memory w.r.t. data cardinality, therefore it is a good candidate for a distributed implementation. But we will go as far as showing that the required computational primitives are indeed *not incompatible* with commonly accepted input-output response models found in the nervous system, for instance in early vision stages. This suggests that the operating mechanism of some areas in the nervous system could be actually implementing a form of spectral embedding for learning representations.

Regarding the landmark set $\{c_j\}$, a basic competitive update rule was used, as described in the experimental section.

As the similarity function $K(\cdot, \cdot)$, up to now we have referred to the Gaussian similarity or heat kernel (1). However, we note that a similar function can be obtained with a standard linear threshold formal unit, modified to include a sum-of-squared-inputs term as follows:

$$r(\mathbf{x}) = w_q \mathbf{x} \cdot \mathbf{x} + \mathbf{w} \cdot \mathbf{x} + w_0 \qquad\qquad a(r) = \frac{1}{1 + e^{-r}}, \qquad (6)$$

where $\mathbf{x} \cdot \mathbf{x} = \sum_{i=1}^{d} x_i^2 = \|\mathbf{x}\|^2$, r is the net stimulus and a the activation value, obtaining what has been termed a *circular perceptron* [22] due to the hyper-spherical shape of its discriminant surface. With suitable constraints on the weights w_q, $\mathbf{w} = [w_1, \ldots, w_d]$, and w_q this model can be interpreted as implementing the formal neural function

$$r(\mathbf{x}) = \frac{\|\mathbf{x} - \mathbf{c}\|^2 + \theta}{\sigma^2} \qquad\qquad a(r) = \frac{1}{1 + e^{-r}} \qquad (7)$$

by means of the following conversions:

$$\begin{cases} w_q = \frac{1}{\sigma^2} \\ \mathbf{w} = -\frac{\mathbf{c}}{\sigma^2} \\ w_0 = \frac{\|\mathbf{c}\|^2 - \theta}{\sigma^2} \end{cases} \Leftrightarrow \begin{cases} \sigma = \frac{1}{\sqrt{w_q}} \\ \mathbf{c} = -\frac{\mathbf{w}}{w_q} \\ \theta = \frac{\|\mathbf{w}\|^2 - w_0}{w_q} \end{cases}$$

A network including these "circular neurons" has been shown [23] to be equivalent in several respects to a vector quantization network. The presence of a quadratic term introduces a (biologically plausible [11]) dependence of the output response on the overall input intensity, not only on the net input. Figure 5 shows the excellent degree of coincidence between the circular activation function and the Gaussian one, attainable with suitable parameter values.
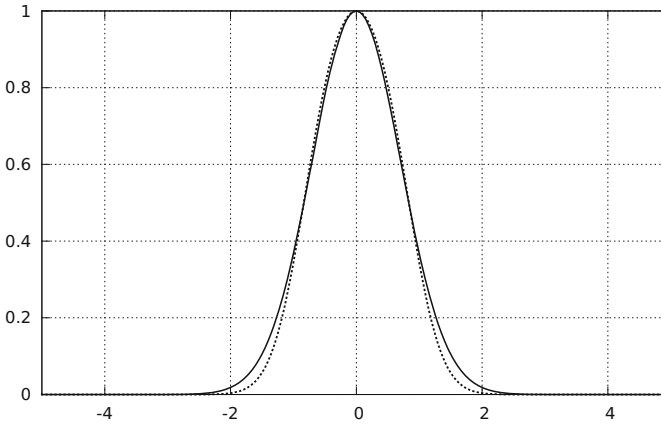
Fig. 5. Comparison of circular unit (dotted) and Gaussian (continuous) activations as functions of their net stimulus.

Competitive learning rules and orthogonalization can also be explained by means of Heeger and Carandini's normalization model [5], which explains experimental data that indicate the presence of a net inhibitory effect of neurons within a group, even in the absence of inhibitory synapses. This effect in turn can be explained in the framework of retrograde signaling [1], neural backpropagation [25], and neuromodulation. These considerations support and reinforce the common idea that both competitive updating rules and the Hebbian learning rule are biologically plausible and indeed may be a model of some mechanisms of synaptic plasticity in the central nervous system.

7 Concept Formation

During the experiments it has been observed that, after an initial phase of "ramping up", adaptation proceeds quite smoothly. As soon as some structure appears in the set of landmarks, the second layer can start learning meaningful eigencomponents. These usually do not change abruptly; it appears that the evolution stays smooth as long as there is no change in attraction basins of landmarks in the first layer. This makes the selection of prototypes in the first layer not critical, as also observed in [7] with experiments on the landmark MDS, a method that shares some elements with the present one.

By performing several experiments it has additionally been observed that, for a suitable selection of σ, or equivalent parameters in case of different formulations of the prototype units, for well separated clusters the embedding tends to be binary, i.e., most coordinates of the embedded point are zero and only one is significantly different from zero. This tendency to form sparse representation is a known property of the Laplacian eigensystem. Moreover, except for possible arbitrary axis permutations, the locus

of embedding points tends to stay about the same regardless of the actual location of landmarks, which depends on the random initialization.

The proposed model therefore shows some very interesting properties:

- Similarly to its standard counterparts (spectral clustering methods), it automatically points out interesting structure from fairly complex data distributions, with higher flexibility than prototype-based models.
- It intrinsically produces sparse internal representations starting from distributed intermediate representations, that tend to move from analog to binary.
- Different exemplars produce similar internal models. This may provide them with an ability to develop a "theory of mind", i.e., a shared representation for concepts and consequently the possibility to model other individual's internal state.

It can be noted that all these properties are obtained by only using computational primitives that are commonly accepted as biologically plausible and have been in use for decades.

It is interesting to note how the observed behaviour is also compatible with experiments and theories of concept formation as observed in the human brain, as well as in primates and other mammals. For instance, in the case of vision, the representation shift from distributed to localized is a well-known organizational aspect: the initial representation is obviously completely distributed on the individual receptors of the retina, then it gets organized into receptive fields [13]; in specific areas of the visual cortex it is gradually made more selective, where neurons traditionally termed "complex" and "hypercomplex" cells [14] implement a hierarchy of sparse distributed representations; finally, there is substantial experimental evidence [21] of the existence of "concept cells" in the medial temporal lobe, that implement a completely localist representation individually elicited by abstract cognitive tasks such as recognizing the face, or even just reading the name, of a known person (e.g., "Jennifer Aniston neurons" [20]).

8   Conclusions and Future Work

The work presented in this paper is just an initial proposal, and several issues should be further investigated. Among these, the technical problems of selecting a proper value for σ and a proper scheduling (or online modulation) for the learning steps. This latter point can be an opportunity to develop a supervised version of the network, as well as to incorporate some form of novelty detection to modulate the stability/plasticity dilemma.

A combined landmark-eigensystem learning method can also be investigated, to possibly reduce the number of model parameters.

Finally, the model is especially well suited for being employed in modular and multi-layer, possibly deep structures, since it is unsupervised and for training it only uses information local to each layer.

References

1. Alger, B., Pitler, T.: Retrograde signaling at GABAA-receptor synapses in the mammalian CNS. Trends in Neurosciences 18(8), 333–340 (1995)
2. Anderson, E.: The irises of the gaspe peninsula. Bulletin of the American Iris Society 59, 25 (1935)
3. Asuncion, A., Newman, D.J.: UCI machine learning repository (2007)
4. Bengio, Y., Paiement, J.F., Vincent, P., Delalleau, O., Le Roux, N., Ouimet, M.: Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering. Mij 1, 2 (2003)
5. Carandini, M., Heeger, D.J.: Normalization as a canonical neural computation. Nature Reviews Neuroscience 13(1), 51–62 (2012)
6. Chung, F.R.K.: Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92). American Mathematical Society (February 1997)
7. De Silva, V., Tenenbaum, J.B.: Sparse multidimensional scaling using landmark points. Tech. rep., Technical report, Stanford University (2004)
8. Drineas, P., Mahoney, M.W.: On the nyström method for approximating a gram matrix for improved kernel-based learning. The Journal of Machine Learning Research 6, 2153–2175 (2005)
9. Filippone, M., Camastra, F., Masulli, F., Rovetta, S.: A survey of kernel and spectral methods for clustering. Pattern Recognition 40(1), 176–190 (2008)
10. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the nystrom method. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(2), 214–225 (2004)
11. Heeger, D.J.: Half-squaring in responses of cat striate cells. Visual Neuroscience 9, 427–443 (1992)
12. Homa, D., Cornell, D., Goldman, D., Shwartz, S.: Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. Journal of Experimental Psychology 101(1), 116 (1973)
13. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology 160(1), 106 (1962)
14. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. The Journal of Neurophysiology 28(2), 229 (1965)
15. Nadler, B., Galun, M.: Fundamental limitations of spectral clustering. In: Schölkopf, B., Platt, J., Hoffman, T. (eds.) Advances in Neural Information Processing Systems 19, pp. 1017–1024. MIT Press, Cambridge (2007)
16. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: Analysis and an algorithm. In: Dietterich, T.G., Becker, S., Ghahramani, Z. (eds.) Advances in Neural Information Processing Systems 14. MIT Press, Cambridge (2002)
17. Ning, H., Xu, W., Chi, Y., Gong, Y., Huang, T.S.: Incremental spectral clustering by efficiently updating the eigen-system. Pattern Recognition 43(1), 113–127 (2010)
18. Oja, E.: Neural networks, principal components, and subspaces. International Journal of Neural Systems 01(01), 61–68 (1989)
19. Oja, E., Karhunen, J.: On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix. Journal of Mathematical Analysis and Applications 106(1), 69–84 (1985)
20. Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., Fried, I.: Invariant visual representation by single neurons in the human brain. Nature 435(7045), 1102–1107 (2005)
21. Quiroga, R.Q.: Concept cells: the building blocks of declarative memory functions. Nature Reviews Neuroscience 13(8), 587–597 (2012)
22. Ridella, S., Rovetta, S., Zunino, R.: Circular back–propagation networks for classification. IEEE Transactions on Neural Networks 8(1), 84–97 (1997)

23. Rovetta, S., Zunino, R.: Circular backpropagation networks embed vector quantization. IEEE Transactions on Neural Networks 10(4), 972–975 (1999)
24. Sanger, T.D.: Optimal unsupervised learning in a single-layer linear feedforward neural network. Neural Networks 2(6), 459–473 (1989)
25. Stuart, G., Spruston, N., Sakmann, B., Häusser, M.: Action potential initiation and backpropagation in neurons of the mammalian CNS. Trends in Neurosciences 20(3), 125–131 (1997)
26. Von Luxburg, U.: A tutorial on spectral clustering. Statistics and Computing 17(4), 395–416 (2007)
27. Yan, D., Huang, L., Jordan, M.I.: Fast approximate spectral clustering. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 907–916. ACM (2009)
28. Zelnik-Manor, L., Perona, P.: Self-tuning spectral clustering. In: Advances in Neural Information Processing Systems, pp. 1601–1608 (2004)

Part III
Pattern Recognition

Machine Learning-Based Web Documents Categorization by Semantic Graphs

Francesco Camastra[1], Angelo Ciaramella[1],
Alessio Placitelli[2], and Antonino Staiano[1]

[1] Dept. of Science and Technology, University of Naples "Parthenope", Isola C4,
Centro Direzionale, I-80143, Napoli (NA), Italy
{camastra, angelo.ciaramella, staiano}@ieee.org,
[2] Vitrociset s.p.a., Via Tiburtina, 1020 - 00156 Roma, Italy
alessio.placitelli@gmail.com

Abstract. This work aims to approach web pages categorization by means of semantic graphs and machine learning techniques. We propose to use a semantic graph that can provide a compact and structured representation of the concepts present in a document in order to take into account the semantic information. The semantic graph allows determining a map of the semantic areas contained in the document and their relationships w.r.t. a particular concept or term. The semantic measure between the terms is calculated by using the lexical database (i.e., WordNet). The document categorization is accomplished by a machine learning technique. We compare the performance of both supervised and unsupervised techniques (i.e., Support Vector Machine and Self Organizing Maps, respectively). The proposed methodology has been applied for classification and agglomeration of benchmark and real data. From the analysis of the results it can be shown that the model trained with semantic features obtains satisfactory results, in particular by using the unsupervised machine learning technique.

1 Introduction

With the dramatically quick and explosive growth of information available over the Internet, World Wide Web has become a powerful platform to store, disseminate and retrieve information as well as mine useful knowledge [3]. Information is mostly in the form of unstructured data. As the data on the web has been growing, it has lead to several problems such as increased difficulty of finding relevant information and extracting potentially useful knowledge. Web mining is an emerging research area focused on the application of data mining techniques to discover patterns from the Web. According to analysis targets, Web mining can be divided into three different types, which are Web usage mining, Web content mining and Web structure mining. In this work we address the problem of Web content mining. Web content mining extracts information from different Web sites for its access and knowledge discovery. In particular, we study a novel methodology for Web pages categorization considering the textual content.

In the past 20 years, the number of text documents in digital form has grown exponentially [12]. As a consequence of this exponential growth, great importance has been put on the classification of documents into groups that describe the content of the documents. The function of a classifier is to merge text documents into one or more predefined categories based on their content. Each document can belong to several categories or may present its own category. In [9] the authors review the Web-specific features and algorithms that have been explored and found to be useful for Web page classification. Most approaches described in literature do not consider the semantic information in the document and therefore in some cases may not perform adequately. In [1] an approach to incorporate concepts from background knowledge into document representations for text document classification (by using boosting machine learning technique) has been proposed. To extract concepts from texts, the authors have developed a detailed process, that can be used with any ontology with lexicon. Our work aims to approach web pages categorization by means of semantic graphs and machine learning techniques [1]. The semantic graph allows determining a map of the semantic areas contained in the document and their relationships w.r.t. a particular concept or term. The similarity between the terms is calculated by using the lexical database (i.e., WordNet) and the pages are represented using a TF-IDF (*Term Frequency-Inverse Document*) mechanism.

The paper is organized as follows. In Section 2, we introduce the categorization problem of documents, and, in Section 3 the TF-IDF methodology is presented. In Section 4 we describe the semantic graph and how to use it for the the TF-IDF methodology. In Section 5, the experimental results on benchmark and real data are presented. Finally, some conclusions and future remarks are outlined.

2   Document Categorization

Categorization of documents refers to the problem of automatic classification of a set of documents in *classes* (or *categories* or *topics*). A common approach for text classification is formed by five steps. The first step (*tokenization*) eliminates the punctuation signs in the text. The second step (*stopping*), removes from the text the so-called *stopping words*, i.e., common words (e.g, articles, modal verbs, prepositions) that are widespread in every text and therefore cannot be used for discriminating a text. The third step is the *stemming*, where each term is reduced to own lexical root (or *stem*) by means of a *stemming algorithm* (e.g., *Porter's algorithm*). In the fourth step, the document is represented by means of a vector whose generic i-th coordinate is computed by TF-IDF (*Term Frequency-Inverse Document*) approach [11]. Finally, the document classification is performed by a machine learning technique. The approach described above does not consider the semantic information in the document and therefore in some cases may not perform adequately.

---

[1] The work was made when Alessio Placitelli was M. Sc. Student at University of Naples Parthenope.

3   Scoring

To extract information from a document we compute a score between a query term t and a document d, based on the weight of t in d. The simplest approach is to assign the weight to be equal to the number of occurrences of term t in document d ($tf_{t,d}$, *Term Frequency* of the term t in document d) [6]. Raw term frequency as above suffers from a critical problem: all terms are considered equally important when it comes to assessing relevancy on a query. In fact, certain terms have little or no discriminating power in determining relevance. A mechanism for attenuating the effect of terms that occur too often in the collection to be meaningful for relevance determination. An idea could be to reduce the $tf_{t,d}$ weight of a term by a factor that grows with its collection frequency. It is more commonplace to use for this purpose the document frequency df_t, defined to be the number of documents in the collection that contain a term t. Denoting the total number of documents in a collection by N, we define the inverse document frequency (idf) of a term t as follows:

$$idf_t = \log \frac{N}{df_t}. \tag{1}$$

Thus, the *idf* of a rare term is high, whereas the *idf* of a frequent term is likely to be low. We now combine the definitions of term frequency and inverse document frequency, to produce a composite weight for each term in each document. The TF-IDF weighting scheme assigns to term t a weight TF-IDF in document d given by

$$\text{TF-IDF}_{t,d} = tf_{t,d} \times idf_t. \tag{2}$$

We may view each document as a vector with one component corresponding to each term in the dictionary, together with a weight for each component that is given by equation (2).

4   Semantic Graph

In order to take into account the semantic information, we propose to use a semantic graph that can provide a compact and structured representation of the concepts present in a document. The semantic graph allows determining a map of the semantic areas contained in the document and their relationships w.r.t. a particular concept or term, called *target*. The semantic weight indicates how much the document is relevant w.r.t. the target. The semantic graph is a undirected, fully connected graph, consisting of the terms of the document connected by relations of similarity to a target term.

A semantic graph is computed starting from a single term. Let t be the term whose semantic graph has to be computed and N is the number of the most similar terms in the document, the construction of the semantic graph is performed by means of four well-defined phases: *similarity calculation*, *ranking*, *graph construction* and *semantic weight calculation*. The similarity (s) between the terms
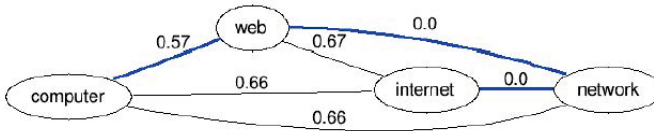
Fig. 1. Minimum spanning tree computed by the algorithm of Kruskal

is calculated by using the lexical database WordNet [7]. Next, the terms more similar to t, are ranked on the basis of a properly chosen similarity metric. Now the top N terms are used as the vertices of the undirected weighted semantic graph. For each pair of vertices an edge is created. The weight of the edge is proportional $(1 - s)$ to the semantic distance between the terms (e.g., Lin similarity in the $[0, 1]$ interval [5]). For instance, consider the construction of a semantic graph related to the target term *computer* based on a document containing the terms: *Internet, www, cat, network, software, computer, web*, and *homepage*. The information contained in the semantic graph can be represented by a single synthetic value called *semantic weight* (w_s). This value is obtained calculating the sum of the reconstructed weights $(1 - w_s)$ for the arcs belonging to the Minimum Spanning Tree (MST) of the semantic graph. The weight indicates that the semantic parsed document, represented through the semantic graph, is relevant to the target word. The higher the value of the weight, the more the document refers to the subject matter from the end target. On the contrary, the smaller this value, the less the document identifies the target. In Figure 1, the minimum spanning tree computed by the algorithm of Kruskal is presented. The final semantic weight is proportional to this estimated weight. Summarizing, the steps of the proposed categorization process are as follows. The first three steps are the same as in usual categorization process (i.e., tokenization, stop words removal, stemming), in the fourth step, to each term it is associated the semantic weight, instead of the usual TF-IDF value. In Figure 2 , the use of a semantic graph in a bag of words mechanism[6], is shown. Finally, using the TF-IDF vectors we have performed the document categorization by means of a machine learning technique. In this specific case, both Support Vector Machine (SVM) [2] and Self Organizing Maps (SOMs) [4] have been applied.

5   Experimental Results

The proposed methodology has been applied for classification (SVM) and agglomeration (SOM) of two different corpora. To evaluate the performance, the results are compared with those obtained by the standard TF-IDF mechanism considering different metrics. For the SOM, we consider the quantization error (QE), the topographic error (TE) and the combined error (CE). Regarding the SVM, the percentage of documents correctly classified is evaluated (for further results as confusion matrix, measures of precision and recall see [8]).
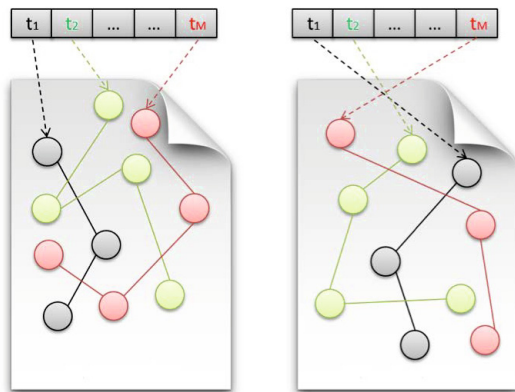
Fig. 2. Bag of words mechanism and semantic graph: extraction of features from documents

In a first phase of validation, the Reuters 21578 corpus has been considered [10]. This corpus was issued by the multinational Reuters in 2000 and made publicly available for research purposes. We used only three categories for a total of 390 documents divided as follows: Cocoa (55 documents), Money Supply (138 documents) and Ship (197 documents). In Tables 1 and 2, we report the results applying SOM and SVM techniques, respectively. Using SOM, in the case of semantic weights, a QE of 41.961, a TE of 0 and a CE of 48.414 are obtained. Instead the SVM, as reported in Table 1, has allowed obtaining a percentage of correct classification of 97.17% (379 documents). In the case of standard TF-IDF, SOM has detected a QE of 43.675, a TE of 0.005 and a CE of 49.820. The percentage of correct classification obtained through SVM is of 93.58% (365 papers).

Successively, a scraping software has been used for analyzing a web page and extract the main content excluding tags, templates and other kind of unnecessary code [8]. In order to build the corpus, the scraper was launched for five days (from March 15, 2013 to March 20, 2013) by performing the scraping of 1995 journalistic news of 5 different categories: *politics*, *sport*, *business*, *science* and *entertainment*. In Table 3, we describe the categories and information sources for the corpus [2]. The dictionary of the processed corpus is composed initially of 27305 terms. The removal of low-frequency terms leads to the elimination of 15619 words, decreasing the size of the dictionary terms to 11686 terms.

The SVM training was performed using a linear kernel and a cost coefficient of $C = 1.0$. The main objective of these experimental results is to compare TF-IDF and SWA approaches. For this reason we chose to use a simple linear kernel and to consider the same weight between the slack variability penalty and the margin in SVM optimization mechanism [2]. A 10-fold cross-validation is performed. On

[2] The corpora are available on request.

Table 1. SVM results: Reuters and scraped corpora

SVM	training set	training set	test set
TF-IDF	93.58%	81.35%	62%
SWA	97.71%	79.79%	70.36%
	Reuters	Real data	Real data

Table 2. SOM results: Reuters and scraped corpora (QE = Quantization Error; TE = Topographic Error; CE = Combined Error)

SOM	QE	TE	CE	QE	TE	CE	Test set	Test set
TF-IDF	41.96%	0.0%	48.41%	75.04%	0.01%	81.74%	83.19%	75.12%
SWA	43.67%	0.0%	49.82%	47.41%	0.04%	58.43%	86.68%	79.37%
	Reuters	Reuters	Reuters	Real data	Real data	Real data	Real data	Feature Selection

Table 3. Information sources used for the corpora (feed RSS)

Category	Information sources
Politics	The Guardian, The Telegraph, The Scotsman, BBC
Sports	The Guardian, The Independent, The Telegraph, Daily Mail, The Express, The Daily Star, The Scotsman, BBC
Business	The Guardian, The Telegraph, Daily Mail, The Express, BBC
Science	The Guardian, The Independent, The Telegraph, Daily Mail, The Express, The Daily Star, The Scotsman (Technology), BBC
Entertainment	The Guardian (Movie), The Telegraph, Daily Mail, The Express, The Scotsman, BBC

Table 4. Test corpus

Category	# of documents	Words (average)	Characters (average)
Politics	47	718	4224
Sport	210	523	3057
Business	165	590	3479
Science	141	584	3585
Entertainment	137	595	3448

the standard TF-IDF vectors, the classification percentage is of 81.35% (on the overall training set). In the unsupervised case, the SOM has a height of 17 and a width of 13 neurons. The estimated QE is 75.047, the TE is of 0.015 while the CE is 81.740. By using the semantic wights the SVM classification is of 80.760%. Instead, SOM has produced a QE of 47.410, a TE of 0.047 and a CE of 58.433. In Figure 3 we show the comparison between the topological mapping of the SOM obtained with TF-IDF and by using semantic weights, respectively. The results and the figures show that by using semantic weights is possible to obtain a more regular topographic map. Moreover, the generated models are evaluated through a test corpus composed by 700 terms (see Table 4 for details). We obtain a percentage of 62% of correct classifications using the TF-IDF features and of 70.36% using
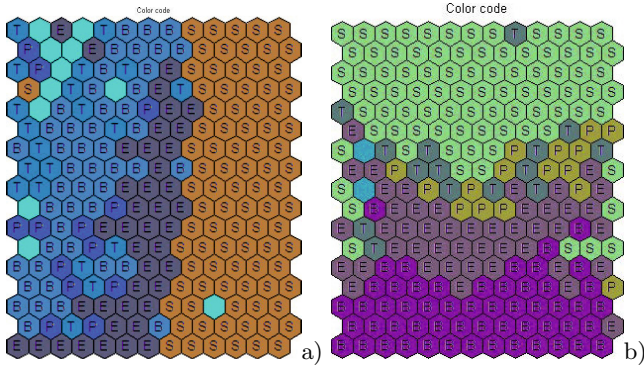
Fig. 3. Topographic result of the SOM: a) with TF-IDF features; b) with semantic features

semantic features. The performance obtained on the trained SOM are of 83.19% of correct classification for TF-IDF and of 86.68% for the semantic case.

Finally, from the previous corpus we generate a reduced corpus obtained by using a semantic feature selection. The semantic feature selection is obtained by using an aggregation approach based on the similarity measures and WordNet. In particular, the words are chosen by using an agglomerating rate and considering the most dissimilar ones for building the bag of words [8]. In this case we note that the best performance are obtained using the SOM with semantic weights ($79, 37\%$ of perfect classification).

6   Conclusions

In this paper, we presented a methodology for web pages categorization by means of semantic graphs and machine learning techniques. The semantic graph allows determining a map of the semantic areas contained in the document and their relationships (i.e., semantic metric) w.r.t. a particular concept or term. The document categorization is accomplished by means of a supervised or unsupervised machine learning technique, Support Vector Machine (SVM) and Self Organizing Maps (SOM), respectively. The model that uses semantic features and trained on the Reuters corpus obtains better results for both SVM and SOM. For the other corpora, the best results are obtained by SOM and semantic weights. We can consider that a category may present concepts strongly correlated with the other categories and this behavior can be better managed by an unsupervised mechanism. We, however, wish to highlight that by using semantic weights a Web Page can also not contain a specific term but it contains correlated concepts. In the next future the authors will focus their attention on the use of different parameters for the machine learning techniques, different semantic metrics and on the categorization of documents also using images, audio and video.

References

1. Bloehdorn, S., Hotho, A.: Boosting for text classification with semantic features. In: Mobasher, B., Nasraoui, O., Liu, B., Masand, B. (eds.) WebKDD 2004. LNCS (LNAI), vol. 3932, pp. 149–166. Springer, Heidelberg (2006)
2. Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning 20(3), 273–297 (1995)
3. Divya, C.: Mining Contents in Web Pages and Ranking of Web Pages Using Cosine Similarity. International Journal of Science and Research (IJSR) 3(4) (2014)
4. Kohonen, T.: The self-organizing map. Proceedings of the IEEE 78(9), 1464–1480 (1990)
5. Lin, D.: An information-theoretic definition of similarity. In: Proceedings of the 15th International Conference on Machine Learning, San Francisco, vol. 1, pp. 296–304 (1998)
6. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press (2008)
7. Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.J.: Introduction to wordnet: An on-line lexical database. International Journal of Lexicography 3(4), 235–244 (1990)
8. Placitelli, A.P.: Categorizzazione di pagine web mediante grafo semantico e tecniche di machine learning, MSc dissertion, University of Naples "Parthenope" (2013)
9. Qi, X., Davison, B.D.: Web Page classification: Features and algorithms. ACM Computing Surveys (CSUR) 41(2), 12 (2009)
10. http://www.daviddlewis.com/resources/testcollections/reuters21578/
11. Salton, G., Buckley, C.: Term-weighting approaches in automatic text retrieval. In: Information Processing and Management, pp. 513–523 (1988)
12. Trstenjaka, B., Mikacb, S., Donkoc, D.: KNN with TF-IDF based Framework for Text Categorization. Procedia Engineering 69, 1356–1364 (2014)

Web Spam Detection Using Transductive–Inductive Graph Neural Networks

Anas Belahcen[1,2], Monica Bianchini[1], and Franco Scarselli[1]

[1] Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche
Università degli Studi di Siena – Siena, Italy
[2] LeRMA – ENSIAS
Mohammed V Souissi University – Rabat, Morocco
`{monica,franco}@diism.unisi.it,`
`belahcen.anas@gmail.com`

Abstract. The Web spam detection problem has received a growing interest in the last few years, since it has a considerable impact on search engine reputations, being fundamental for the increase or the deterioration of the quality of their results. As a matter of fact, the World Wide Web is naturally represented as a graph, where nodes correspond to Web pages and edges stand for hyperlinks. In this paper, we address the Web spam detection problem by using the GNN architecture, a supervised neural network model capable of solving classification and regression problems on graphical domains. Interestingly, a GNN can act as a mixed transductive–inductive model that, during the test phase, is able to classify pages by using both the explicit memory of the classes assigned to the training examples, and the information stored in the network parameters. In this paper, this property of GNNs is evaluated on a well–known benchmark for Web spam detection, the WEBSPAM–UK2006 dataset. The obtained results are comparable to the state–of–the–art on this dataset. Moreover, the experiments show that performances of both the standard and the transductive–inductive GNNs are very similar, whereas the computation time required by the latter is significantly shorter.

1   Introduction

In several application areas, data are naturally represented as graphs or trees, e.g., in computer vision, molecular biology, software engineering and natural language processing. As a matter of fact, nodes in these structures are used to represent objects, while edges determine the relationships between them. For example, the World Wide Web is commonly described by a graph, where nodes represent Web pages and edges stand for hyperlinks. In the Web graph, nodes and edges may have vector labels, collecting the information available about the page contents and the hyperlinks, respectively.

Traditional machine learning approaches try to reduce graphical data into simple representations, as, e.g., a set of vectors. In this way, the topological information may

be lost during the preprocessing step, which can deeply affect the achieved performance. On the other hand, the Graph Neural Network (GNN) model [1] is capable of processing graphs directly, without any preprocessing step. GNNs are supervised neural network models that extend the recursive paradigm, and can be applied on most of the practically useful kinds of graphs, including directed, undirected, labeled and cyclic graphs. GNNs have been successfully employed in several application domains, such as molecule classification, object localization in images, and Web page ranking.

In this paper, we apply GNNs to Web spam detection, i.e., the problem of classifying a Web page as a document containing spam or not. Such a problem has received a growing interest in the last few years, due to its importance for search engines [2–4]. In Web spam detection, the Web graph can be used both for learning and testing. During training, we use a small set of pages, for which the target is known, to learn the GNN parameters. Then, the trained GNN is applied on the whole Web graph, to classify the remaining Web pages. GNNs are well suited for Web spam detection, since they can learn to automatically classify pages, exploiting both the information available on the page contents and on the Web connectivity.

Interestingly, a GNN can operate using two modalities: it can act as a pure inductive model or as a mixed transductive–inductive model. In the former case, during the test phase, the GNN completely relies on its parameters to classify Web pages. The actual classification of a training page is not explicitly memorized, so that such page can be even misclassified by the GNN. With the transductive–inductive modality, the classification of the training pages is explicitly added to the Web graph. In this way, the GNN operates also as a transductive model, that classifies test pages by using the information already available for the training pages, and by diffusing such an information through the Web graph.

In order to evaluate our approach, we tested GNNs on a well–known benchmark for Web spam detection, the WEBSPAM–UK2006 dataset [5], finding promising preliminary results.

2   The Graph Neural Network Model

Graph Neural Networks (GNNs) are a supervised connectionist model capable of solving classification and regression problems on graphical domains. One of the major advantages of GNNs is their capacity of processing graphs directly (without preprocessing), which preserves the information collected into the graph topology.

In fact, graph nodes are used to represent concepts, while edges determine the relationships between them. Each concept or node in the graph is defined by its features, and also by the information contained in its neighborhood. Based on these two information sources, the GNN calculates a state x_n, for each node n, which contains the node representation (see Fig. 1). Then, using this state, the GNN produces an output that denotes the classification decision on that node. Formally, the output of a GNN is defined by the following equations:

$$x_n = f_w\big(l_n, l_{co[n]}, x_{ne[n]}, l_{ne[n]}\big), \tag{1}$$
$$o_n = g_w(x_n, l_n),$$

where f_w and g_w are parametric functions, implemented by two feedforward neural networks, which express the dependence of the state at each node on the state of its neighborhood, and the dependence of the node output on its state, respectively.
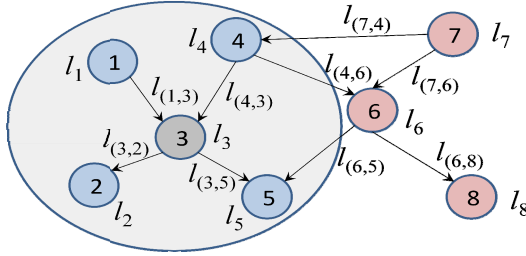
Fig. 1. A graph and, in evidence, the neighborhood of a node. The state x_3 of node 3 depends on the information contained in its neighborhood. The transition and the output functions are, respectively defined as $x_3 = f_w\big(l_3, l_{(1,3)}, l_{(3,2)}, l_{(4,3)}, l_{(3,5)}, x_1, x_2, x_4, x_5, l_1, l_2, l_4, l_5\big)$, $o_3 = g_w(x_3, l_3)$

Moreover, $l_n, l_{co[n]}, x_{ne[n]}, l_{ne[n]}$ represent the label of n, the labels of its attached edges, and the states and the labels of the nodes in its neighborhood, respectively.

In order to compute the output defined by Eq. (1), the Banach Fixed Point Theorem suggests the following classic iterative scheme:

$$x_n(t + 1) = f_w\big(l_n, l_{co[n]}, x_{ne[n]}(t), l_{ne[n]}\big), \tag{2}$$
$$o_n(t) = g_w(x_n(t), l_n),$$

for each node n. Intuitively, the computation described by Eq. (2) can be interpreted as the activity of a network consisting of units which compute f_w and g_w. Such a network, built by replacing each node of the graph with a unit computing f_w (see Fig. 2), will be called the encoding network. Each unit stores the current state $x_n(t)$ at node n and, when activated, it calculates $x_n(t + 1)$ (Fig. 2). The output at n is produced by another unit which implements g_w.
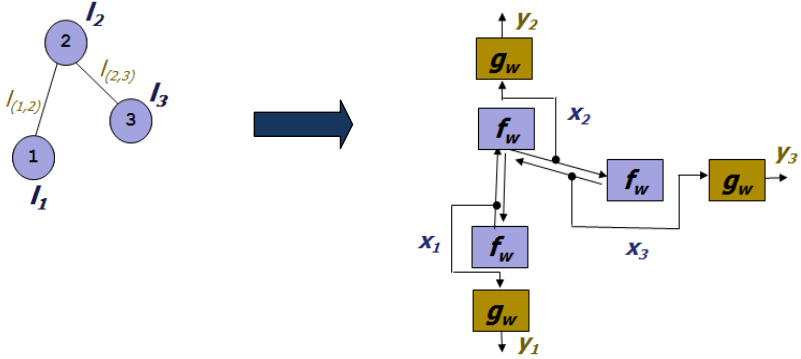
Fig. 2. The graph (on the left) and the corresponding encoding network (on the right). Graph nodes (circles) are replaced by ad hoc units computing f_w and g_w (squares). When f_w and g_w are implemented by feedforward neural networks, the encoding network is a recurrent network.

More details on the GNN training algorithm and output computation can be found in [1]. Here, it suffices to say that both training and test sets consist of a labelled graph which, in our application, is a portion of the Web graph. For the training set, also targets for some nodes are provided, which define the actual class of these nodes, i.e. whether corresponding pages are spam or not. The training procedure adapts the network parameters in order to produce the correct outputs on the supervised pages, while the test procedure uses the trained GNN to classify the remaining pages.

As mentioned in Section 1, GNNs can be exploited either as a common parameterized inductive model or as a mixed transductive−inductive model. In the inductive setting, the network is fed with the Web graph, using supervised pages to adapt the GNN parameters. Hence, in this way, the information contained in the training set is used to approximate a classification function that can be used to directly classify the nodes of the Web graph. On the other hand, in the transductive−inductive model, during training, a subset of the supervised pages is assigned a label enriched with their class membership, whereas the remaining (the class membership label is unset) are used for training − i.e. they contribute to the calculus/optimization of the error function. Instead, during testing, a component of the label of each training page explicitly specifies whether such a page is spam or not (the class membership label is unset for unsupervised pages), so that the information available on the training pages *is directly diffused* through the Web graph.

3   The WEBSPAM−UK2006 Dataset

In order to assess our approach, we evaluate the GNN model on the WEBSPAM−UK2006 dataset. Actually, the dataset was adopted in 2007 by the Web Spam Challenge, a competition held annually during the International Workshop on

Adversarial Information Retrieval on the Web. The Web graph is a crawl of the .uk domain that includes 77.9 million pages and over 3 billion links in 11,402 hosts. The labeling was at the host level, i.e., the assessors labeled the hosts as normal or spam. Such a benchmark is particularly suited for our purpose both because it has been used by several research groups and because it is sufficiently large to produce significant results and, at the same time, not too huge to prevent a wide experimentation.

3.1   Features

Data are represented by the following features: (1) link−based features, which include, f.i., the indegree and the outdegree of hosts and their neighbors, PageRank and TrustRank; (2) content−based features, which include, f.i., the fraction of anchor and visible text, the compression rate, the corpus precision (the fraction of words in a page that belong to the set of popular terms), and the corpus recall (the fraction of popular terms that appear in the page).

3.2   Feature Preprocessing

The WEBSPAM−UK2006 dataset includes 41 link−based and 96 content−based features, which are used as node labels. Due to the high number of features, we use a feedforward neural network in order to summarize and compress them into a single one. More precisely, different configurations were used for GNNs, as it follows.

- **Link and content−based features directly**: The most significant link and content−based features are selected, using a correlation−based feature selector [6], and integrated as the node label.
- **Link−based feature**: A feedforward neural network is employed in order to compress all the link−based features into a single output. This output is then used as the node label.
- **Content−based feature**: A feedforward neural network is employed in order to compress all the content−based features into a single output. This output is then used as the node label.
- **Compressed and uncompressed features:** Link and content−based features, already compressed by feedforward networks, in addition to some features directly selected (in particular, the PageRank and the TrustRank of the host and of its maximum scored page) are collected together and then used as the node label.

3.3   Teams Participating to the 2007 Web Spam Challenge

We compare our results with those gained by the six teams participating to the 2007 Web Spam Challenge. The competition attracted three teams from academic institutions (Hungarian Academy of Sciences, University of Waterloo, and Chinese Academy of Sciences) and three teams from industry research laboratories (Genie

Knows, Microsoft, and France Télécom). Results obtained by competing teams are shown in Table 3. Notice that, after the competition, other groups have worked on these benchmarks, but their results are difficult to be comparatively evaluated, due to the fact that, in most cases, the original splitting between training and test sets has not been used.

4   Experimental Results

In this section, we present the experimental results obtained by GNNs. The experiments are divided into two parts. The first one uses the original training/test splitting, already fixed in the challenge, whereas in the second one the splitting is randomly constructed from all the dataset pages. The choice of testing the approach on a random splitting is a common procedure and it is motivated by the presence of different the data distributions in the original training and test sets. Besides, for each splitting, an inductive learning model and a mixed transductive−inductive learning approach were used. As mentioned before, in the transductive−inductive setting, the training pages were divided into two groups, in order to define the error function and to simulate the transductive inference, respectively. In our experiments, two equal−size groups were randomly defined. Finally, the performance was measured by the area under the ROC curve, the F−measure, and the accuracy.

4.1   The Random Splitting

The WEBSPAM−UK2006 dataset was randomly split into training (2228 hosts), validation (1000 hosts), and test (2518 hosts) sets. The results are divided according to whether GNNs are used as an inductive or a mixed transductive−inductive learning model. Table 1 shows the performance obtained by different GNN configurations.

Each row represents a different simulation, as described below:
- In the first experiment, the most significant link and content−based features were chosen, using a correlation−based feature selector [6]. In fact, according to a preliminary experiment, ten link−based and two content−based features were selected. The performance was the lowest compared to the other configurations.
- In the second and in the third experiment, a feedforward neural network was used, in order to compress all the features into a single output (each type of features has its own output). Exploiting this idea, the performance of the model increases.
- In the last configuration, we combine link and content−based features, already compressed by feedforward networks, with directly selected features (i.e., PageRank and TrustRank of the host and of its maximum PageRank page). With this experiment, we obtain the highest performance.

For most of the experiments, the results obtained in the transductive−inductive learning framework are slightly better. In Table 3, we compare the results of our best GNN configuration with those produced by the other teams, proving that it gains very similar performance to the winner.

Table 1. Performances of different GNN configurations with random splitting

	Accuracy		F-Measure		ROC	
Configurations	Transd.-Ind.	Induc.	Transd.-Ind.	Induc.	Transd.-Ind..	Induc.
Link and content directly	89,48%	89,27%	0.7550	0.7528	0.9467	0.9446
Link based (FFNN)	90,30%	90,27%	0.7680	0.7763	0.9506	0.9516
Content based (FFNN)	90,50%	90,11%	0.7532	0.7531	0.9417	0.9387
Link and content (FFNN) and directly selected	**94,08%**	**93,96%**	**0.8534**	**0.8499**	**0.9681**	**0.9717**

4.2   The Original Splitting

The splitting adopted in the Challenge was also used for the experiments. In this case, the training set includes 8415 hosts (7472 normal, 767 spam, and 175 undecided), while the test set contains 2247 hosts (651 normal, 1346 spam, and 250 undecided). The architecture used to address this classification problem is shown in Fig. 3.
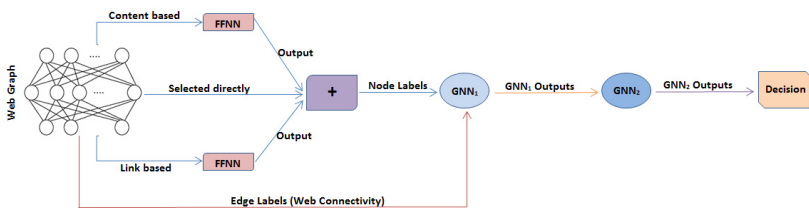
Fig. 3. The configuration adopted in the original splitting.

In this case, content−based and link−based features compressed into single features (by feedforward networks) are used, in addition to features directly selected, to construct the whole labels for the GNN processing. The produced output will be used in a second GNN. The output of the second GNN will be the decision on the hosts, classified as spam or normal. As in the random splitting experiments, this GNN configuration can be used as an inductive or a mixed transductive−inductive model. The obtained results are shown in Table 2.

Table 2. Performance comparison with the original splitting

	Accuracy		F-Measure		ROC	
Configurations	Transd.	Induc.	Transd.	Induc.	Transd.	Induc.
Link and content (FFNN) and directly selected	**89,53%**	**89,23%**	**0.9219**	**0.9215**	**0.9496**	**0.9502**

Based on the experiments, we observe that the standard inductive and the transductive−inductive models are comparable in terms of performance.

Nevertheless, some more experiments are worth carrying out in order to clearly establish the GNN ability in addressing the proposed problem. In the following, we compare the performance obtained in our experiments with respect to those of the other competing teams, and also evaluate training times of both the transductive−inductive and the inductive models, with respect to the random and the original splitting.

4.3   Performance Comparison

The experiments, based on both original and random splitting, show that our results are comparable to the best results obtained so far on the WEBSPAM−UK2006.

Table 3. Comparative results

Participants	F1	ROC
Abou *et al.* (Genie Knows)	0.81	0.80
Benczur *et al.* (Hungarian Academy of Sciences)	0.91	0.93
Cormack (University of Waterloo)	0.67	0.96
Fetterly *et al.* (Microsoft)	0.79	-
Filoche *et al.* (France Télécom)	0.88	0.93
Geng *et al.* (Chinese Academy of Sciences)	0.87	0.93
Random splitting	**0.85**	**0.97**
Original splitting	**0.92**	**0.95**

4.4   Training Time Comparison

Even if the two learning frameworks show comparable performances, they significantly differ in terms of training time (see Fig. 4).
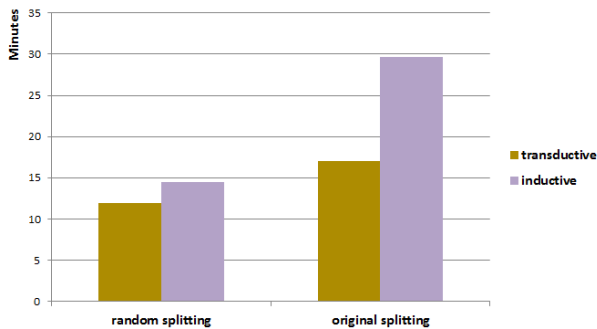
Fig. 4. Comparison between the transductive−inductive and the inductive configurations with respect to the training time

Actually, the training time for the inductive model is greater than that for the transductive–inductive approach, with the random splitting, of about 20%, while, with the original splitting, it increases of 76%, which means that the transductive–inductive model is as efficient as the inductive model, but it is certainly very less expensive from the computational point of view.

5   Conclusions

According to experiments conducted on the WEBSPAM–UK2006 dataset, the results obtained using the random and the original splitting can be compared, in terms of performance, to the state–of–the–art results. Besides, the experiments were realized based on both transductive–inductive and inductive frameworks, which show different training times, clearly assessing the advantages of the transductive–inductive approach from the computational point of view.

References

1. Scarselli, F., Gori, M., Tsoi, A.-C., Hagenbuchner, M., Monfardini, G.: The graph neural network model. IEEE Trans. on Neural Networks 20(1), 61–80 (2009)
2. Di Noi, L., Hagenbuchner, M., Scarselli, F., Tsoi, A.-C.: Web Spam Detection by Probability Mapping GraphSOMS and Graph Neural Networks. In: Diamantaras, K., Duch, W., Iliadis, L.S. (eds.) ICANN 2010, Part II. LNCS, vol. 6353, pp. 372–381. Springer, Heidelberg (2010)
3. Gyöngyi, Z., Garcia-Molina, H.: Web spam taxonomy. Adversarial information retrieval on the Web (2005)
4. Castillo, C., Donato, D., Becchetti, L., Boldi, P., Leonardi, S., Santini, M., Vigna, S.: A reference collection for Web spam. ACM SIGIR Forum 40(2), 11–24 (2006)
5. Web spam challenge (2007), `http://webspam.lip6.fr/wiki/pmwiki.php?n=Main.PhaseIResults`
6. Hall, M.A.: Correlation–based Feature Subset Selection for Machine Learning. Hamilton, New Zealand (1998)

Hubs and Communities Identification
in Dynamical Financial Networks

Hassan Mahmoud[1], Francesco Masulli[1,2], Marina Resta[3], Stefano Rovetta[1],
and Amr Abdulatif[1]

[1] DIBRIS, Dipartimento di Informatica, Bioingegneria, Robotica e Ingegneria dei
Sistemi, University of Genoa, via Dodecaneso 35, 16146, Genoa, Italy
[2] Center for Biotechnology, Temple University, Philadelphia, USA
[3] Dipartimento di Economia, University of Genoa, via Vivaldi 5, 16126, Genoa, Italy
{hassan.mahmoud, francesco.masulli, marina.resta,
stefano.rovetta}@unige.it, 3955680@studenti.unige.it

Abstract. In this study we aim at identifying companies influencing
the performance of the stock market sector. We propose an approach
for constructing the similarity between stock company profiles based on
the estimates of the log return similarity of stock prices and on Fuzzy
Spectral Modularity community detection method to infer the network
hubs and significant communities and we applied it to the Italian stock
market store. Experimental results show that companies in the same
sector highly affect the price change of each other. Moreover, We notice
a robust temporal stability of detected communities, and the short time
correlation computed with the fuzzy rand index is strong.

Keywords: Communities, Dynamical Financial networks, Correlation
networks, Spectral clustering, Fuzzy clustering, Stability.

1 Introduction

Community discovery is difficult task since communities are hidden behind com-
plicated relationships. Moreover, it is unclear how to extract coverage in real
world networks. Several attempts were devoted to characterize overlapping com-
munities [14], however none of them is able to efficiently infer the hidden struc-
ture of the network.

In financial stocks, in order to identify groups of assets highly affecting the
price of each other compared to the rest of the network (i.e., communities) several
approaches have been proposed able to identify the structure of the empirical
correlation matrix of asset profiles, and model their hierarchies in terms of hier-
archical trees and networks [13,15,19]. The process of communities identification
needs to make use of statistically reliable information, often the correlation ma-
trix is sensitive to several factors, such as the heterogeneity of sampling, the
interaction with environment, and the non-stationariety of data sources.

The Italian stock market network we will analyze in this work showed a
structure evolving through time progression in which there is a possibility for

some assets observations to be missed or for others to appear and affect the stock market. Finding a suitable community detection technique in complex real world networks is an open problem due to various challenges given by clustering approaches. Among them, there are the following: initialization criteria (e.g., choosing an initial number of clusters is required in partitional clustering like K-Means [8], while it is not needed in hierarchical clustering), accuracy (e.g., a main drawback of hierarchical clustering is the possible misclassification of some nodes [11], while removing edges may result in singleton clusters in graph bisection approach), stability (e.g., results may differ depending on the specific similarity measure used and on the random initialization of cluster centers in partitional clustering).

This paper is organized as follows: Sec. 2 illustrates the proposed method to identify financial communities and hubs using the Fuzzy c-means Spectral Modularity community detection method we proposed in [9]; Sec. 3 applies the fuzzy Rand index for measuring the stability of the overlapping evolutionary communities obtained in each time window (30 days); Sec. 4 shows the experimental results obtained on the Italian Stock Market data; Finally Sec. 5 shows the conclusions.

2   Financial Communities Identification

In this paper we consider correlation estimators based on the Euclidean distance between the log returns daily closure prices of assets pairs. After obtaining the asset similarities, we apply the Fuzzy c-means Spectral Modularity *(FSM)* community detection method we proposed in [9] to infer the overlapping asset communities, and identify hub assets. Finally, we measured the strength of our proposed *FSM* in quantifying the Italian stock market network using fuzzy Rand index [18].

The approach we propose for analyzing how assets interact in financial networks includes the following steps (also summarised in Fig.1):

1. Normalize the data.
2. Define the initial time t_0 and the length l of the temporal window.
3. Within each temporal window measure the Log–Return Similarity *(LRS)* of stock prices at closure between each pair of following days defined as:

$$r_h(t) = \ln \frac{p_h(t)}{p_h(t-1)} = \ln p_h(t) - \ln p_h(t-1), \qquad (1)$$

   where $p_h(t)$ refers to closure price of asset h at time t.
4. Calculate the pairwise Euclidean distance d_{hi} between the profiles of assets h and i (temporal windows) for each pair of assets (excluding the cases with missing observations, and replacing undefined values with zero in each asset, hence the cardinality is always the same):

$$d_{hi} = \sqrt{\sum_{p=0}^{l-1} (r_h(t_p) - r_i(t_p))^2}. \qquad (2)$$

5. Estimate the similarity between asset profiles [16] as:

$$sim_{hk} = \exp\left(\frac{-d_{hk}}{s}\right), \tag{3}$$

where s is the dispersion, it depends on the data distribution. We estimated s experimentally using histogram analysis instead of choosing it randomly.
6. Apply the FSM-community detection method to the asset profile similarities matrix $\Sigma = [sim_{hk}]$ to infer overlapping financial communities.

The art of identifying nodes having more influence over the network structure than others is referred to as *node centrality* study. A vertex with high centrality (hub) implies that it lies on considerable fractions of shortest paths connecting vertexes. As a consequence, various *centrality measures* have been developed in network analysis: among them the most known (and used) are: centrality degree, closeness, betweenness, and modularity [2,7].

Network modularity is used for measuring the strength of community structure in networks. High network modularity implies the existence of dense connections within communities, and of sparse links between them. Although modularity shows a resolution limit specially in case of detecting small communities, it has the advantages of not requiring prior knowledge about the number or sizes of communities, and it is capable of discovering network partitions composed of communities having different sizes. Moreover, degree, and closeness are local measures which limit their efficiency in case of evolving networks [10,11].

The ***Fuzzy c-means Spectral Modularity* (FSM) - *community detection method*** [9], derived by the Ng et al. [12] spectral clustering algorithm [20,5,3]. The main improvement introduced in the *FSM* is the application of the Fuzzy c-means *(FCM)* algorithm [1,4] for clustering in the affine subspace spanned by the first k^* eigenvectors. The *FCM* allows an instance to belong to two or more clusters at the same time, with possibly different membership degrees. This feature supports the detection of overlapping communities and can allow to understand the role that each node may play in different communities.

3   Measuring Evolving Communities Stability Using Fuzzy Rand Index

To measure the stability in our experiments, we used the fuzzy Rand index as defined in [6], in that paper the Rand index is viewed as distance function $D_{RI} = 1 - RI$.

$$a = (1 - |u - v|) \times u \times v; \tag{4}$$

$$b = (1 - |u - v|) \times (1 - u \times v); \tag{5}$$

$$c = max((u - v), 0); \tag{6}$$

$$d = max((v - u), 0). \tag{7}$$
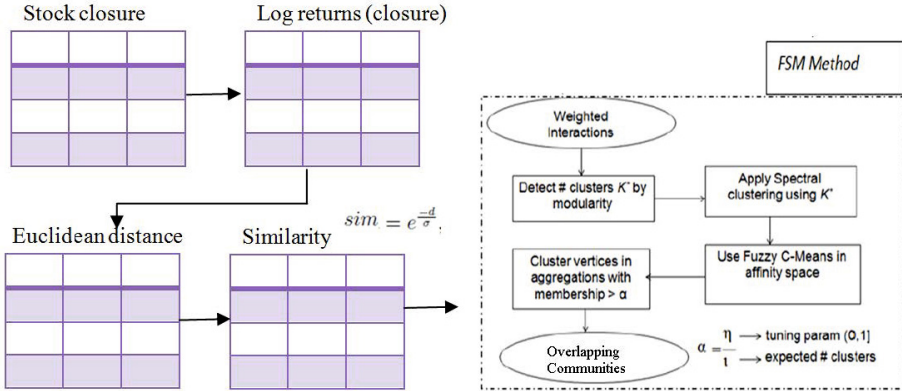
Fig. 1. The Log-Return Similarity approach

where $u = E_{P_1}(x, x^{'})$, $v = E_{P_2}(x, x^{'})$ are a fuzzy equivalence relation on X in terms of a similarity measure on the associated membership vectors $P(X) = P_1(X), P_2(X), ...P_k(X) \in [0,1]^k$ given by:

$$E_{P_1}(x, x^{'}) = 1- \parallel P(x) - P(x^{'}) \parallel, \tag{8}$$

where $\parallel . \parallel$ is a proper distance on $[0,1]^k$

The distance measure on fuzzy partitions is then defined as the normalized sum of degrees of discordance:

$$D(P_1, P_2) = \frac{\sum_{(x,x')\in C} |E_{P_1}(x, x^{'}) - E_{P_2}(x, x^{'})|}{n(n-1)/2}) \tag{9}$$

Hence, the fuzzy Rand index is given by:

$$RI_f = 1 - D(P_1, P_2) \tag{10}$$

4   Experimental Study

The approach described in previous section was applied to the closure prices of the Italian Stock Exchange observed in the period from 15 March 2004 to 15 March 2014. Our dataset contains 171 assets actively traded on Milan Stock Exchange (MSE), in addition to 13 evolving assets depicted in Fig. 2, classified into 37 categories.

In order to discuss the result we have obtained, we now focus on the first time window we analyzed, consisting of a 30 days window (as done in [13]) from 15 March 2004 to 15 June 2004. The software in use was developed in Matlab R2009b[C] under Windows 7[C] 32 bits. The experiments were performed on a laptop with 2.00 GHz dual-core processor and 3.25 GB of RAM. We used a time

Table 1. Network characteristics of Italian Stock market between 15/3/2004 and 15/3/2005

	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.
Assets	171	171	174	174	174	176	178	179	179	179	180	182
Edge	29241	29241	30276	30276	30276	30976	31684	32041	32041	32041	32400	33124
Avg. path	0.043	0.012	0.013	0.003	0.088	0.067	0.023	0.027	0.128	0.142	0.178	0.054
Modularity	0.384	0.384	0.469	0.479	0.434	0.425	0.404	0.389	0.384	0.411	0.407	0.416
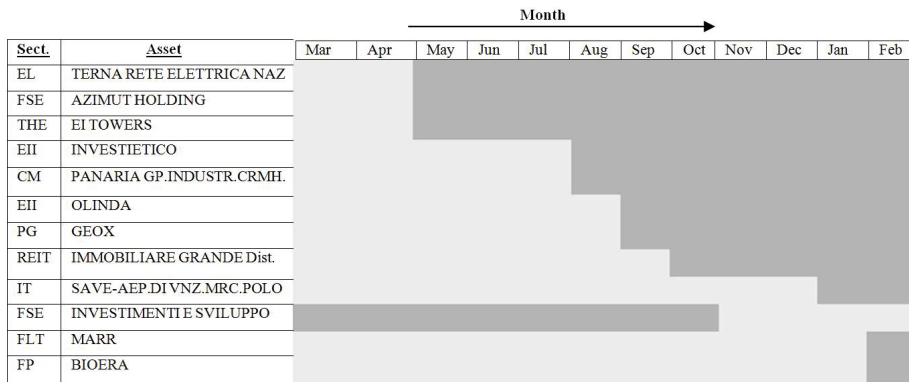
Fig. 2. Monthly stock market evolution through one year, light cell indicates that this asset observations are missed during this month, while dark cells refers to new assets involved in market.

window of 30 days, a fuzziness parameter $m = 2$, and a dispersion parameter $s = .5$ experimentally evaluated.

We used modularity maximization to obtain the number of clusters. We report the modularity values of the communities identified in Tab. 1, with together the number of assets, the number of edges, and the average path length *(APL)* in each time timestamp (30 days). The network *(APL)* is the average number of steps along the shortest paths for all possible pairs of network nodes. and for assets x, and $x^{'}$ having distance $d(x, x^{'})$ in a financial network with n assets is given by $\frac{2}{n(n-1)} \sum d(x, x^{'})$, s.t., $x \neq x^{'}$. Moreover, we notice that the *(APL)* is considered as a measure of the efficiency of information transport on the network.

Fig. 3 lists the composition of the five communities obtained by running the LRS approach, and shows the obtained hubs (in bold) and bridge (denoted as "fuzzy") assets belonging to more than one community. Note that assets belonging to the same category show higher tendency to be grouped together. Moreover, we found some sectors affecting others, such as electricity (EL), industrial engineering (IE), and Industrial transportation (IT) sectors, as depicted in Fig. 4 that shows a detailed three-dimensional representation of the number of assets for each category to the five communities.

To study the stability of obtained communities, we constructed the contingency matrix containing the fuzzy Rand index RI_f between each pair of times-

C1

Asset Name	sector	Fuzzy
'A2A'	'EL'	
'ENEL'	'EL'	
'FALCK RENEWABLES'	'EL'	
'SOGEFI'	'AUP'	
'DAVIDE CAMPARI MILANO'	'BEV'	
'AMUNDI RE EUROPA'	'EII'	
'AMUNDI RE ITALIA'	'EII'	
'BANCA CARIGE'	'BKS'	
'BANCA PPO.ETRURIA LAZIO'	'BKS'	
'BANCA PPO.DI SONDRIO'	'BKS'	
'BANCA PPO.EMILIA ROMAGNA'	'BKS'	
'BCA.PICCOLO COT.VALTELL'	'BKS'	
'ASTALDI'	'CM'	
'CALTAGIRONE'	'CM'	
'ITALCEMENTI FABBRICHE RIUNITE'	'CM'	
'CEMBRE'	'EEE'	
'EL EN'	'EEE'	
'IRCE'	'EEE'	
'DEA CAPITAL'	'FSE'	
'BONIFICHE FERRARESI'	'FP'	
'SNAM'	'GWM'	
'ATLANTIA'	'HGHC'	
'ASTM'	'IT'	
'SIAS'	'IT'	
'CAIRO COMMUNICATION'	'MTEL'	
'MEDIACONTECH'	'MED'	1,5,2
'UNIPOL GRUPPO FINANZIARI'	'NLI'	
'VITTORIA ASSICURAZIONI'	'NLI'	
'ENI'	'OG'	
'ERG'	'OG'	
'CIA'	'REIS'	
'NOVA RE'	'REIS'	
'RISANAMENTO'	'REIS'	
'BENI STABILI'	'REIS'	
'AUTOGRILL'	'TLE'	1,4,5
'SS LAZIO'	'TLE'	

C2

Asset Name	sector	Fuzzy
'ACEA'	'EL'	
'FRENI BREMBO'	'AUP'	
'FINMECCANICA'	'AED'	
'KINEXIA'	'AEN'	
'BANCA MONTE DEI PASCHI'	'BKS'	
'BANCA POPOLARE DI MILANO'	'BKS'	
'BANCA PROFILO'	'BKS'	
'BANCA DI SARDEGNA RSP'	'BKS'	
'BNC.DI DESIO E DELB.'	'BKS'	
'CREDITO EMILIANO'	'BKS'	
'MEDIOBANCA BC.FIN'	'BKS'	
'UNIONE DI BANCHE ITALIAN'	'BKS'	
'UNICREDIT'	'BKS'	
'BUZZI UNICEM'	'CM'	
'CEMENTIR HOLDING'	'CM'	
'DATALOGIC'	'EEE'	
'PRIMA INDUSTRIE'	'EEE'	
'MITTEL'	'EEE'	
'CENTRALE DEL LATTE DI TRO.'	'FSE'	
'IREN'	'FP'	
'CENTRO HL DISTRIBUZIONE'	'GWM'	2,3
'AMPLIFON'	'GR'	
'DE LONGHI'	'HCES'	
'VINCENZO ZUCCHI'	'HGHC'	
'AEROPORTO DI FIRENZE'	'HGHC'	
'CATTOLICA ASSICURAZIONI'	'IT'	
'CALTAGIRONE EDITORE'	'LINS'	
'MONDO TV'	'MED'	
'POLIGRAFICI EDITORIALE'	'MED'	
'TELECOM ITALIA MEDIA'	'MED'	
'UNIPOLSAI'	'MED'	
'ASSICURAZIONI GENERALI'	'NLI'	
'GIOVANNI CRESPI'	'NLI'	
'CSP INTERNATIONAL'	'PG'	
'BRIOSCHI SVILUPPO IMMBL'	'PG'	
'CAD IT'	'REIS'	2,1
'DADA'	'SCS'	
'ENGR.INGEGNERIA INFORMA'	'SCS'	
'TXT E-SOLUTION'	'SCS'	
'ESPRINET'	'THE'	
'FNM'	'TLE'	
'GRANDI VIAGGI'	'TLE'	
'JUVENTUS FOOTBALL CLUB'	'TLE'	
'SNAI'	'TLE'	

C3

Asset Name	sector	Fuzzy
'ISAGRO'	'CHE'	2,1
'SOL'	'CHE'	
'CREDITO BERGAMASCO'	'BKS'	
'BEGHELLI'	'EEE'	
'GEFRAN'	'EEE'	
'BANCA INTERMOBILIARE'	'FSE'	
'ACQUE POTABILI'	'GWM'	
'ACSM-AGAM'	'GWM'	
'HERA'	'GWM'	
'CICCOLELLA'	'GR'	3,5
'BORGOSESIA RSP'	'IE'	3,1
'MONRIF'	'MED'	
'SEAT PAGINE GIALLE'	'MED'	3,2,5
'SAIPEM'	'MM'	
'BASICNET'	'PG'	
'RATTI'	'PG'	
'RECORDATI INDUA.CHIMICA'	'PHB'	3,1
'GABETTI PROPERTY SLTN.'	'REIS'	
'BASTOGI'	'SCS'	
'FIERA MILANO'	'SSER'	
'POLIGRAFICA S F'	'SSER'	
'AMBIENTHESIS'	'SSER'	
'OLIDATA'	'THE'	
'AS ROMA'	'TLE'	
'GTECH'	'TLE'	3,2

C4

Asset Name	sector	Fuzzy
'ALERION CLEAN POWER'	'EL'	4,1
'K R ENERGY'	'EL'	4,1,3
'IMMSI'	'AUP'	
'PININFARINA'	'AUP'	
'BOERO BARTOLOMEO'	'CM'	
'GRUPPO CERAMICHE RICCHET'	'CM'	
'TREVI FIN INDUSTRIALE'	'CM'	
'VIANINI INDR.'	'CM'	
'VIANINI LAVORI'	'CM'	
'BANCA IFIS'	'FSE'	
'LVENTURE GROUP'	'FSE'	
'INVESTIMENTI E SVILUPPO'	'FSE'	
'LA DORIA'	'FP'	
'RENO DE MEDICI'	'GI'	
'SORIN'	'HCES'	
'BIESSE'	'IE'	
'FIDIA'	'IE'	
'AUTOSTRADE MERIDIONALI'	'IT'	
'FEDON (PAR)'	'PG'	
'AEDES LIGURE LOMBARDA'	'PHB'	
'BEE TEAM'	'REIT'	
'EXPRIVIA'	'SCS'	
'FULLSIX'	'SCS'	
'SINTESI SOCIETA DI INVMI'	'SSER'	

C5

Asset Name	sector	Fuzzy
'FIAT'	'AUP'	
'PIRELLI'	'AUP'	
'BANCA FINNAT EURAMERICA'	'BKS'	
'BANCA PPO.DI SPOLETO'	'BKS'	
'BANCO POPOLARE'	'BKS'	
'INTESA SANPAOLO'	'BKS'	
'IMPREGILO'	'CM'	
'ITALMOBILIARE'	'CM'	
'SAES GETTERS'	'EEE'	
'RETI TELEMATICHE ITALIAN'	'FSE'	
'TELECOM ITALIA'	'FLT'	
'ARENA'	'FDR'	
'EDISON RSP'	'GWM'	5,2
'CIR CIE.INDI.RIUN.'	'GI'	
'COFIDE GRUPPO DE BENEDET'	'GI'	
'DMAIL GROUP'	'GR'	
'INDESIT COMPANY'	'HGHC'	
'INTERPUMP GROUP'	'IE'	
'INTEK GROUP'	'IE'	
'PREMUDA'	'IT'	
'DIGITAL BROS'	'LG'	5,1
'MEDIOLANUM'	'LINS'	
'ACOTEL GROUP'	'MTEL'	
'CLASS EDITORI'	'MED'	
'GRUPPO EDIT.L"ESPRESSO'	'MED'	
'MEDIASET'	'MED'	
'ARNOLDO MONDADORI EDI.'	'MED'	
'RCS MEDIA GROUP'	'MED'	
'INTEK GROUP'	'MED'	
'LUXOTTICA'	'PG'	
'STEFANEL'	'PG'	
'TOD"S'	'PG'	
'PRELIOS'	'REIS'	5,1
'REPLY'	'SCS'	
'TAS TGA.AVANZATA SISTEMI'	'SCS'	
'TISCALI'	'SCS'	
'IT WAY'	'THE'	

Fig. 3. Communities identified using *FSM* over a sample 30 days long time window. Hubs are labeled by bold, while for fuzzy assets the indices of their overlapping communities are listed.
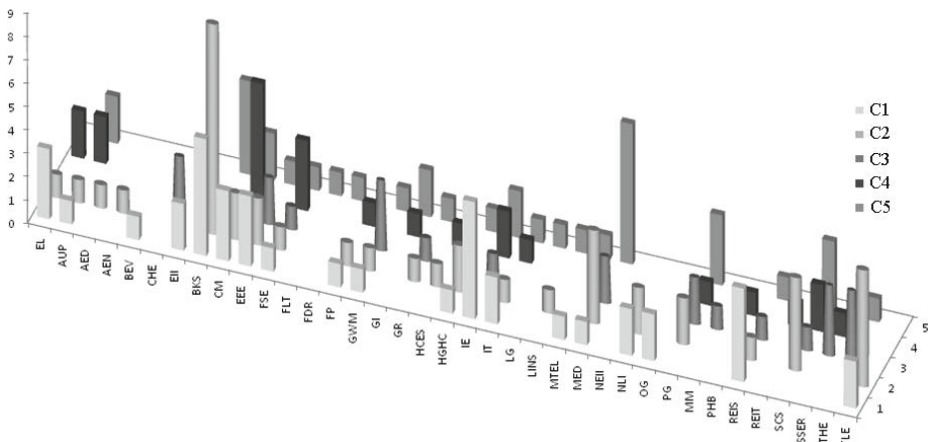
Fig. 4. Distribution of Italian stock 37 sectors in 5 clusters, as resulting from the *LRS* method over a sample 30 days long time window.

tamps (30 days) (see Fig. 5). The RI_f shows small variance all over the network timestamps. Moreover, the variance is proportional to the network temporal evolution. In facts, generally the closer the time stamps are, the higher the fuzzy Rand index between the communities we identified. This is due to having smaller change in the operating assets during adjacent months compared to far months, hence the communities detected are similar.

The visualization procedure works as follows:

1. Perform repeated *FSM* for detecting communities in each time stamp (t) (e.g., month).
2. Evaluate the similarity measure (or its average, in the case of multiple starts) of each pair of detected communities.
3. Arrange the similarity values in a $N_{steps} \times N_{steps}$ similarity matrix.
4. Convert the similarity matrix to a heat map image and visually analyze it (see Fig. 5).
5. Rank the columns or rows index corresponding to maximum stability.
6. Retrieve the values corresponding to these indices, corresponding to most correlated market behaviors at these time stamps to that t.

We highlight that our approach is general and other indices may be used (e.g., the fuzzy Jaccard index, as we used in [17]).

It is worth to note that our experimental results confirm as those observed in [19] and [13], but are more robust to network evolution and missing observations because they are not sensitive to noise artifacts, and moreover, support overlapping communities.
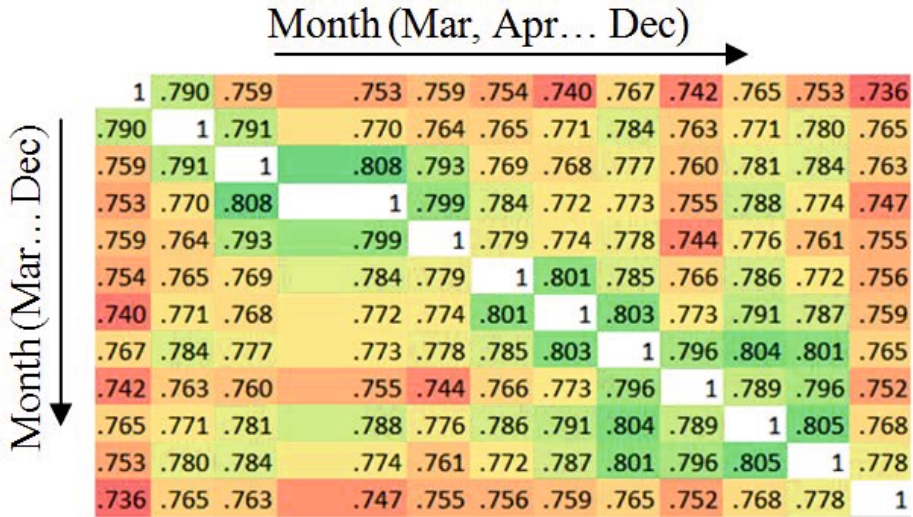
Fig. 5. Heat map of fuzzy Rand contingency matrix between *FSM* dynamic assets monthly memberships during one year.

5   Conclusions

This paper proposes a novel approach to financial stocks analysis, and to infer significant communities on them based on graph theory and on fuzzy spectral clustering, that we applied to the Italian stock market data.

To obtain the asset similarities we used a window of 30 daily measures of similarities based on the temporal profile distances of the log return of stock prices (LRS approach). We then applied the Fuzzy c-means Spectral Modularity community detection method *FSM* proposed in [9] to infer the overlapping asset graph, hence identifying hub assets.

Our approach could infer robust communities regardless of the temporal variant structure of the analyzed financial market interaction changes due to possibility that some assets disappear and new assets appear during temporal progression. We noticed high correlation between adjacent months having few changes in network structure than those with larger time gap due to the analyzed network dynamics.

The experimental results on the Italian stock market that led us to report five significant communities, and to notice that assets from the same category mostly affect each other and hence have high tendency to be grouped together. Moreover, we notice that some sectors deeply affect each others (namely, electricity, industrial engineering, and industrial transportation sectors).

Next step of this work includes the study of our procedure by replacing the correlation estimators proposed in this paper, with other correlation estimators, such as Fourier estimator, maximum likelihood correlation estimator, and dynamical estimator.

References

1. Bezdek, J.C.: Pattern recognition with fuzzy objective function algorithms. Kluwer Academic Publishers, Norwell (1981)
2. Brandes, U.: A faster algorithm for betweenness centrality. Journal of Mathematical Sociology 25, 163–177 (2001)
3. Donath, W.E., Hoffman, A.J.: Lower bounds for the partitioning of graphs. IBM Journal of Research and Development 17(5964), 420–425 (1973)
4. Dunn, J.C.: Some recent investigations of a new fuzzy partitioning algorithm and its application to pattern classification problems, pp. 1–15 (1974)
5. Filippone, M., Camastra, F., Masulli, F., Rovetta, S.: A survey of kernel and spectral methods for clustering. Pattern Recognition 41, 176–190 (2008) ISSN: 0031–3203
6. Hüllermeier, E., Rifqi, M.: A Fuzzy Variant of the Rand Index for Comparing Clustering Structures. In: IFSA/EUSFLAT Conf., pp. 1294–1298 (2009)
7. Koschützki, D., Lehmann, K.A., Peeters, L., Richter, S., Tenfelde-Podehl, D., Zlotowski, O.: Centrality indices. In: Brandes, U., Erlebach, T. (eds.) Network Analysis. LNCS, vol. 3418, pp. 16–61. Springer, Heidelberg (2005)
8. Lloyd, S.P.: Least square quantization in PCM, Bell Telephone Laboratories. Murray Hill (1957); Reprinted in: IEEE Transactions on Information Theory 28(2), 129–137 (1982)
9. Mahmoud, H., Masulli, F., Rovetta, S., Russo, G.: Community detection in protein-protein interaction networks using spectral and graph approaches. In: Formenti, E., Tagliaferri, R., Wit, E. (eds.) CIBB 2013. LNCS (LNBI), vol. 8452, pp. 62–75. Springer, Heidelberg (2014)
10. Newman, M.E.J.: Detecting community structure in networks. The European Physical Journal B-Condensed Matter 38, 321–330 (2004)
11. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Physical Review E 69(2), 026113 (2004)
12. Ng, J., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. In: Proceedings of Neural Information Processing Systems, pp. 849–856 (2002)
13. Onnela, J.P., Kaski, K., Kertész, J.: Clustering and information in correlation based financial networks. The European Physical Journal B-Condensed Matter and Complex Systems 38(2), 353–362 (2004)
14. Palla, G., Derenyi, I., Farkas, I., Vicsek, T.: Uncovering the overlapping community structure of complex networks in nature and society. Nature 435, 814 (2005)
15. Resta, M.: On a data mining framework for the identification of frequent pattern trends. In: Perna, C., Sibillo, M. (eds.) Mathematical and Statistical Methods for Actuarial Sciences and Financial Markets, pp. 173–176. Springer International Publishing
16. Rovetta, S., Masulli, F., Mahmoud, H.: Neighbor-based similarities. In: Masulli, F. (ed.) WILF 2013. LNCS (LNAI), vol. 8256, pp. 161–170. Springer, Heidelberg (2013)
17. Rovetta, S., Masulli, F.: Visual stability analysis for model selection in graded possibilistic clustering. Information Sciences 279, 37–51 (2014)
18. Rand, W.M.: Objective criteria for the evaluation of clustering methods. Journal of the American Statistical association 66(336), 846–850 (1971)
19. Tumminello, M., Coronnello, C., Lillo, F., Micciche, S., Mantegna, R.N.: Spanning trees and bootstrap reliability estimation in correlation-based networks. International Journal of Bifurcation and Chaos 17(07), 2319–2329 (2007)
20. Von Luxburg, U.: A tutorial on spectral clustering. Statistics and Computing 17, 395–416 (2007)

Video-Based Access Control
by Automatic License Plate Recognition

Emanuel Di Nardo[1], Lucia Maddalena[2], and Alfredo Petrosino[1]

[1] University of Naples Parthenope,
Department of Science and Technology, Naples, Italy
[2] National Research Council,
Institute for High-Performance Computing and Networking, Naples, Italy
`emanuel.dinardo@gmail.com, alfredo.petrosino@uniparthenope.it,`
`lucia.maddalena@na.icar.cnr.it`

Abstract. We report an access control system based on automatic license plate recognition, consisting of three main modules for acquisition, extraction, and recognition. The basic idea is to couple the online learning of a neural background model with a stopped foreground subtraction mechanism to efficiently provide a subset of relevant video frames where to look for. Another key point is the use of matching the entire license plate ROI with those stored in a database of authorized license plates, based on suitable features and validation tests. Experimental results confirm that the proposed system attains overall performance comparable with that of the state-of-the-art ALPR methods.

Keywords: Automatic License Plate Recognition, Access Control System, Neural-based Vehicle Detection.

1 Introduction

Automatic license plate recognition (ALPR) consists in extracting vehicle license plate information from images or image sequences taken by fixed or mobile cameras, identifying their unique associated identities [11]. Examples of applications include access control, where the plate number captured by a fixed camera is used to automatically allow the entrance in restricted areas to registered users, low-enforcement, where roadside cameras are adopted to detect vehicles violating traffic laws, and road patrolling, where vehicles equipped with installed or handheld cameras are adopted to monitor vehicular traffic [14].

ALPR is widely regarded to be a solved problem, even though the proposed systems often are only applicable under restricted illumination, view-point, and plate specification conditions, or require specialized hardware [27]. In this work, we propose an access control system (ACS) based on ALPR, designed in order to provide as much as possible recognition accuracy, at the same time relying as much as possible on off-the-shelf non-specialized hardware. Therefore, the reference setting includes a fixed, standard-resolution video camera, positioned at the entrance of a restricted access area.

The paper is organized as follows. In Section 2, we present a fairly compact overview of the approaches to ALPR, providing links to appropriate references for extensive surveys. Section 3 describes the basic building blocks of the proposed system. In Section 4, we present results achieved with the proposed system, also providing performance comparisons with other existing systems. Section 5 includes conclusions and further research directions.

2   Related Work

Most of the modern ALPR systems described in the literature can be reconducted to a three-step scheme that includes acquisition, extraction, and recognition.

The *acquisition* step is aimed at acquiring vehicle images using a camera and determining when the subsequent steps must be activated. Indeed, the continuous monitoring of the scene under surveillance is a computationally demanding task, that could also lead to incoherent results. Therefore, most of the modern ALPR systems [19], [4], [1] implement an acquisition step that detects the presence of vehicles in the monitored area. Acquisition can be achieved through specialized sensors (usually infrared or ultrasound sensors) or through methodologies that detect new objects in the scene, usually based on background subtraction.

The *extraction* step (often referred to as "localization" or "detection") performs an automatic selection of the license plate region of interest (ROI), in order to limit the image area for applying the subsequent step. This not only reduces precessing times, but also avoids in the recognition step the presence of disturbing objects, that could generate confusion. Generally, extraction exploits license plate features in order to distinguish it by other scene objects. These features can include image edges, texture, color, spatial measurements, presence of characters, or a combination of them. For extensive and up-to-date reviews of license plate extraction approaches, the interested reader is referred to [2], [11]. The task is hindered by several issues, since license plates may be different from state to state (in terms of dimensions, color, number and distribution of characters), the presence of other text areas in the scene can generate confusion, illumination conditions as well as plate dirtiness can strongly influence extraction accuracy.

The *recognition* step allows the system to identify the license plate included in the detected ROI. Recognition is very often achieved by segmenting each single character and applying optical character recognition algorithms in each of the segmented regions, as witnessed by the abundant literature reported in [2], [11]. Very rarely, recognition is achieved by extracting features by the entire license plate, which are then matched between the current frame and images included into a license plate dataset [7], [10]. Among the main issues of the recognition step are the invariance to license plate rotation and scaling, as well as to illumination conditions, which can be better handled by the second approach.

3   The Proposed System

The proposed ALPR system follows the three-step scheme described in the previous section in order to automatically recognize a license plate in a dataset of license plates allowed to enter a restricted access area. Our choices for the three steps are detailed in the following.

3.1   Acquisition

The acquisition module is based on foreground detection, achieved by neural-based background subtraction, and stopped foreground detection, in order to detect cars that stop in the monitored area, causing a trigger alarm to be issued for opening the barrier in case the license plate of the car is recognized.

For moving object detection we adopt the *3dSOBS+* algorithm [23], based on the neural background model $\mathcal{B}_t$ automatically generated and updated at each time t by a self-organizing method. The algorithm is shown to accurately handle most of the well known issues related to background maintenance for moving object detection (moving backgrounds, gradual illumination variations, shadows cast by moving objects) and to be robust against false detections for different types of videos taken with stationary cameras.

For the detection of stopped objects, we adopt the *SFS* algorithm proposed in [22]. The basic idea consists of keeping a model $\mathcal{F}_t$ of moving foreground pixels, that is similar to the neural model adopted for background pixels. At each time t, foreground pixels are classified as stopped pixels if their moving foreground model holds the same features for at least τ consecutive frames, with τ stationary threshold whose choice is application dependent. The model for stopped pixels is moved to a stopped foreground model $\mathcal{S}_t$, while remaining foreground pixels are classified as moving pixels.

3.2   Extraction

In order to extract the license plate ROI, we rely on Radon projections of the image edges, also exploiting a priori information on license plates, including their usual aspect ratio, the color contrast between characters and background, as well as the presence of characters in the searched area.

For each sequence image I, after median filtering pre-processing, image edges are extracted through the Sobel operator. Image projections $P_x(x)$ and $P_y(y)$ are computed in the horizontal and vertical directions, respectively:

$$P_x(x) = \sum_{j=0}^{h-1} I(x,j), \quad P_y(y) = \sum_{i=0}^{w-1} I(i,y), \tag{1}$$

where $w \times h$ is the size of I. Then, projection peaks and extremes of the peaks region that identify the license plate ROI are detected. Specifically, in the case of horizontal projections, peaks x_p are computed as

$$x_p = \arg \max_{0 \leq x < w} P_x(x) \tag{2}$$

and extremes x_l and x_r are computed as:

$$x_l = \max_{0 \leq x \leq x_p} \{x|P_x(x) \leq c * P_x(x_p)\}, \quad x_r = \max_{x_p \leq x < w} \{x|P_x(x) \leq c * P_x(x_p)\},$$
(3)

with $c \in [0,1]$ constant value. Analogous formulas hold for vertical projections.
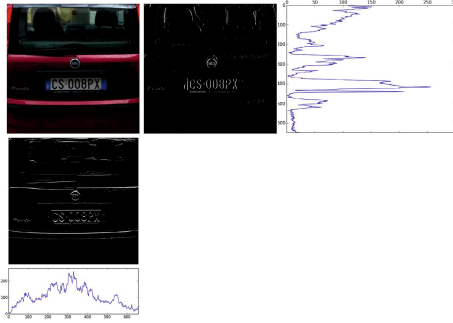
Fig. 1. License plate extraction by horizontal and vertical projections of image edges

In order to make sure that all possible rectangular ROIs are taken into account, we select n_p highest local maxima in each projection direction, for a total of n_p^2 ROIs. A further euristic postprocessing of the extracted ROIs is carried out, aimed at ensuring that each of them really includes a license plate and pruning the others:

1. *Refinement according to the shape*: If the size of a ROI is too much higher or lower than the expected license plate area A, the detected region is likely linked to one of the excess local maxima taken into account. Moreover, a license plate should have a higher number of holes as compared to other scene object; therefore, a region is pruned if its Euler number is less than a fixed number n_E.
2. *Aspect Ratio*: a ROI is discarded if its aspect ratio r is too different from the expected aspect ratio. In order to take into account the acquisition noise and possible adverse illumination conditions, a ROI is pruned only if its aspect ratio is outside the range $[r - \delta, r + \delta]$, with δ experimentally chosen threshold.
3. *Brightness analysis*: a further test is based on the total brightness reflected by the plate surface. Usually, license plates have dark characters on a light background, thus showing overall high brightness. Therefore, after converting the ROI into the HSV color space, we compute the histogram H of the brightness component B and choose the smallest (b_{min}) and largest (b_{max}) non-empty classes of H, and their average b_{med}:

$$b_{min} = \arg\min_{b \in B}(H(b)|H(b) \neq 0), \quad b_{max} = \arg\max_{b \in B}(H(b)|H(b) \neq 0), \quad (4)$$

$$b_{med} = \frac{b_{min} + b_{max}}{2}. \tag{5}$$

The value β given by the difference of the sums of the two identified areas of the histogram:

$$\beta = \left(\sum_{b_1=b_{med}}^{b_{max}} H(b_1) \right) - \left(\sum_{b_2=b_{min}}^{b_{med-1}} H(b_2) \right) \tag{6}$$

provides an indication wether the ROI has a sufficient brightness ($\beta > 0$) to be considered as including a license plate or should be pruned. Analogous reasoning can be applied to the case of light characters over a dark background.

4. *Characters presence*: The presence of characters in a selected ROI is checked in order to discard ROIs including less than n_c characters. After contrast enhancement, the detection of characters is performed through horizontal projections of license plate ROI's edges, in a way similar to what has been done for extracting the possible ROIs in the entire image, leading to a segmentation of characters into ROI's blocks (see Fig. 2).
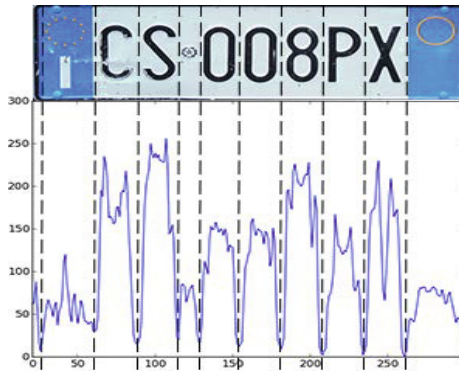
Fig. 2. Projection-based character segmentation into ROI's blocks

3.3   Recognition

In our ACS application context, the proposed recognition module relies on matching the extracted license plate ROI (testing dataset) with those stored in a database of authorized license plates (training dataset), based on suitable features and validation tests. As it will be shown also through experimental results (Section 4), this approach makes the recognition step robust to illumination and position variations, to plate surface irregularities, and to partial occlusions.

Features representing the extracted license plate ROIs are based on Affine-SIFT (ASIFT) [29], a fully affine invariant image comparison method that is robust not only to translation, rotation, and scaling, but also to image distortions arising by the camera orientation. Similarly to the well known SIFT [21], it produces a 128-dimensional feature vector characterizing each keypoint, but it tends to use a higher number of keypoints.

Feature matching to analyze the similarity of each feature vector T_k in the test image with feature vectors D_i in the training dataset is based on nearest neighbor search using the Euclidean distance [20], that identifies the training image having the nearest feature vector D_1. The adopted space partitioning technique is the Randomized KD-Tree [8], [26], that iteratively subdivides the search space into sub-regions that contain half the points of the original region, using more than one search tree.

Three validation tests follow, in order to exclude from the matching results those keypoints whose feature vectors have no good match in the training set:

1. The first validation test considers the Nearest Neighbor Distance Ratio (NNDR) [20], [25], that compares the closest feature vector D_1 with the second closest feature vector D_2 belonging to a different class:

$$\frac{||D_1 - T_k||_2}{||D_2 - T_k||_2} < \rho_1, \tag{7}$$

   with $\rho_1 \in [0, 1]$. NNDR discards a match if the L_2 distance from the nearest matched feature vector is not significantly different from that of a different license plate.

2. The shape validation test relies on Hu moments [15], adopted to describe the shape of objects related to the matched keypoints in a way that is invariant to scaling, rotation, and translation. If the objects are not enough similar according to these moments, the match is discarded. Specifically, for each couple (T, D) of testing and training keypoints, the seven Hu invariant moments $h_j^T, h_j^D, j = 1, \ldots, 7$ are computed on the contours of the corresponding objects in the testing and training images. The match will be discarded if these contours are too different, i.e., if

$$Diss(T, D) = \max_{j=1,\ldots,7} \frac{\left|m_j^T - m_j^D\right|}{\left|m_j^T\right|} > \rho_2, \tag{8}$$

   with

$$m_j^T = sign(h_j^T) * \log h_j^T, \quad m_j^D = sign(h_j^D) * \log h_j^D \tag{9}$$

   and $\rho_2 \in [0, 1]$.

3. As a last validation step, the homography between training and testing matched images is computed by RANSAC [12] to further prune outliers, i.e., those keypoints whose re-projection error is greater than ρ_3 pixels.

At the end, a testing license plate k is recognized as license plate j of the training dataset if:

$$j = \arg\max_i FM(T_k, D_i) \quad \text{AND} \quad \max_i FM(T_k, D_i) \geq 3,$$

where $FM(T_k, D_i)$ indicates the number of matching testing/training features that have passed the three validation steps.

4   Experimental Results

4.1   Data

For testing the proposed ACS, we produced the *ACS Video Dataset*[1], including eight home-made color videos of size 1280×720, for a total of 5900 frames. These are typical ACS videos, taken from three different view-points and under different illumination conditions. Example frames of each video, identified by the license plate number, are reported in Fig. 3.

| BL021TA | CS008PX | EH246ZK | ER984ZN |
| BD691JJ | CM640GG | DP756YZ | DW072YY |

Fig. 3. Example frames from the ACS Video Dataset

In order to focus the attention on the only image area where to look for license plates, for each of the three different view-points we defined a *search area* (see white pixels in the masks of Fig. 4) where the proposed ACS is applied.
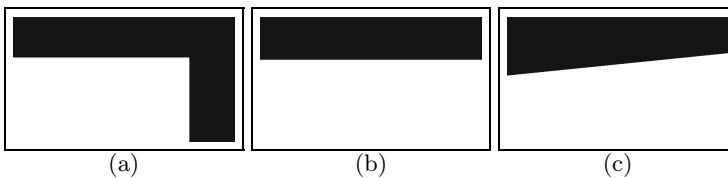
Fig. 4. Search areas for: (a) BL021TA and CS008PX; (b) EH246ZK and ER984ZN; (c) BD691JJ, CM640GG, DP756YZ, and DW072YY

For the recognition phase, we further produced the *ACS Recognition Dataset*[1], an image database of fifty different license plates used for recognition. Example

---

[1] The ACS Video Dataset and the ACS Recognition Dataset are available for download at http://cvprlab.uniparthenope.it.

(a)   (b)   (c)   (d)

Fig. 5. Example images from the ACS Recognition Dataset. Similar license plates can be observed in (c) CS008PX (original) and (d) CS000PX (digitally modified).

images are reported in Fig. 5. To better test the recognition performance, this database includes also cases of very similar license plates, such as the one in Fig. 5-(d) that has been obtained by digitally modifying the digit 8 in the original license plate of Fig. 5-(c).

4.2   Acquisition Results

Fig. 6 shows the results of the acquisition step described in Section 3.1 on video BL021TA of the ACS Video Dataset. As soon as single pixels are detected as moving and similar to the foreground for τ consecutive frames, they are classified as stopped (red pixels in Figs. 6-(a) and (e)), and moved from the moving foreground model $\mathcal{F}_t$ (Fig. 6-(d) and (h)) to the stopped foreground model $\mathcal{S}_t$ (Fig. 6-(c) and (g)). Further foreground pixels previously covered by the barrier have not yet reached the stationary threshold in frame $t = 420$, and are still stored in the moving foreground model $\mathcal{F}_t$ (Fig. 6-(d)).

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)
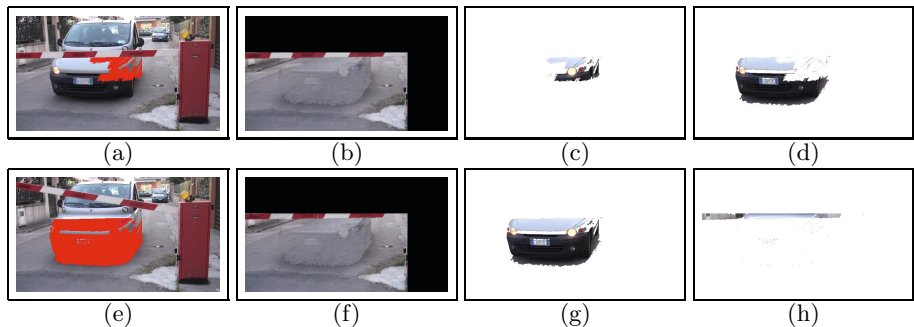
Fig. 6. Acquisition step on video BL021TA of the ACS Video Dataset, frame $t = 350$ (first row) and $t = 420$ (second row): stopped foreground pixels (first column); representations of the background model $\mathcal{B}_t$ (second column), stopped foreground model $\mathcal{S}_t$ (third column), and moving foreground model $\mathcal{F}_t$ (fourth column)

In Table 1, we report results of the acquisition module on each sequence of the ACS Video Dataset, obtained choosing a stationary threshold $\tau = 80$ (values for all remaining parameters have been chosen as in [22,23] for all the video sequences). The second column reports the number i_S of the frame in which the

Table 1. Results of stopped foreground detection on the ACS Video Dataset

Video	Start i_S of stopped event	GT $(i_S + \tau)$	Stopped event trigger issued
CS008PX	172	252	228
DW072YY	164	244	145
BL021TA	288	368	344
BD691JJ	249	329	222
DP756YZ	533	613	433
CM640GG	306	386	245
EH246ZK	210	290	262
ER984ZN	246	326	296

car begins stopping, the third column reports the number of the frame where the stopped object event should be detected (the "Ground Truth" - GT), while the fourth column reports the number of the frame where the stopped object event has been detected. It can be observed that the acquisition module triggers the stopped alert about one second earlier than expected. Indeed, the pixel-based approach starts signaling stopped foreground pixels of the uniformly colored auto body before the full auto front side stops. Even though this anticipation has proved to be beneficial to the system, providing further initial frames where to look for possible license plates, region-level post-processing of the stopped foreground masks could easily help in detecting only the complete object as stopped, based on the pixel-wise information.

Further experimental results concerning moving and stopped object detection accuracy on publicly available sequences can be found in [22,23].

4.3   Extraction Results

The extraction step for the ACS Video Dataset has been performed on all sequence frames where stopped foreground objects have been signaled by the acquisition step. Examples of extracted license plate ROIs are reported in Fig. 7, where we can observe high accuracy in the identification of ROI borders. Only few extracted ROIs have been partially detected (e.g., Fig. 7-(e) includes only some of the license plate digits) or are completely wrong (e.g., Fig. 7-(f) does not include any license plate). These partial/complete failures are due to the camouflage of the car with the background, to the license plate orientation, or to the illumination conditions, that can negatively influence the segmentation, the projections, or the extraction of the license plate.

In Table 2, for each sequence of the ACS Video dataset we report results of the extraction module in terms of number of correct (third column), partial (fourth column), and wrong (fifth column) extracted ROIs as compared to the total number of extracted ROIs (second column). In all the experiments, values for the extraction parameters (see Section 3.2) have been chosen based on a priori information on Italian license plates and on experiments as follows: c=0.15

| (a) | (b) | (c) | (d) | (e) | (f) |

Fig. 7. Extraction step on ACS Video Dataset: examples of correct ((a)–(d)), partial (e), and wrong (f) extracted ROIs

in Eq. (3); n_p=3 for the number of highest local maxima in each projection direction; expected license plate area A in the range of $[20,150] \times [20,150]$ pixels and threshold $n_E = 3$ for the Euler number (postprocessing step 1); expected license aspect ratio $r = 3.27$ (Italian license plates standard dimensions are width 360mm and height 110mm), with δ=1 (postprocessing step 2); minimum number of character ROIs n_C=3 (postprocessing step 4).

It should be pointed out that, although few incomplete or wrong ROIs were extracted, the extraction step succeeded in extracting a more than sufficient number of correct license plate ROIs for the subsequent recognition module.

Table 2. Results of ROI extraction on the ACS Video Dataset

Video	Extracted ROIs	Correct ROIs	Partial ROIs	Wrong ROIs
CS008PX	204	194	10	0
DW072YY	92	24	66	2
BL021TA	86	86	0	0
BD691JJ	85	75	9	1
DP756YZ	350	329	21	0
CM640GG	339	311	28	0
EH246ZK	336	336	0	0
ER984ZN	125	124	0	1
Avg.		91.5%	8.3 %	0.2 %

In order to provide comparisons of the extraction results with those of other existing approaches, we considered the software JavaANPR [24], a system for ALPR in still images. In Table 3, for each license plate we report its results on sixty selected sequence frames of each video of the ACS Video Dataset. Here, we can observe that JavaANPR accuracy, in terms of correct/wrong extracted ROIs, is quite low for this dataset, achieving on average 42.5% of correctly extracted ROIs (as compared to the average 91.5% of the proposed ACS reported in Table 2).

4.4   Recognition Results

Examples of recognition results of testing license plates in the ACS Recognition Dataset are reported in Fig. 8. Here, we can observe that matched keypoints (green circles) perfectly match (green lines connecting them).

Table 3. Extraction results of the software JavaANPR [24] on sixty selected sequence frames of each video of the ACS Video Dataset

Video	Correct ROIs	Partial/Wrong ROIs
CS008PX	10	50
DW072YY	3	57
BL021TA	36	24
BD691JJ	21	39
DP756YZ	17	43
CM640GG	35	25
EH246ZK	32	28
ER984ZN	0	60
Avg.	42.5%	57.5%

Fig. 8. Recognition step: license plates extracted by the ACS Video Dataset (top of each figure) and correctly matched with license plates of the ACS Recognition Dataset (bottom of each figure). Green circles indicate keypoints common to the matched images and green lines connect matched keypoints.

Table 4 provides results of the proposed recognition step, also comparing them with those obtained by an analogous recognition module, but based on SIFT, rather than ASIFT, features. For all the experiments, values for recognition parameters ρ_1, ρ_2, and ρ_3 (see Section 3.3) have been experimentally fixed as 0.8, 0.2, and 5, respectively. We can observe that the recognition module perfectly recognizes all the correctly extracted license plates, notwithstanding the very similar license plates included into the ACS Recognition Dataset (Fig. 5). Such good results are strictly linked to the choice of the ASIFT feature descriptors, as verified by comparison with the well known SIFT descriptors.

4.5 Further Comparisons

In order to further compare the accuracy of the proposed ACS with that of other existing systems, Table 5 reports the performance of recently proposed ALPR systems, each achieved on a different dataset. Here, the Extraction Rate refers to the percentage of correctly extracted licence plate ROIs and the Recogni-

Table 4. Results of license plate recognition using SIFT and ASIFT features

Plate	Total ROIs	SIFT			ASIFT		
		Correct recogn.	Wrong recogn.	Non recogn.	Correct recogn.	Wrong recogn.	Non recogn.
CS008PX	194	85	10	99	194	0	0
DW072YY	24	24	0	0	24	0	0
BL021TA	86	86	0	0	86	0	0
BD691JJ	75	74	1	0	75	0	0
DP756YZ	329	321	0	8	329	0	0
CM640GG	311	309	0	2	311	0	0
EH246ZK	336	336	0	0	336	0	0
ER984ZN	124	124	0	0	124	0	0
Avg.		91.9%	0.7 %	7.4 %	100%	0 %	0 %

tion Rate refers to the percentage of correct plate recognitions (resulting by the product of character segmentation and character recognition rates for methods performing these two sub-steps), as reported by the respective authors. The System Performance indicates the percentage of licence plates correctly recognized by the system, obtained as:

$$SystemPerformance = ExtractionRate \times RecognitionRate.$$

Table 5 helps us to conclude that the proposed ACS achieves the highest Recognition Rate but almost the lowest Extraction Rate, even though the System Performance is comparable with that of recently proposed approaches. Further work will be devoted to enhance the ROI extraction module.

Table 5. Performance comparison of different systems

Method	Extraction Rate	Recognition Rate	System Performance	Plate Format
[5] (2008)	91.70%	79.25%	72.67%	Turkish
[6] (2009)	97.30%	95.70%	93.10%	Chinese
[16] (2009)	98.40%	97.30%	95.70%	Motorcycle
[17] (2009)	95.90%	92.30%	88.52%	Multinational
[13] (2010)	88.10%	98.25%	86.56%	Greek
[18] (2010)	97.30%	86.48%	84.14%	Iranian
[28] (2010)	96.80%	90.00%	87.50%	Taiwanese
[27] (2011)	98.30%	95.20%	93.50%	Multinational
[9] (2011)	91.00%	95.50%	86.90%	Iranian
[3] (2014)	96.80%	97.52%	94.40%	Iranian
Proposed	91.50%	100.00%	91.50%	Italian

5   Conclusions

In this paper we propose an access control system based on automatic license plate recognition, consisting of three main modules for acquisition, extraction, and recognition. We show how the online learning of a neural background model, coupled with a stopped foreground subtraction mechanism, can be exploited for acquisition, in order to activate the subsequent modules and provide a subset of relevant video frames where to look for. To extract the license plate ROI, we rely on Radon projections of the image edges, also exploiting a priori information on license plates. The recognition module, instead of segmenting characters and then recognizing each of them, relies on matching the entire license plate ROI with those stored in a database of authorized license plates, based on suitable features and validation tests. Experimental results show that, although the extraction module could be improved, the 100% success rate of the recognition module, that does not require online training, makes the proposed system attain overall performance comparable with that of the state-of-the-art ALPR methods.

References

1. Anagnostopoulos, C.N.: License plate recognition: A brief tutorial. IEEE Intelligent Transportation Systems Magazine 6(1), 59–67 (2014)
2. Anagnostopoulos, C.N., Anagnostopoulos, I., Psoroulas, I., Loumos, V., Kayafas, E.: License plate recognition from still images and video sequences: A survey. IEEE Transactions on Intelligent Transportation Systems 9(3), 377–391 (2008)
3. Ashtari, A., Nordin, M., Fathy, M.: An iranian license plate recognition system based on color features. IEEE Transactions on Intelligent Transportation Systems (2014) (to appear)
4. Bailey, D., Irecki, D., Lim, B.K., Yang, L.: Test bed for number plate recognition applications. In: Proceedings of the First IEEE International Workshop on Electronic Design, Test and Applications, pp. 501–503 (2002)
5. Caner, H., Gecim, H., Alkar, A.: Efficient embedded neural-network-based license plate recognition system. IEEE Transactions on Vehicular Technology 57(5), 2675–2683 (2008)
6. Chen, Z.X., Liu, C.Y., Chang, F.L., Wang, G.Y.: Automatic license-plate location and recognition based on feature salience. IEEE Transactions on Vehicular Technology 58(7), 3781–3785 (2009)
7. Comelli, P., Ferragina, P., Granieri, M., Stabile, F.: Optical recognition of motor vehicle license plates. IEEE Transactions on Vehicular Technology 44(4), 790–799 (1995)

8. Dasgupta, S., Sinha, K.: Randomized partition trees for exact nearest neighbor search. CoRR abs/1302.1948 (2013)

9. Dashtban, M.H., Dashtban, Z., Bevrani, H.: A novel approach for vehicle license plate localization and recognition. Int. J. Comput. Appl. 26(11), 22–30 (2011)

10. Dlagnekov, L., Belongie, S.: Recognizing cars. Tech. Rep. CS2005-0833, CSE, UCSD (2005)

11. Du, S., Ibrahim, M., Shehata, M., Badawy, W.: Automatic license plate recognition (ALPR): A state-of-the-art review. IEEE Transactions on Circuits and Systems for Video Technology 23(2), 311–325 (2013)

12. Fischler, M.A., Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24(6), 381–395 (1981)

13. Giannoukos, I., Anagnostopoulos, C.N., Loumos, V., Kayafas, E.: Operator context scanning to support high segmentation rates for real time license plate recognition. Pattern Recognition 43(11), 3866–3878 (2010)

14. Hsu, G.S., Chen, J.C., Chung, Y.Z.: Application-oriented license plate recognition. IEEE Transactions on Vehicular Technology 62(2), 552–561 (2013)

15. Hu, M.K.: Visual pattern recognition by moment invariants. IRE Transactions on Information Theory 8(2), 179–187 (1962)

16. Huang, Y.P., Chen, C.H., Chang, Y.T., Sandnes, F.E.: An intelligent strategy for checking the annual inspection status of motorcycles based on license plate recognition. Expert Systems with Applications 36(5), 9260–9267 (2009)

17. Jiao, J., Ye, Q., Huang, Q.: A configurable method for multi-style license plate recognition. Pattern Recognition 42(3), 358–369 (2009)

18. Kasaei, S.H., Kasaei, S.M., Kasaei, S.A.: New morphology-based method for robust iranian car plate detection and recognition. Int. J. Comput. Theory Eng. 2(2), 264–268 (2010)

19. Kim, K.K., Kim, K., Kim, J., Kim, H.: Learning-based approach for license plate recognition. In: Proceedings of the 2000 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing X, vol. 2, pp. 614–623 (2000)

20. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)

21. Lowe, D.: Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2, pp. 1150–1157 (1999)

22. Maddalena, L., Petrosino, A.: Stopped object detection by learning foreground model in videos. IEEE Trans. Neural Net. and Learn. Sys. 24(5), 723–735 (2013)

23. Maddalena, L., Petrosino, A.: The 3dSOBS+ algorithm for moving object detection. Computer Vision and Image Understanding 122(0), 65–73 (2014)

24. Martinsky, O.: Algorithmic and mathematical principles of automatic number plate recognition systems (2006), http://javaanpr.sourceforge.net/

25. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(10), 1615–1630 (2005)

26. Muja, M., Lowe, D.G.: Fast approximate nearest neighbors with automatic algorithm configuration. In: VISAPP International Conference on Computer Vision Theory and Applications, pp. 331–340 (2009)

27. Thome, N., Vacavant, A., Robinault, L., Miguet, S.: A cognitive and video-based approach for multinational license plate recognition. Machine Vision and Applications 22(2), 389–407 (2011)
28. Wang, M.L., Liu, Y.H., Liao, B.Y., Lin, Y.S., Horng, M.F.: A vehicle license plate recognition system based on spatial/frequency domain filtering and neural networks. In: Pan, J.-S., Chen, S.-M., Nguyen, N.T. (eds.) ICCCI 2010, Part III. LNCS, vol. 6423, pp. 63–70. Springer, Heidelberg (2010)
29. Yu, G., Morel, J.M.: A fully affine invariant image comparison method. In: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2009, pp. 1597–1600 (April 2009)

Part IV
Signal Processing

On the Use of Empirical Mode Decomposition (EMD) for Alzheimer's Disease Diagnosis

Domenico Labate[1,2], Fabio La Foresta[1], Giuseppe Morabito[3],
Isabella Palamara[1], and Francesco Carlo Morabito[1]

[1] Department of Civil Engineering, Energy, Environment and Materials (DICEAM)
Mediterranea University of Reggio Calabria, Reggio Calabria I-89060, Italy
[2] DIMES – University of Calabria, Cosenza, Italy
[3] University of Pavia, Pavia, Italy
{domenico.labate,fabio.laforesta,
isabella.palamara,morabito}@unirc.it,peppe_mb}@hotmail.it

Abstract. Alzheimer's Disease (AD) is considered one of the most common form of dementia; it involves a progressive decline in cognitive function because of pathological modifications or damage of the brain. One of the major challenges is to develop tools for early diagnosis and disease progression. Electroencephalogram represents potentially a noninvasive and relatively non-expensive approach for screening of dementia and AD. It provides a method to objectively quantify the cortical activation patterns but it is usually considered insensitive in the early AD. This study introduces a novel method where electroencephalographic recordings (EEG) are subjected to Empirical Mode Decomposition (EMD), which decomposes a signal into components known as Intrinsic Mode Functions (IMFs). The results, suggest that, the IMFs may be used to determine the particular frequency bandwidths in which specific phenomena occur.

Keywords: EEG, Alzheimer's Disease, Classification, EMD.

1 Introduction

The brain is a highly complex and non linear system. Alzheimer Disease (AD) is the most common neurodegenerative disorder. It that involves a progressive decline in cognitive function due to atrophy of the brain as well as alteration of connectivity profiles. AD manifests itself through a slowly progressive impairment of mental functions whose course lasts several years [1-3]. Clinically, the evaluation of memory decline is evaluated by neuropsychiatric tests (Mini Mental State Examination, MMSE) but age and education can compromise results. Some images techniques like Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), Magnetic Resonance Imaging (MRI), are useful to observe structural or functional changes in neurodegenerative disorders. Unfortunately these methods are restricted due the high cost and the related dangers to the exposure to contrast agent. On the other hands the electroencephalogram (EEG) is a non invasive and simple technique and represents a powerful and relatively cheaper approach for screening of dementia and AD. In recent years, many authors have studied the characteristic of

EEG to improve diagnostic power by making use of many different signal processing techniques [4-9]. In particular, the EEG traces of AD patients typically shows three kinds of abnormalities [10-14]:

- i)   slowing, i.e., the increase of the relative power of the low frequency bands and a reduction of the individual mean and peak alpha frequency;
- ii) a reduction of complexity;
- iii) an altered synchrony of the EEG channels recordings.

The use of wavelet transform has been demonstrated useful for processing of EEG signals, also because of the ability of developing a time-frequency tiling of the original time-series. As a by-product, these kinds of decompositions are suitable to highlight and thus cancel some kinds of artifacts, invariably present in EEG traces. Unfortunately, there is no general consensus on the basic wavelet function to be used and this approach requires detailed tailoring and expertise [15]. Furthermore, the suitable wavelet basis can be different for various patients or disease's stage. In this paper, a relatively novel technique, namely, the Empirical Mode Decomposition (EMD), is applied in order to exploit a natural decomposition of the EEG recordings in frequency bands. EMD is an adaptive and fully data-driven technique which obtains the oscillatory modes present in the data, thus producing a variable number of components. This technique is able to cope with possible non linearity and non-stationarity of this physiological signal.

The paper is organized as follows: in Section 2 the theoretical basis of the applied technique are provided; then, in Section 3, the experimental data are described and the obtained results are discussed. Conclusive remarks are provided in Section 4.

2   Methodology

EMD allows to decompose any time series, by means of a process called the *sifting* algorithm [16], into a finite set of oscillatory components by exploiting both local temporal and structural characteristics of the data. [15-19]These components, called "intrinsic mode functions" (IMFs), represent the oscillation modes embedded in the data. The IMFs act as a naturally derived set of basis functions for the signal.

This decomposition does not require any conditions about the stationarity and the linearity of the time-series. The principle of EMD is to locally estimate a signal $x(t)$ as a sum of a local trend, that represents the low frequency part named *residual*, and a local detail component, that represents the high frequency named *Intrinsic Mode Function* (IMF).

$$x(t) = \sum_{i=1}^{n} h_i(t) + r_n(t) \tag{1}$$

where $h_i(t)$ denote the set of IMFs and $r_n(t)$ is the trend within the data, also referred to as the last IMF or residual.

By design, an IMF satisfies two basic properties:

- in the complete data set, the number of extrema and zero crossing are exactly equal or they differ at most by one;
- at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The first condition is similar to the narrow-band requirement and the second one is necessary to ensure that the instantaneous frequency will not have redundant fluctuation as induced by asymmetric waveform [12].
The IMFs extraction, from real world signals, is based on the sifting algorithm [16] as shown in Fig. 1. The algorithm steps are :

1. Detect the extrema (both local maxima and minima) of $x(t)$;
2. Connect local maxima and minima with a spline and let $e_{min}(t)$ and $e_{max}(t)$ the spline that forms the upper and lower envelope of the signal;
3. Compute the local mean envelope : $r(t)= (e_{min}(t) + e_{max}(t))/2$;
4. IMF should have zero local mean so subtract the mean envelope from the original series $d(t)= x(t)-r(t)$ to obtain a proto-IMF;
5. Decide if the proto-IMF $d(t)$ is an IMF by checking the two basic condition described above;
6. If $d(t)$ is an IMF has to be subtracted from the original data and the residual is a new data to fed back to step 1 of algorithm;
7. The sifting procedure ends when the residual of step 5 is a constant, monotonic function. The last residual is considered the trend.

In EMD procedure it is important to focus both on the choice of extrema, in order to avoid over-sampling issue, and on boundary conditions for the analysis of discrete time sequences. IMFs form a complete and "nearly" orthogonal basis for the original signal, in fact, different components can have parts with similar frequencies, at different time duration, but locally any two IMFs tend to be orthogonal.
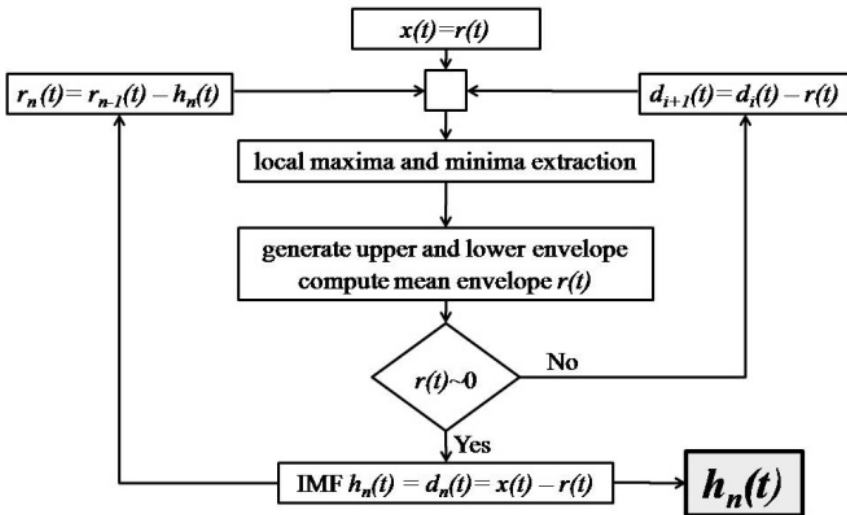
Fig. 1. Block diagram of Intrinsic Mode Function extraction

3   Experimental Results

3.1   Data Description and Acquisition

The analysis was conducted on an experimental EEG database which refers to three different groups of subjects (male and aged between 60–75 years): Healthy Control (HC), Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD).-

The inclusion criteria for enrollment of patients for statistical analysis are mainly standard and at the first level are based on Mini Mental State Examination. All patients are enrolled from "IRCCS - Centro Neurolesi" of Messina, Italy, within an ongoing cooperation agreement. The EEG recordings have been collected according to the sites defined by the standard 10–20 international system, channels (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz), at a sampling rate of 256 Hz. The data are band-pass filtered between 0.5 and 32 Hz, so including the relevant bands for AD diagnosis. In the course of the experimental activity, EEG was recorded in rest condition with closed eyes (under vigilance control).

3.2   Simulation Results

The analysis here presented has been carried out by using codes written by some of the authors in MATLAB environment with specific instructions without using any available tool-boxes. Different IMFs capture the properties of the original signal at different time scale and are presumably generated from different physiological mechanisms. As shown in Figure 2, with reference to the three different classes of subjects, the EEG signal is decomposed by IMFs components. Five components have been shown in the Fig. 2. The finest time scale is shown in the 1st IMF, and the largest is in the fifth one. The frequency gradually decreases moving to lower IMFs. In Fig. 3 are shown the Power Spectral Density (PSD) of the extracted IMFs. This is useful to highlight the well-known "slowing" effect related to the disease. This behavior can be clearly reflected on IMF2. Thus, IMFs power density displays evident variations across the three different classes of subjects. This behavior is, also, well shown in Figure 4 by using a logarithmic scale.

To better highlight different behaviors of the extracted IMFs, four bands are used to categorize the relative PSD: δ (0-4 Hz); θ (4-8 Hz); α (8-13 Hz) and β (13-30 Hz). As shown in Figure 5 the PSDs are distributed differently for the three class of subjects. In δ-band the PSDs of IMF2 and IMF3 are higher in AD-subjects then both HC and MCI; in θ-band the PSD of IMF1 and IMF2 shows the same behavior unlike the PSD of IMF3 is lower than both HC and MCI. Increasing frequency (α-band and β-band) the IMF's PSD of HC manifests higher values, and this behavior is coherent with slowing phenomena.

4 Conclusion

The results here presented suggest that an adaptive data-driven method, such as EMD, can show the dynamics of EEG for the three different classes of subject corresponding to HC, MCI, and AD patients. EMD can be considered a suitable tool for diagnosis and progression of Alzheimer's disease. A detailed analysis of IMFs may yield the possibility of analyzing the basic dynamics characteristic of the three different classes of subjects. This may offer a novel quantitative element to clinicians for evaluating the conversion of MCI patients to AD. In the future, a comparison with other time-frequency techniques will be carried out to understand the relative merits and limitations of different approaches.
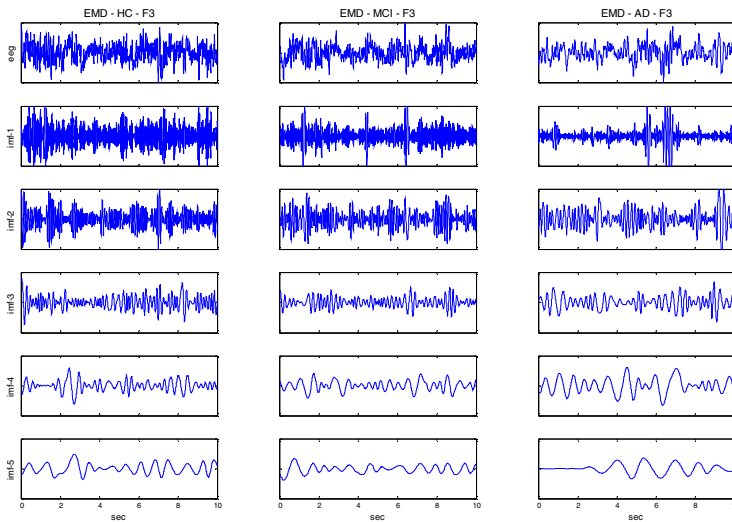
Fig. 2. Empirical Mode Decomposition of EEG recording related to three class of subject (HC, MCI and AD) for F3 electrode. EEG epochs of 10 seconds duration were processed by EMD.
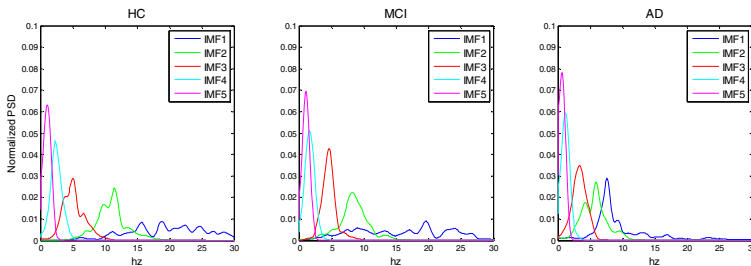
Fig. 3. Power Spectral Density of IMFs related to three class of subjects (HC, MCI and AD)
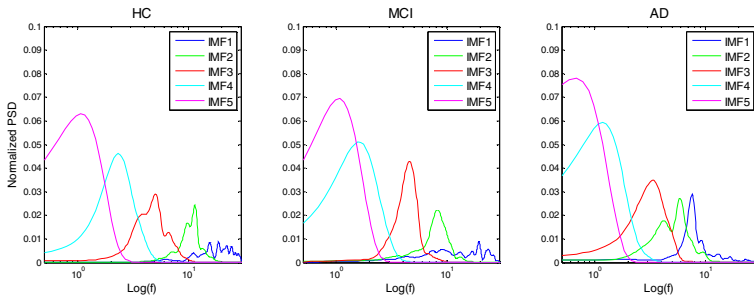
Fig. 4. Power Spectral Density of IMFs for the three classes of subjects (HC, MCI and AD) in logarithmic scale
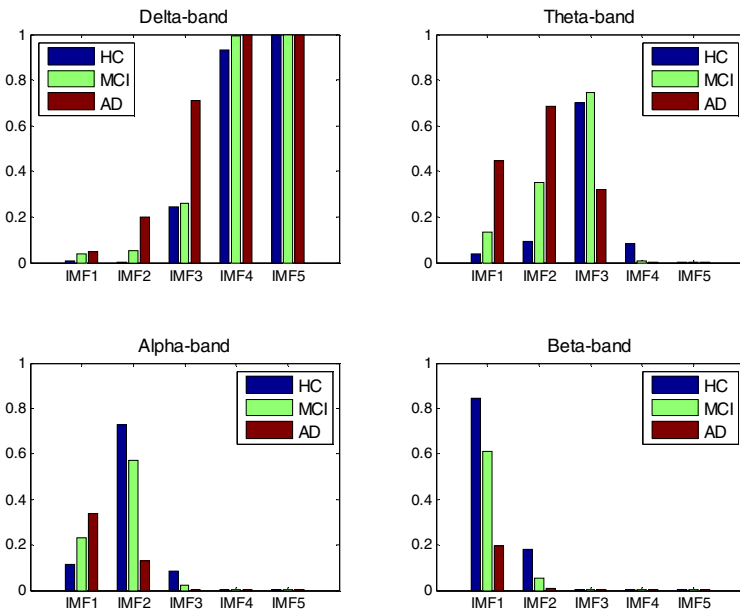
Fig. 5. Normalized Power Spectral Density of IMFs for the three classes of subjects (HC, MCI and AD) categorized into four bands: δ (0-4 Hz); θ (4-8 Hz); α (8-13 Hz) and β (13-30 Hz)

References

1. Jeong, J.: EEG Dynamics in patients with Alzheimer's disease. Clinical Neurophysiology 115, 1490–1505 (2004)
2. Delbeuck, X., van Der Linden, M., Collette, F.: Alzheimer's Disease as a disconnection syndrome, 121, 1438–1446 (2003)
3. Fouquet, et al.: Cerebral imaging and physiopathology of Alzheimer's disease. Psychol Neuropsychiatr Vieil 5, 269–279 (2007)
4. Goldeberger, A.L., Amaral, L.A.N., Hausdorff, J.M., Ivanov, P.C., Peng, C.K., Stanley, H.E.: Fractal dynamics in physiology: Alterations with disease and aging. Nat. Acad. Sci. 99, 2466–2472 (2002)
5. Dauwels, J., Vialatte, F., Latchoumane, C., Jeong, J., Cichocki, A.: EEG synchrony analysis for early diagnosis of Alzheimer's disease: a study with several synchrony measures and EEG data sets. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 2224–2227 (2009)
6. Morabito, F.C., Labate, D., La Foresta, F., Bramanti, A., Morabito, G., Palamara, I.: Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG. Entropy 14(7), 1186–1202 (2012), doi:10.3390/e14071186
7. Inuso, G., La Foresta, F., Mammone, N., Morabito, F.C.: Brain activity investigation by EEG processing: Wavelet analysis, kurtosis and Renyi's entropy for artifact detection. In: Proceedings of the 2007 International Conference on Information Acquisition, ICIA 2007, Jeju City, South Korea, pp. 195–200 (2007), doi:10.1109/ICIA.2007.4295725
8. Azzerboni, B., Finocchio, G., Ipsale, M., La Foresta, F., McKeown, M.J., Morabito, F.C.: Spatio-temporal analysis of surface electromyography signals by independent component and time-scale analysis. In: Proceedings of The Annual International Conference of the IEEE Engineering in Medicine and Biology, vol. 1, pp. 112–113 (2002)
9. Calcagno, S., La Foresta, F., Versaci, M.: Independent component analysis and discrete wavelet transform for artifact removal in biomedical signal processing. American Journal of Applied Sciences 11(1), 57–68 (2014)
10. Mammone, N., Inuso, G., La Foresta, F., Versaci, M., Morabito, F.C.: Clustering of entropy topography in epileptic electroencephalography. Neural Computing and Applications 20(6), 825–833 (2011)
11. Dauwels, J., Srinivasan, K., et al.: Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin? Intl. J. of Alzheimer's Disease (2011)
12. Vialatte, F.B., Cichocki, A., Dreyfus, G., Musha, T., Shishkin, S.L., Gervais, R.: Early Detection of Alzheimer's Disease by Blind Source Separation, Time Frequency Representation, and Bump Modeling of EEG Signals. In: Duch, W., Kacprzyk, J., Oja, E., Zadrożny, S. (eds.) ICANN 2005. LNCS, vol. 3696, pp. 683–692. Springer, Heidelberg (2005)
13. Labate, D., La Foresta, F., Morabito, G., Palamara, I., Morabito, F.C.: Entropic measures of EEG complexity in alzheimer's disease through a multivariate multiscale approach. IEEE Sensors Journal 13(9), 3284–3292, Article number 6552994 (2013)
14. Labate, D., La Foresta, F., Palamara, I., Morabito, G., Bramanti, A., Zhang, Z., Morabito, F.C.: EEG Complexity Modifications and Altered Compressibility in Mild Cognitive Impairment and Alzheimer's Disease. In: Bassis, S., Esposito, A., Morabito, F.C. (eds.) Recent Advances of Neural Networks Models and Applications. SIST, vol. 26, pp. 163–173. Springer, Heidelberg (2014)
15. Labate, D., La Foresta, F., Occhiuto, G., Morabito, F.C., Lay-Ekuakille, A., Vergallo, P.: Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: A comparison. IEEE Sensors Journal 13(7), 2666–2674 (2013), doi:10.1109/JSEN.2013.2257742

16. Mandic, D., Souretis, G., Leong, W.Y., Looney, D., Van Hulle, M.M., Tanaka, T.: Complex Empirical Mode Decomposition for Multichannel Information Fusion. In: Signal Processing Techniques for Knowledge Extraction and Information Fusion, pp. 243–260 (2008)
17. Huang, N.E., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. Royal Soc. London A 454, 903–995 (1998)
18. Campolo, M., Labate, D., La Foresta, F., Morabito, F.C., Lay-Ekuakille, A., Vergallo, P.: ECG-derived respiratory signal using Empirical Mode Decomposition. In: Proceedings of the 2011 IEEE International Symposium on Medical Measurements and Applications (MeMeA 2011), article number 5966727 (2011)
19. Rutkowski, T.M., Cichocki, A., Tanaka, T., Ralescu, A.L., Mandic, D.P.: Clustering of Spectral Patterns Based on EMD Components of EEG Channels with Applications to Neurophysiological Signals Separation. In: Köppen, M., Kasabov, N., Coghill, G. (eds.) ICONIP 2008, Part I. LNCS, vol. 5506, pp. 453–460. Springer, Heidelberg (2009)

Effects of Artifacts Rejection
on EEG Complexity in Alzheimer's Disease

Domenico Labate[1,2], Fabio La Foresta[1], Nadia Mammone[1],
and Francesco Carlo Morabito[1]

[1] Department of Civil Engineering, Energy, Environment and Materials (DICEAM)
Mediterranea University of Reggio Calabria, Reggio Calabria I-89060, Italy
[2] DIMES – University of Calabria, Cosenza, Italy
{domenico.labate,fabio.laforesta,
nadia.mammone,morabito}@unirc.it

Abstract. EEG complexity analysis has recently been shown to help to diagnose Alzheimer's Disease (AD) in the early stages. The complexity study is based on the processing of continuous artifact-free Electroencephalography (EEG). Therefore, artifact rejection is normally required because artifacts might mimic cognitive or pathologic activity and therefore bias the neurologist visual interpretation of the EEG. Furthermore, the EEG complexity analysis is strongly altered by artifacts. In this paper, we evaluate the effects of artifacts rejection by a promising technique, Automatic Wavelet-Independent Component Analysis (AWICA), on the EEG Complexity in AD patients. We also investigate the EEG complexity before and after artifact rejection through some measures based on Shannon's Entropy, Renyi's Entropy and Tsallis's Entropy.

Keywords: Alzheimer's Disease, EEG Complexity, Artifact Rejection.

1 Introduction

Electroencephalography (EEG) is a de facto standard methodology for recording the electrical activity generated by populations of neurons of the cerebral cortex. The major advantage of EEG is being a noninvasive way of recording the neurophysiological activity of patients: for this reason, since its discovery, it has been widely used to investigate the neurological diseases. EEG relies essentially on a multichannel cap that records the bioelectric signals generated by the brain through a set of scalp electrodes, according to the international 10-20 system (Fig. 1). From an information processing perspective, it represents a multivariate, non-stationary, nonlinear time series. Many authors agree that entropy has achieved a large consensus as an indicator of complexity of nonlinear signals.

This assumption is the basis of the complexity study of EEG that aims to differentiate among different brain states through the estimation of entropic measures. In particular, there are dynamical changes of EEG related to normal aging and some others that might reveal aging pathologies. Recent works have shown that the EEG complexity analysis could detect markers of Alzheimer's Disease (AD) in the EEG even in the

early stages [1]-[9]. In fact, the change in the complexity of EEG fluctuations seem to be linked to the evolution of AD disease: this link is not clear yet but there is a increasing evidence that the evolution of AD affects the shape of EEG. This would have a strong impact on the health system since EEG is a cheap and reproducible way to plan and carry out a screening and a follow-up of population at risk. The EEG complexity study is based on the processing of continuous artifacts-free recordings. Unfortunately, the EEG traces are often contaminated by artifacts, signals with non-cerebral origin that overlap to the brain waves. They are generated by different bioelectrical sources such as scalp muscles, eye movements and blinks, sweating, breathing, heart beat, or electrical line noise. The presence of artifacts in the EEG is troublesome and misleading because they can overlap to EEG and heavily obscure the brain waves that the physician needs to examine in order to come up with a reliable diagnosis. Furthermore, if visual inspection is not the final purpose of our analysis but EEG is meant to be processed by any algorithm, artifacts may distort EEG so that the output of the algorithm is not correct. Moreover, even though the physician decided to discard the EEG artifact-laden segments, this would introduce unacceptable discontinuities in the EEG.

In recent years, many authors [10]-[14] dealt with the problem of automatic EEG artifact rejection in order to skip the visual inspection and the subsequent manual artifact rejection from the EEG traces. Recently, Mammone *et al.* [11], introduced a promising automatic method for artifact rejection (AWICA) based on the joint use of Discrete Wavelet Transform (DWT) and Independent Component Analysis (ICA). AWICA is based on the projection of of the single EEG signal into the four frequency bands (delta, theta, alpha and beta) that are then passed through ICA.

In this paper we evaluate the effects of artifact rejection by AWICA on the EEG Complexity in AD patients. In particular, we show that most of artifacts can be removed by AWICA. We also investigate the EEG complexity before and after artifacts rejection through entropic measures based on Shannon's Entropy (SE), Renyi's Entropy (RE) and Tsallis's Entropy (TE).
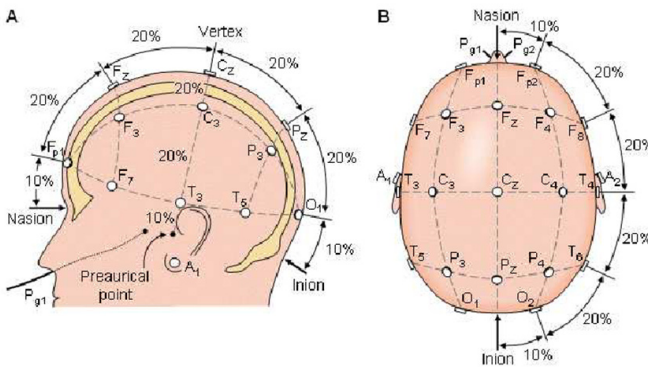
Fig. 1. The international 10-20 system seen from (A) left and (B) above the head. A = Ear lobe, C = central, Pg = nasopharyngeal, P = parietal, F = frontal, Fp = frontal polar, O = occipital

The paper is organized as follows: in Section 2 and 3 we provide a basic description respectively of the Alzheimer's Disease and the AWICA methodology to perform artifacts rejection. In Section 4 we show the results about the effects of artifacts rejection on EEG complexity. Conclusive remarks are provided in Section 5.

2   Effects of Alzheimer's Disease on EEG

Neurodegenerative diseases such as Alzheimer's Disease (AD) have long been the focus of bioengineering researches. Recently, the number of people suffering from AD is estimated in 35 million and the number is expected to raise to 110 million by the year 2050. As the number of elderly population affected by AD rises, the need for making available to the community innovative, accurate, inexpensive and non-invasive diagnostic techniques for early screening of population at risk is becoming a relevantly urgent public health concern [15], [16].

Many researches have shown that the EEG of patients suffering from AD start to modify well in advance of the clinical diagnosis. Furthermore, there are even conditions and diseases that can mimic AD symptoms which are instead reversible. Early diagnosis would be of great importance. As of today, a definitive diagnosis of Alzheimer's is possible only by postmortem necropsy. It is usually diagnosed clinically from the patient history, collateral history from relatives, and clinical observations, based on the presence of characteristic neurological and neuropsychological features and the absence of alternative conditions [17]. Advanced medical imaging with computed tomography (CT) or magnetic resonance imaging (MRI), as well as with single photon emission computed tomography (SPECT) or positron emission tomography (PET) can be used to help exclude other cerebral pathology or subtypes of dementia. However, this kinds of diagnostics are not suitable for screening of large populations. A non-invasive alternative clinical diagnosis is represented by EEG. AD is known to have three main effects on EEG [4], [18]:

1. slowing, i.e. the increase of the relative power of the low frequency bands (delta, 0.5-4 Hz, and theta, 4-8 Hz), coupled with a reduction of the mean frequency (this can be measured by standard Fourier analysis);
2. complexity reduction, by implicitly hypothesizing that regularity of the AD patients' EEG is higher than age-matched controls;
3. loss of synchrony of the electrodes' time series reading: this effect on synchrony can be measured by both nonlinear and linear indices.

Recent studies also give a dynamical description of AD development, data from AD patients showed a loss of complexity over the wide range of time scales, indicating a destruction of nonlinear structures in brain dynamics [1], [9]. These studies was conducted only on artifacts-free EEG segments by cutting the entire artifactual EEG segments. In the next sections, we try to perform the EEG complexity analysis after artifacts rejection procedure and we evaluate the effects through entropic measures.

3   AWICA Methodology for Artifacts Rejection

The AWICA methodology is based on the exploitation of the different information content in the four frequency bands (rhythms) obtained by the DWT step that precedes ICA [11]. The method consists in a two-step artifact identification procedure based on the estimation of kurtosis and Renyi's entropy [19]. The DWT allows to completely recover the neural components of the EEG channels corrupted by the artifacts outside of the contaminated frequency range. AWICA also mostly preserves the cerebral activity because of the increased redundancy of the input to the ICA-step.

The block diagram of AWICA is depicted in Fig. 2: (1) the first level is a decomposition through the Discrete Wavelet Transform (DWT) that partitions each channel of the original dataset into the four major bands of brain activity; each rhythm of each channel is represented by a Wavelet Component (WC). (2) Once the raw data recordings have been so projected into the dimensional space, the Wavelet Components (WCs) linked to artifactual events are automatically identified by means of a quantitative measure and (3) passed through ICA in order to concentrate the artifactual content in a few independent components. (4) Then, the artifactual Wavelet Independent Components (WICs) are automatically selected and rejected. (5) Two reconstruction steps are then performed: the inverse ICA and the inverse DWT, so that the artifact-free EEG dataset is eventually reconstructed (for more details see [11]).

As shown in the next section, AWICA methodology was successfully applied on the artifact-corrupted EEG segments. In addition, the possible alteration of EEG Complexity was evaluated.
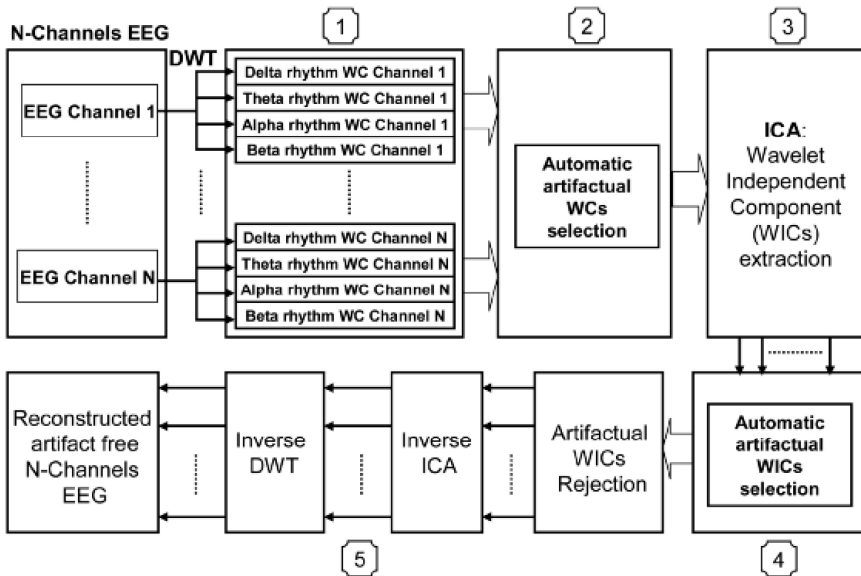
Fig. 2. Block diagram of WICA processing system for EEG artifacts rejection

4   Results

The analysis was conducted on an experimental EEG database which refers to three different groups of subjects (male and aged between 60–75 years): Mild Cognitive Impaired (MCI) patients, AD patients and age-matched healthy elderly control (HC). The inclusion criteria for enrollment of patients for statistical analysis are mainly standard and at the first level are based on Mini Mental State Examination. The EEG database has been made available by the IRCCS "Centro Neurolesi" of Messina, Italy, within an ongoing cooperation agreement. The EEG recordings have been col-lected according to the sites defined by the standard 10–20 international system, 19-channels (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz), at a sampling rate of 256 Hz. The data are band-pass filtered between 0.5 and 32 Hz, so including the relevant bands for AD diagnosis. In the course of the experimental ac-tivity, EEG was recorded in rest condition with closed eyes (under vigilance control). The continuous EEG, whose length is 210 seconds, was partitioned into 21 windows of 10 seconds each. Fig. 3 shows the 12nd window, that was the only time window corrupted by artifacts (on the 7th channel). The artifact was successfully removed by AWICA methodology and the artifact-free continuous EEG was reconstructed. Mammone *et al.* [11] have shown that the artifact rejection performed by AWICA did not introduce significant alterations in the spectrum and the temporal correlation of the EEG. But the AWICA methodology effects on the EEG complexity have not been evaluated. Morabito *et al.* [4] have shown that entropy can be successfully employed to estimate the EEG complexity. Thus, the entropic indexes based on Shannon's Entropy (SE), Renyi's Entropy (RE) and Tsallis's Entropy (TE) were estimated to evaluate the effects on the EEG complexity (for more details about entropic indexes see [1]).

In Fig. 4 we compare the normalized RE, SE and TE of each channels of the 12th window before (red line) and after (green line) artifact rejection with the mean values of normalized RE, SE and TE of each channels computed on the artifact-free win-dows (blue line). All entropic indexes of P3 (7th channel) are corrupted by the artifact. The AWICA artifact rejection is able to remove the alteration of all entropic indexes restoring the measures into mean-value range.

5   Conclusions

The recently introduced entropic complexity measures has been shown to be capable of processing EEG data as an enabling tool for distinguish among different brain states. These indexes are also able to capture the typical "slowing effect" related to Alzheimer's disease. It has been shown that entropic indexes are particularly suitable for monitoring the changes in the elderly brain by distinguish between physiological ageing and pathological dementias. In this paper we have also evaluated the effects of the artifacts rejection by AWICA methodology on the entropic indexes that are em-ployed in the EEG complexity analysis. The results confirm that AWICA is able to rectify the entropy value alteration.
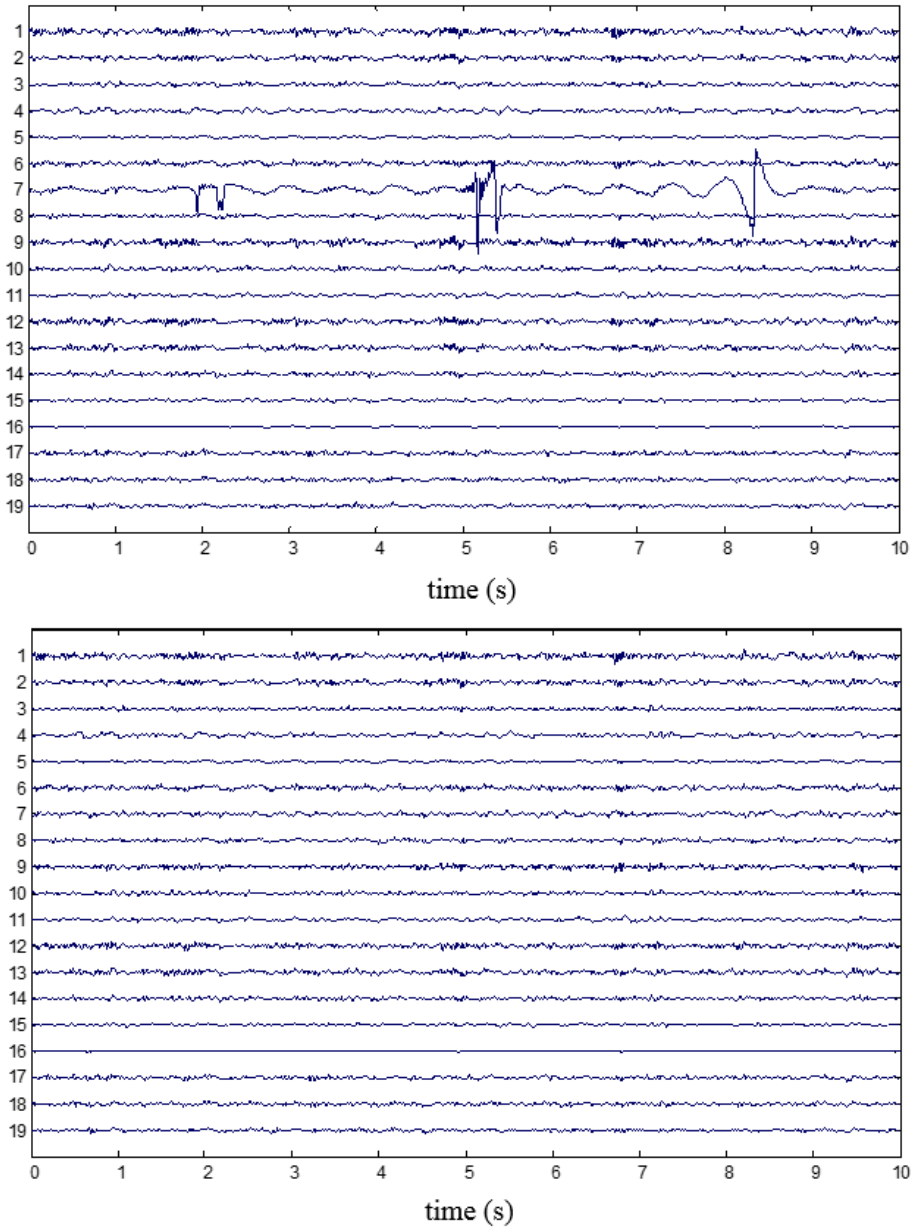
Fig. 3. Multichannel real EEG recorded from an AD patient. According to the international 10-20 system, the montage is Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz. In particular, the 12nd window is shown, that was the only one corrupted by artifacts. (Top) The EEG showing an artifact on the 7th electrode (P3). (Bottom) The same EEG segment after artifact rejection through AWICA.
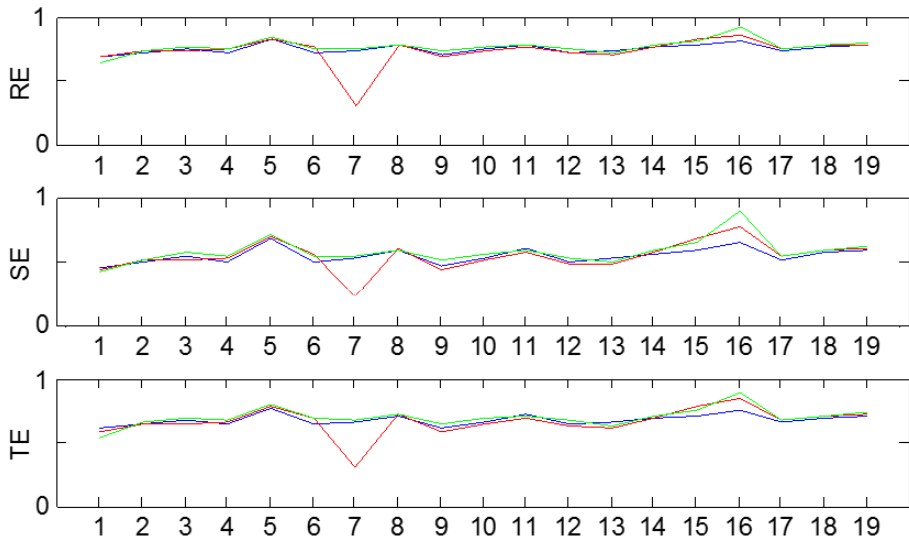
Fig. 4. Evaluation of EEG complexity by normalized Entropic Measures. Comparison of normalized RE, SE and TE of each channels of the 12nd window before (red line) and after (green line) artifact rejection with the mean values of normalized RE, SE and TE of each channels computed on the artifact-free windows (blue line).

References

1. Labate, D., La Foresta, F., Morabito, G., Palamara, I., Morabito, F.C.: Entropic measures of EEG complexity in alzheimer's disease through a multivariate multiscale approach. IEEE Sensors Journal 13(9), 3284–3292 (2013)
2. Labate, D., La Foresta, F., Palamara, I., Morabito, G., Bramanti, A., Zhang, Z., Morabito, F.C.: EEG complexity modifications and altered compressibility in mild cognitive impairment and Alzheimer's Disease. In: Bassis, S., Esposito, A., Morabito, F.C. (eds.) Recent Advances of Neural Networks Models and Applications. SIST, vol. 26, pp. 163–173. Springer, Heidelberg (2014)
3. Ahmed, M.U., Mandic, D.P.: Multivariate Multiscale Entropy Analysis. IEEE Signal Processing Letters 19(2) (2012)
4. Morabito, F.C., Labate, D., La Foresta, F., Bramanti, A., Morabito, G., Palamara, I.: Multivariate Multi-Scale Permutation Entropy for Complexity Analysis of Alzheimer's Disease EEG. Entropy 14, 1186–1202 (2012)
5. Ahmed, M.U., Mandic, D.P.: Multivariate Multiscale Entropy: A tool for complexity analysis of multichannel data. Phys. Rev., E (2011)

6. Mammone, N., Inuso, G., La Foresta, F., Versaci, M., Morabito, F.C.: Clustering of entropy topography in epileptic electroencephalography. Neural Computing and Applications 20(6), 825–833 (2011)
7. Gómez, C., Hornero, R.: Entropy and complexity analyses in Alzheimer's Disease: An MEG study. Open Biomed. Eng. J. 4, 223–235 (2010)
8. Abasolo, D., Hornero, R., Espino, P., Alvarez, D., Poza, J.: Entropy analysis of the EEG background activity in Alzheimer's Disease patients. Phys. Meas. 27, 241–252 (2006)
9. Morabito, F.C., Labate, D., Bramanti, A., La Foresta, F., Morabito, G., Palamara, I., Szu, H.H.: Enhanced compressibility of EEG signal in alzheimer's disease patients. IEEE Sensors Journal 13(9), 3255–3262 (2013)
10. Calcagno, S., La Foresta, F., Versaci, M.: Independent component analysis and discrete wavelet transform for artifact removal in biomedical signal processing. American Journal of Applied Sciences 11(1), 57–68 (2014)
11. Mammone, N., La Foresta, F., Morabito, F.C.: Automatic artifact rejection from multichannel scalp EEG by wavelet ICA. IEEE Sensors Journal 12(3), Article number 5713804, 533–542 (2012)
12. Makarovand, V.A., Castellanos, N.P.: Recovering EEG brain signals: Artifact suppression with wavelet enhanced Independent Component Analysis. J. Neurosci. Methods 158(2), 300–312 (2006)
13. La Foresta, F., Morabito, F.C., Azzerboni, B., Ipsale, M.: PCA and ICA for the extraction of EEG dominant components in cerebral death assessment. In: Proceedings of The 2005 International Joint Conference on Neural Networks, vol. 4, Article number 1556301, pp. 2532–2537 (2005)
14. Azzerboni, B., Finocchio, G., Ipsale, M., La Foresta, F., McKeown, M.J., Morabito, F.C.: Spatio-temporal analysis of surface electromyography signals by independent component and time-scale analysis. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology, vol. 1, pp. 112–113 (2002)
15. Mattson, M.: Pathways towards and away from Alzheimer disease. Nature 430 (2004)
16. Boller, F., Forbes, M.M.: History of dementia and dementia in history: an overview. J. Neurol. Sci. 158(2), 125–133 (1998)
17. Jeong, J.: EEG Dynamics in patients with Alzheimer's disease. Clinical Neurophysiology 115, 1490–1505 (2004)
18. Dauwels, J., Srinivasan, K., et al.: Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin? Intl. J. of Alzheimer's Disease 2011
19. Inuso, G., La Foresta, F., Mammone, N., Morabito, F.C.: Brain activity investigation by EEG processing: Wavelet analysis, kurtosis and Renyi's entropy for artifact detection. In: Proceedings of the 2007 International Conference on Information Acquisition (ICIA 2007), article number 4295725, pp. 195–200 (2007)

Denoising Magnetotelluric Recordings Using Self-Organizing Maps

Luca D'Auria[1], Antonietta M. Esposito[1], Zaccaria Petrillo[1], and Agata Siniscalchi[2]

[1] Istituto Nazionale di Geofisica e Vulcanologia,
Sezione di Napoli Osservatorio Vesuviano, Napoli, Italy
{luca.dauria,antonietta.esposito,zaccaria.petrillo}@ingv.it
[2] Università degli Studi di Bari "Aldo Moro",
Dipartimento di Scienze della Terra e Geoambientali, Bari, Italy
agata.siniscalchi@uniba.it

Abstract. A novel approach for processing magnetotelluric data in urban areas is presented. The magnetotelluric (MT) method is a valid technique for geophysical exploration of the Earth's interiors. It provides information about the rocks' resistivity and in particular, in volcanology, it allows to delineate the complex structure of volcanoes possibly detecting magmatic chambers and hydrothermal systems. Indeed, geological fluids (e.g. magma) are characterized by resistivity of many orders of magnitude lower than the surrounding rocks. However, the MT method requires the presence of natural electromagnetic fields. So in urban areas, the noise strongly influences the MT recordings, especially that produced by trains. Various denoising techniques have been proposed, but it is not always easy to identify the noise-free intervals. Thus, in this work a neural method, the Self-Organizing Map (SOM), is proposed to perform the clustering of impedance tensors, computed on a Discrete Wavelet (DW) expansion of MT recordings. The use of the DW transform is motivated by the need of analyzing MT recordings both in time and frequency domain. The SOM is principally tested on synthetic dataset. Then, as a further validation of the method, it is applied on real data recorded at volcano Etna, Sicily. In both cases, the obtained results have shown the SOM capability of greatly reducing the effect of the noise on the retrieved apparent resistivity curves.

Keywords: magnetotelluric method (MT), denoising, SOM networks.

1 The Magnetotelluric (MT) Method

The Magnetotelluric (MT) method [22] is based on the study of the interaction between natural low frequency electromagnetic waves and the rocks in the Earth's interiors. The simultaneous recording at the Earth surface of the electric and the magnetic fields allows the determination of the resistivity of rocks inside the Earth.

MT signals span a wide range of frequencies: the sources of the high frequency ($1 \div 10^5$ Hz) are lightning while the low frequency source ($10^{-5} \div 1$ Hz) is the

interaction of the solar wind with the magnetosphere and the ionosphere. Both sources realize on the Earth surface a quasi-plane wave orthogonally incident.

Figure 1 illustrates a schematic representation of the physics of magnetotelluric exploration while figure 2 shows how the MT instrumentation operates.
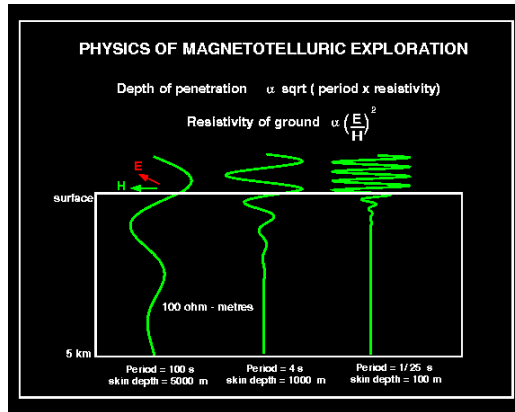
Fig. 1. A scheme of physics of magnetotelluric exploration
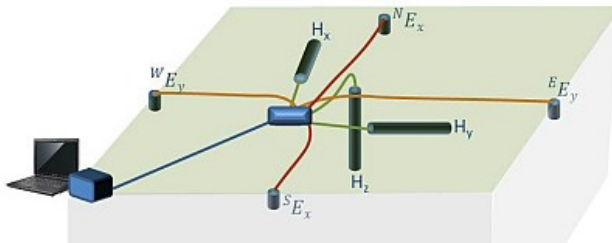
Fig. 2. The MT instrumentation

The physical basis of the method relies on the relationship between the electric and the magnetic fields measured on the Earth surface:

$$E = ZH,$$

where E is the electric field vector, H is the magnetic field vector and Z is the impedance tensor.

Due to the high atmosphere resistivity, the vertical component of the electric field is close to zero. Hence only the horizontal components of the vectors E and H are usually considered. This implies that a 2x2 complex matrix represents the impedance tensor Z.

Natural oscillations of the magnetic field, triggered by various sources (e.g. solar wind transients, thunders, etc) provide a broad spectrum of natural signals useful for MT studies. They excite currents within the Earth (called *telluric currents*, from

which the name magnetotelluric method comes), which are the source of an electric field, whose intensity depends upon the resistivity of the rocks. The longer the period of the oscillation, the deeper is the resistivity investigated. Signals with a period of 10^5s allow the determination of the Earth resistivity up to tens of kilometers beneath the surface. The other end of the spectrum signals with a period of 10^{-2}s determines accurately the resistivity of the rocks up to depth of few hundreds meters. The full determination of the resistivity as a function of the depth requires studying the whole range of MT signal frequencies.

The basic assumption of the MT method is that the electromagnetic field, used for the analysis, consists in downgoing plane waves with nearly vertical incidence. This assumption is usually valid when dealing with natural sources. However, in urbanized areas, artificial noise could affect MT analysis dramatically, especially when dealing with periods longer than 10^2s (i.e. with depth higher than few kilometers). The main noise sources are trains, whose electromagnetic field is so powerful to disturb MT recording many kilometers away from their path.

Different techniques for MT recording denoising have been proposed. They are usually based on searching for specific signatures in the MT recordings indicating the presence of noise. For instance, recently Escalas et al. (2013) proposed a technique based on the analysis of the polarization of the electric field to detect and remove linearly polarized portions of MT signals, which are likely to be affected by noise due to human settlements and infrastructures therein.

In this work a novel technique based on the clustering of impedance tensors, computed on a Discrete Wavelet (DW) expansion of MT recordings, is proposed. The DW transform allows analyzing MT recordings both in time and frequency domains [11].

In the following, the data processing performed on synthetic and real data is described and the applied clustering technique is illustrated. Finally the SOM results are discussed.

2   Data Processing

The MT method has been applied on synthetic data generated by an automatic procedure, assuming the use of a single recording station. The signals are corrupted by adding different noise levels in order to provide different noise conditions. As an example, figure 3 shows a MT synthetic noisy signal and its components in the magnetic and electric filed. Figure 4 visualizes instead a real MT signal recorded at the volcano Etna, Sicily.

To exploit the information about the phase difference of the signals we use their analytic representation:

$$A = s(t) + iH[s(t)]$$

where $s(t)$ is the time domain signal, H is the Hilbert transform and i is the imaginary unit. Once DW transformed, the electric and magnetic components of the MT signal are processed separately for each wavelet scale. Given a set of coefficients

the impedance tensor is determined through the least squares solution of the linear system of equations in (1):

$$E_{xi} = Z_{xx} H_{xi} + Z_{xy} H_{yi}$$
$$E_{yi} = Z_{yx} H_{xi} + Z_{yy} H_{yi},$$

(1)

where the index i runs over the DW coefficients of the analytic representation of the fields for a given wavelet scale. However, the application of the previous procedure to a noisy dataset is likely to lead to unreliable results.
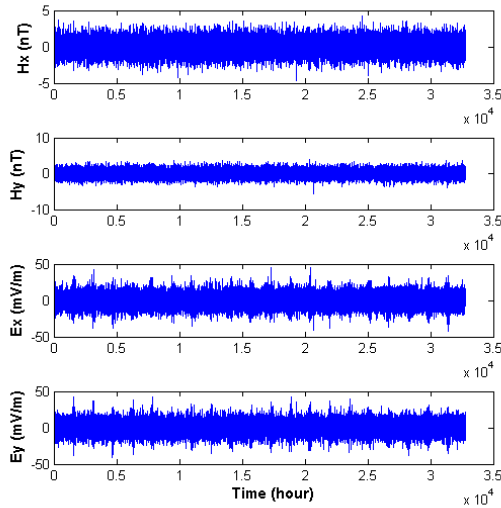
Fig. 3. An example of a synthetic MT signal where noise (N/S 50%) has been added to almost 30% of it. *Hx* and *Hy* are the components of the magnetic field while *Ex* and *Ey* the components of the electric field.

Being the most relevant noise sources transient, the basic idea of the method is to apply a clustering procedure of the retrieved impedance tensors over subsets of the DW coefficients for each wavelet scale. A set of 2^s DW coefficients for a wavelet scale s can be partitioned in different ways. We have applied a Monte Carlo technique selecting N random subsets of k coefficients from the whole set. For each set we applied the least square approach of eq. (1) to determine a set of impedance tensors.

Finally, the obtained impedance vectors have been normalized using a logistic transformation which scales all possible values between [0,1] before being processed by the SOM.
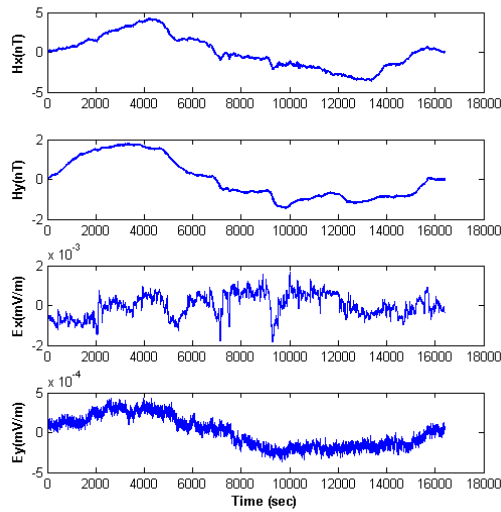
Fig. 4. A real MT signal recorded at volcano Etna (Sicily). As for the synthetic one, *Hx* and *Hy* and *Ex* and *Ey* are the components of the magnetic and electric field respectively.

3   SOM-Based Clustering

The clustering process is primarily carried out for the purpose of grouping data with similar features. In this works, among the existing methods of cluster analysis [8], the Self-Organizing Map (SOM) [10,15,16] technique was selected, and motivations are reported in the following. Compared to a standard iterative partitioning method such as K-means, the SOM is similar in many respects. However, as suggested in [8] its use is preferable to the K-means algorithm, since it does not require to fix the number K of clusters in advance, which would not be possible in our case since no prior signal's knowledge is available [21]. In addition, the K-means algorithm is sensitive to noisy data and outliers and it does not provide an immediate interpretation of the obtained results.

In contrast, with appropriate training parameters, the SOM algorithm produces a low-dimensional plot as a visual representation of the clustering and can handle very large and complex data sets. For these reasons it has been proposed (see Bação et al., 2005) as the most convenient clustering method and as reliable substitute for the classic K-means. Furthermore, in [23] is reported a comparison between the neural approach and a number of more conventional clustering methods where it has been shown that the SOM network performance was similar to or better than that obtained through the other methods. Finally, several comparisons between SOM and K-means performance have been reported [2,19,23]. Conclusions seem to be ambivalent as different authors point to different conclusions, and no definitive results have emerged. Some authors [2,9,23] suggest that SOM performs equal or worst than

statistical approaches, while other authors conclude the opposite [19, 20]. As pointed out by Everitt et al. (2001) it should be evident that there is not a clustering method suitable for all cases, but particular methods will be best for particular types of data and application areas.

The SOM technique has already been applied in Geophysics, for the analysis and clustering of seismic data [3,5,6,7,13,18,12,14,17].

In our tests, the SOM Toolbox package for Matlab [15] has been considered. The batch training algorithm was chosen between the two available iterative versions, being much faster [15,16]. In the batch mode, the whole data set is presented to the SOM. In each training step, the data set is partitioned according to the Voronoi regions of the map weight vectors. Then, the new weight vectors are calculated as:

$$m_i(t+1) = \frac{\sum_{j=1}^{n} h_{ic}(t)x_j}{\sum_{j=1}^{n} h_{ic}(t)} \qquad (2)$$

where c is the index of the BMU of sample vector $\mathbf{x}_j$. The new weight vector is a weighted average of the data samples, where the weight of each data samples is the neighbourhood function value $h_{ci}(t)$ at its BMU c.

The net parameters have been set according to Kohonen et al. (1996). Before the training, a random initialization at small values of the prototypes has been adopted in order to exhibit the self-organizing capability of the SOM. The selected map topology presents a local hexagonal structure and a global toroid shape shown as a sheet paper to have a direct visual interpretation of the cluster configuration. Then, the *Gaussian neighborhood function* has been chosen, which reveals how strongly neurons are connected to each other. Such function is expressed as:

$$h_{c,i} = \eta_t * \exp(-d_{c,i}^2 / 2\sigma_t^2) \quad (3),$$

where $d_{c,i}$ is the distance between the positions of the units c and i on the map grid, while η_t and σ_t are the *learning rate* and the *neighborhood radius* at step t respectively. The first one controls the attraction strength of the input vector, while the second regulates the number of the attracted vectors other than the winning node. Both parameters are time-decreasing functions and change their values during training. Thus, as suggested in [10, 16], η_t starts with a value close to 1 then decreases until a value close to zero. In the same way, initially σ_t begins with a large value in order to include all neurons for any winning node and successively it decreases until to 1. Both of these values for η_t and σ_t remain constant during the convergence phase. In this way, initially a rough representation of the input data distribution is provided by the map, while at the end the prototypes are settled to their final values and the final map is shown.

The SOM has been applied on the impedance tensors computed on a Discrete Wavelet (DW) expansion of MT recordings. For this purpose, each tensor (a 2x2 complex matrix) has been parameterized as a vector of 8 real elements. For each

wavelet scale (i.e. each frequency band) the SOM clustering has been performed. Consistent clusters are formed only by the noise-free portions of the signals.

4 Results

For the impedance tensors clustering, a SOM map with 16 (4x4) nodes has been used, with an hexagonal structure and a toroid shape, plotted as a sheet in order to represent immediately the clusters' structure. Figure 5 shows, as an example, one of the 12 resulting SOM maps obtained from synthetic data under examination.
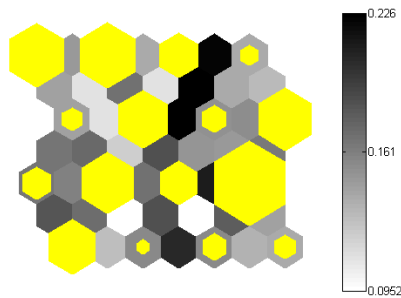
Fig. 5. An example of 4x4 SOM map for the wavelet scale with a characteristic period of $5x10^4$s. The yellow hexagons are the individual nodes (clusters) while the gray ones indicate the Euclidean distances between the nodes according to the gray level scale on the right. The node size specifies the number of signals in it included.

The yellow hexagons indicate the individual nodes and their size represents the data density i.e. how many signals each node contains. The gray hexagons describe the Euclidean distances between the nodes using a gray level scale as reported in [15,16].

Figures 6 and 7 illustrate the results obtained by denoising the same synthetic noisy MT signal with a conventional processing algorithm and the proposed SOM method respectively. Blue curves are the pseudo-resistivity curves used to compute the synthetic data. For this particular synthetic MT signal, noise (N/S 50%) has been added to almost 30% of it. In Figure 6 it can be seen that the conventional method undergoes to significant errors when periods are greater than 10^4s. Figure 7 shows the pseudo-resistivity curve after the application of the SOM clustering procedure to the same MT signal of figure 6. It is possible to observe how the SOM results are more consistent even for higher periods.

Finally, figure 8 illustrates the results obtained on a real MT signal, recorded at the volcano Etna, both with classical preprocessing (cyan and magenta dots) and with the SOM (red and blue circles). It is possible to observe from the figure how the Ryx component related to the first method is more noisy and presents unrealistic oscillations, while the SOM graphic shows a much more plausible trend.
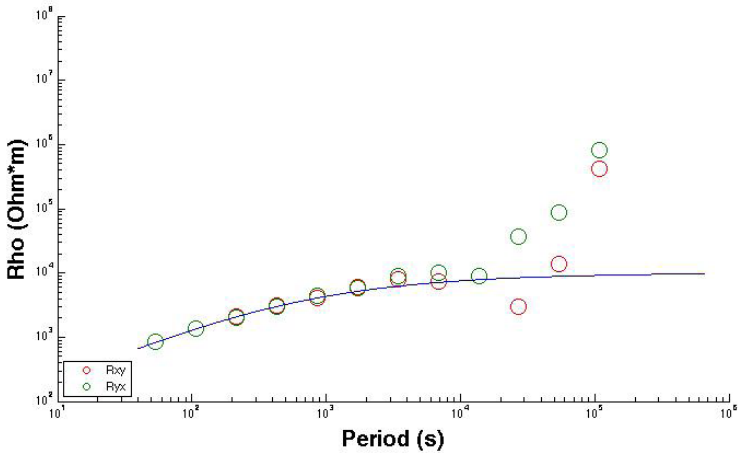
Fig. 6. The results obtained with the conventional processing for a synthetic signal with a noise (N/S 50%), applied over about 30% of the signal. The blue curve is the pseudo-resistivity model. *Rxy* (red circles) and *Ryx* (green circles) indicate the resistivity computed using the two antidiagonal component of the impedance tensor.
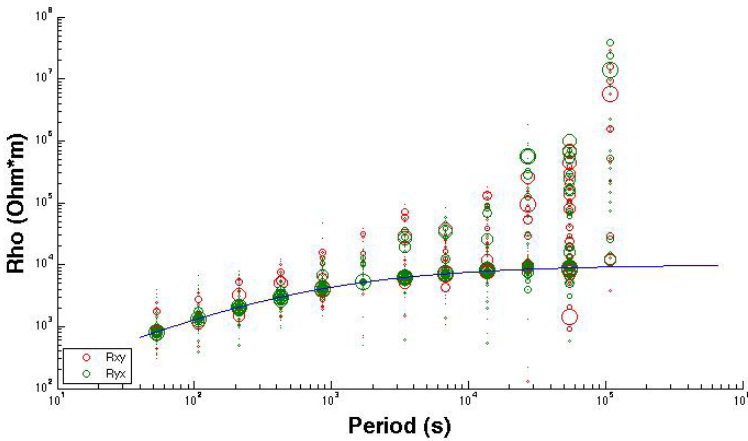
Fig. 7. The results obtained after applying the SOM clustering procedure to the same data reported in figure 6.
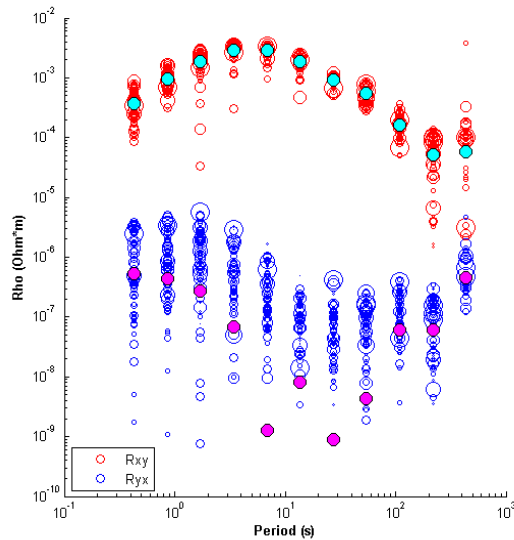
Fig. 8. The results obtained on a real MT signal, recorded at volcano Etna, both with the conventional processing (cyan and magenta dots) and with the SOM (red and blue circles). The Ryx component associated to the first method is clearly more noisy than that relating to the SOM that exhibits a more consistent behavior.

5   Conclusions

A new method to process the magnetotelluric recordings affected by noise due to human settlements and infrastructures therein has been presented. It is based on the use of a neural network, the SOM [15,16], able to cluster the data identifying noise-free intervals in the recordings. The method has been tested on synthetic data and, as a further validation, on a real data, recorded at volcano Etna (Sicily). A Discrete Wavelet transform was applied to the data in order to analyze MT recordings both in time and frequency domain. The MT signal is generally very complex and it is not easy to indentify the noise-free intervals when no a priori information on it is available.

Observing the SOM results and comparing them with those obtained through a conventional processing, it is possible to conclude that the proposed method shows a greater capability of reducing the noise's effect on the retrieved pseudo-resistivity curves, especially for longer periods (i.e. greater investigation depths). Thus, the SOM-based clustering analysis is able to identify tensors related to noise-free intervals.

Finally, the proposed method can be considered as novel and pioneering since it is the first in literature to use neural networks for denoising MT signals offering performance comparable or even better than other available denoising techniques [4].

References

1. Bação, F., Lobo, V., Painho, M.: Self-organizing Maps as Substitutes for K-Means Clustering. In: Sunderam, V.S., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) ICCS 2005. LNCS, vol. 3516, pp. 476–483. Springer, Heidelberg (2005)
2. Balakrishnan, P.V., Cooper, M.C., Jacob, V.S., Lewis, P.A.: A study of the classification capabilities of neural networks using unsupervised learning: a comparison with k-means clustering. Psychometrika 59(4), 509–525 (1994)
3. Calabrese, L., Campanella, G., Proverbio, E.: Use of Cluster Analysis of Acoustic Emission Signals in Evaluating Damage Severity in Concrete Structures. J. Acoustic Emission 28 (2010)
4. Escalas, M., Queralt, P., Ledo, J., Marcuello, A.: Polarisation analysis of magnetotelluric time series using a wavelet-based scheme: A method for detection and characterisation of cultural noise sources. Physics of the Earth and Planetary Interiors 218, 31–50 (2013)
5. Esposito, A.M., Scarpetta, S., Giudicepietro, F., Masiello, S., Pugliese, L., Esposito, A.: Nonlinear Exploratory Data Analysis Applied to Seismic Signals. In: Apolloni, B., Marinaro, M., Nicosia, G., Tagliaferri, R. (eds.) WIRN 2005 and NAIS 2005. LNCS, vol. 3931, pp. 70–77. Springer, Heidelberg (2006)
6. Esposito, A.M., Giudicepietro, F., D'Auria, L., Scarpetta, S., Martini, M., Coltelli, M., Marinaro, M.: Unsupervised neural analysis of very-long-period events at Stromboli volcano using the self-organizing maps. BSSA 98, 2449–2459 (2008), doi:10.1785/0120070110.
7. Esposito, A.M., D'Auria, L., Giudicepietro, F., Caputo, T., Martini, M.: Neural analysis of seismic data: applications to the monitoring of Mt. Vesuvius. Special Issue "Mt. Vesuvius monitoring: the state of the art and perspectives", Annals of Geophysics 56(4), S0446 (2013), doi:10.4401/ag-6452
8. Everitt, B., Landau, S., Leese, M.: Cluster Analysis. Oxford University Press, New York (2001)
9. Flexer, A.: On the use of self-organizing maps for clustering and visualization. In: Żytkow, J.M., Rauch, J. (eds.) PKDD 1999. LNCS (LNAI), vol. 1704, pp. 80–88. Springer, Heidelberg (1999)
10. Floreano, D., Matiussi, C.: Manuale sulle Reti Neurali. Edizione Mulino, Bologna (1995)
11. Foufoula-Georgiou, E., Kumar, P.: Wavelets in Geophysics, 373 pages. Academic Press, San Diego (1994) ISBN: 978-0-12-262850-4
12. Marroquín, I.D., Brault, J.-J., Hart, B.S.: A visual data-mining methodology for seismic-facies analysis: Part 1 - Testing and comparison with other unsupervised clustering methods. Geophysics 74(1), 1–11 (2009), http://dx.doi.org/10.1190/1.3046455, doi:10.1190/1.3046455
13. Köhler, A., Ohrnberger, M., Scherbaum, F.: Unsupervised pattern recognition in continuous seismic wavefield records using Self-Organizing Maps. Geophys. J. Int. 182, 1619–1630, Journal compilation © 2010 RAS (2010), doi:10.1111/j.1365-246X.2010.04709.x
14. Köhler, A., Chapuis, A., Nuth, C., Kohler, J., Weidle, C.: Autonomous detection of calving-related seismicity at Kronebreen, Svalbard. The Cryosphere 6, 393–406 © Author(s) CC Attribution 3.0 License (2012), http://www.the-cryosphere.net/6/393/2012/, doi:10.5194/tc-6-393-2012

15. Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J.: SOM_PAK: The self-organizing map program package. Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo, Finland (1996),
    http://www.cis.hut.fi/research/som_lvq_pak.shtml
16. Kohonen, T.: Self-Organizing Maps, 2nd edn. Series in Information Sciences, vol. 30. Springer, Heidelberg (1997)
17. Marcilio Castro de Matos, M., Osorio, P.L.M., Johann, P.R.S.: Unsupervised seismic facies analysis using wavelet transform and self-organizing maps. Geophysics 72(1), 9–21 (2007)
18. Marciclio Castro de Matos, M.C., Marfurt, K.J., Johann, P.R.S.: Seismic interpretation of Self-organizing maps using 2D color displays. Revista Brasileira de Geofísica 28(4), 631–642 © 2010 Sociedade Brasileira de Geofísica (2010), http://www.scielo.br/rbg
19. Openshaw, S., Openshaw, C.: Artificial intelligence in geography. John Wiley & Sons, Chichester (1997)
20. Openshaw, S., Blake, M., Wymer, C.: Using neurocomputing methods to classify Britain's residential areas. In: Fisher, P. (ed.) Innovations in GIS, vol. 2, pp. 97–111. Taylor and Francis (1995)
21. Rokach, L., Maimon, O.: Clustering Methods. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook, 2nd edn., vol. XX, 1285 p. Springer, New York (2010), doi:10.1007/978-0-387-09823-4, ISBN 978-0-387-09823-4
22. Vozoff, K.: The magnetotelluric method in the exploration of sedimentary basins. Geophysics 37(1), 98–141 (1972)
23. Waller, N.G., Kaiser, H.A., Illian, J., Manry, M.: A comparison of the classification capabilities of the 1-dimensional Kohonen neural network with two partitioning and three hierarchical cluster analysis algorithms. Psychometrika 63(1), 5–22 (1998)

Integration of Audio and Video Clues for Source Localization by a Robotic Head

Raffaele Parisi, Danilo Comminiello, Michele Scarpiniti, and Aurelio Uncini

DIET Dept., University of Rome "Sapienza", Rome, Italy
`raffaele.parisi@uniroma1.it`

Abstract. In this work the first step of an integration process between audio and video information for the localization of speakers in closed environments is presented. The proposed metod is based on binaural source localization followed by face recognition and tracking and was realized and implemented in a real environment. Some preliminary results demonstrated the effectiveness of this approach.

Keywords: Binaural source localization, face detection and tracking, audio and video integration.

1 Introduction

Binaural localization consists in estimating the position of a sound source in a generic environment by use of a single pair of microphones. This approach gets inspiration from biological organisms, where the auditive system works by integrating information acquired by the body, the outer ear and the inner ear [1].

Different models of binaural localization are available [2]. A popular approach is based on combined use of *Interaural Level Difference* (ILD) and *Interaural Time Difference* (ITD) [3]. These cues can separately give information about the source position in different range of frequencies and can be fruitfully combined so as to generate an effective binaural localization algorithm [3].

The exploitation of audio signals is just one side of a localization system based on proper integration of audio and video clues. As a matter of fact, in biology the two senses of hearing and vision cooperate in order to augment the information acquired on the surrounding environment. Of course this is a fundamental task, both for hunting and for escaping from hunters.

Some works exist that deal with the fusion of audio and video signals at different levels and for different applications [4] [5] [6] [7] [8]. As far as we know, there are not works explicitly dealing with the topic of integration of binaural audio signals and video signals.

In this paper some preliminary results toward effective integration of audio and video signals in a robotic head are described. Fig. 1 shows the robotic head that was realized in the ISPAMM Laboratory of the DIET Dept. at the University of Rome "La Sapienza". The device is equipped with two omnidirectional microphones and two cameras. Two stepper motors can rotate the head

and move the eyes. These stepper motors are controlled using the Arduino Uno board. The Arduino Uno is a very common microcontroller board: it has 14 digital input/output pins that can be used to control some external devices, and a USB connection, used to load the control software from a personal computer.

In this preliminary setup, two main tasks were implemented:

1. a binaural source localization procedure. The joint ILD/ITD estimation was employed to localize the speaker in terms of angular distance from the center. The main issue of this approach is the presence of reverberation, that actually reduces the accuracy of the estimation.
2. A face detector/tracking procedure. It is possible to find and to track a human face in images captured by the cameras. In this way it is possible to track the movement of the speaker and to correct the sound localization errors due to reverberation.

Experimental results in a real environment demonstrated the effectiveness of this preliminary idea, as a first step toward a full integration of audio and video information. In the following the main steps of the developed procedure are described.
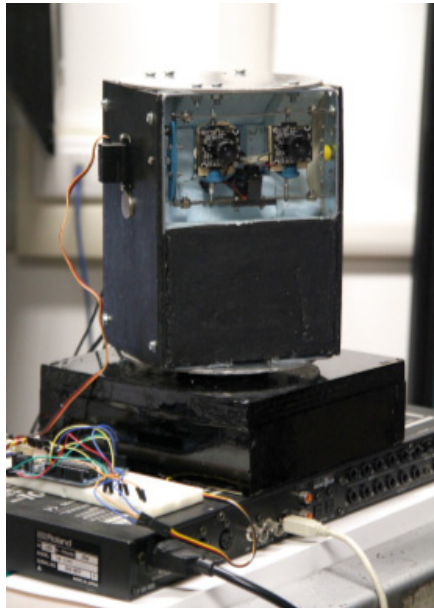
Fig. 1. The artificial head described in this paper

2   Description of the Audio System

In this section we briefly recall the main concepts of localization of audio sources by binaural processing. Binaural perception was studied by Lord Rayleigh at the

beginning of the $20th$ century [1]. From that time on several models of the human binaural system have been proposed. An estensive description was presented in [9].

Binaural localization can be realized by using the Interaural Level Difference (ILD) and the Interaural Time Difference (ITD) in a joint way. ILD is proportional to the difference in the sound levels reaching the left and right ear, while ITD is the measure of the time difference of arrival of a signal to each ear. These cues can be used to obtain information about the source position in different ranges of frequencies. In fact, independent use of ILD and ITD does not yield robust source position estimators [3], since ITD is affected by ambiguity due to an *a priori* unknown phase unwrapping factor, while ILD estimates display a significant standard deviation. Localization of sources can be realized by properly combining ILD and ITD. In the following we briefly describe a possibile approach [3].

The binaural model of received signals is

$$x_l[n] = h_l[n] * s[n] + \eta_l[n], \tag{1}$$
$$x_r[n] = h_r[n] * s[n] + \eta_r[n], \tag{2}$$

where l and r refer to the left and right ear respectively. In this equation $h_i[n]$ $(i = l, r)$ is the impulse response, $s[n]$ is the source signal while $\eta_i[n]$ represent an additive uncorrelated noise term. In the following description noise will be considered negligible, a simplifying assumption which is true in many practical situations.

As in [3], ILD and ITD for the generic n-th time-frame are

$$ILD^n(\omega, \theta, \phi) = 20log_{10} \left| \frac{X_r^n(\omega, \theta, \phi)}{X_l^n(\omega, \theta, \phi)} \right|, \tag{3}$$

$$ITD^n(\omega, \theta, \phi) = \frac{1}{\omega} \left(\angle \frac{X_r^n(\omega, \theta, \phi)}{X_l^n(\omega, \theta, \phi)} + 2\pi p \right). \tag{4}$$

In these equations ω is frequency, θ and ϕ are the elevation and azimuth angles respectively, $X_r^n(\omega, \theta, \phi)$ and $X_l^n(\omega, \theta, \phi)$ are the Short Time Fourier Transforms (STFTs) of the right and left ear signals and p is the *phase unwrapping factor*, which is unknown *a priori* and needs to be estimated.

The new joint ILD and ITD localization method [3] is based on comparison between the particular estimated pair (ILD, ITD) and a reference set of pairs contained in a data lookup matrix. This matrix is constructed by exploiting the fact that Head Related Transfer Functions (HRTFs) are stationary and can be used in calculating two different ITD and ILD reference sets that depend on azimuth and frequency alone. Equations (3) and (4) in this case can be written as

$$ILD(\omega, \phi) = 20log_{10} \left| \frac{HRTF_r(\omega, \phi)}{HRTF_l(\omega, \phi)} \right|, \tag{5}$$

$$ITD(\omega, \phi) = \frac{1}{\omega} \left(\angle \frac{HRTF_r(\omega, \phi)}{HRTF_l(\omega, \phi)} + 2\pi p \right). \tag{6}$$

In these equations $HRTF_r$ and $HRTF_l$ are the HRTF functions on the right and left ears respectively. By assumption the value of the unwrapping factor p does not change dramatically across azimuth [3]. Smoothing across azimuth with a constant Q filter was performed on the ILD lookup set in order to better represent the limits of human interaural level difference perception. More specifically, a Gaussian filter was employed, as indicated in the CIPIC database [10].

Comparison between the ILD and ITD lookup sets and the estimated ILD and ITD allows to estimate the azimuth of the sound source. In particular ILD is exploited to find the correct value of the unwrapping factor p and to select the azimuth value minimizing the difference between the ITD-only and ILD-only estimates. This p-estimation procedure was repeated for each available time frame. A time average across frames was performed and the results graphed. The final azimuth estimations selected were those displaying a minimum in the difference function that was consistent across frequencies.

As an example, fig. 2 shows the results obtained in simulations with the source placed at different azimuth angles. Joint exploitation of ILD and ITD allows to obtain an azimuth estimate which is correct over the whole frequency band and for different positions of the source.
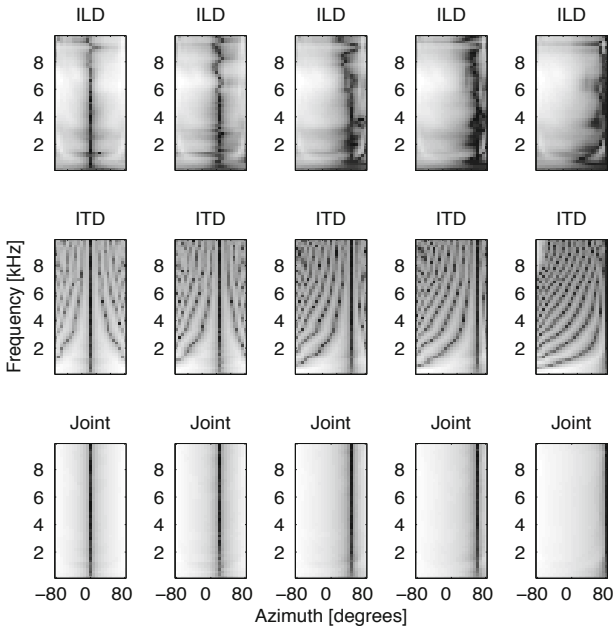
Fig. 2. Source azimuth estimate in an anechoic room and Gaussian noise: ILD, ITD and joint ILD-ITD methods. Columns from left to right refer to source azimuth angles of 0°, 20°, 45°, 60° and 80° respectively. Darkest pixels are lowest in value.
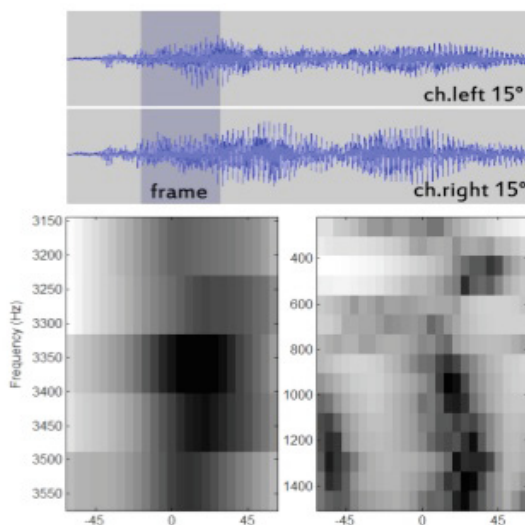
Fig. 3. Source azimuth estimate in a real room and a female speaker placed at the azimuth angle of 15°: ILD and ITD estimates in different ranges of frequencies.

Fig. 3 shows the results obtained in a real environment with a female speaker, speaking from an azimuth angle of 15°, in terms of the ILD and ITD estimates in different ranges of frequencies. Slight reverberation is present. It is clear that in the presence of reverberation [11], commonly assumed in closed environments, proper prefiltering techniques should be adopted [12] [13] [14]. An example is cepstral prefiltering [15].

3   Description of the Video System

In this preliminary study, the video information was used for localizing and tracking the head of a speaker, after she/he has been localized by using the audio information. The main task in this process is the localization of the face of the speaker in the image acquired.

3.1   Face Detection

The face recognition task was realized by using the *Viola-Jones method* [16]. This technique was one of the first methods introduced for detecting the presence of objects in images and it is currently used for the detection of faces. It is based on classification of specific features rather than on the intensity values of the image pixels. Namely the steps of the classification process are:

1. extraction of *Haar features*. Haar features are basically determined by computing the sum and/or the differences of the pixels within two rectangular regions of the image.

2. Construction of the *integral image*. The integral image is an intermediate representation of the original image. Namely, the generic point (x, y) of the integral image is defined as the sum of the pixels above and to the left of (x, y).

3. *AdaBoost.* The AdaBoost (short for *Adaptive Boosting*) is a machine learning meta-algorithm used to improve the performance of learning algorithms [17]. It is based on the combination of various weak classifiers in order to obtain a final robust classifier and it is employed in the Viola-Jones method

4. *Chain classifier.* The Viola-Jones method is based on a cascade of AdaBoost classifiers in order to classify portions of images. As a consequence of this processing phase, the performance of the detection task is increased, while reducing the computation time required.

3.2　Face Tracking

Once the region containing the face has been detected, the next step is to move the image of the face to the central position of the video image. This task can be realized by a feedback loop where a pair of proportional controls is employed to progressively reduce the difference between the position of the detected face and the center of the video image. To this goal, the *tilt* and *pan* angles of the head are used. Figure 4 shows the scheme of the head control unit.
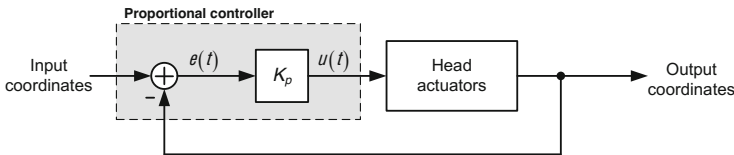
Fig. 4. The head control loop

4　Experiments

The head was equipped with two omnidirectional microphones AKG C562M. Signals were acquired through an Edirol UA-1000 acquisition board. Figure 5 shows the configuration of the testbed, with five possible positions of the source.

The control of the servomotors was realized by an Arduino board[1]. The face-tracking algorithm was written in C++ by using the functions available at the OpenCV website[2]. Figure 6 shows in detail the Arduino board used for processing of the video part.

The face recognition and tracking algorithm was used to localize the face of the speaker after the audio localization task and to move it to the center of the

---

[1] `www.arduino.cc`
[2] `www.opencv.org`
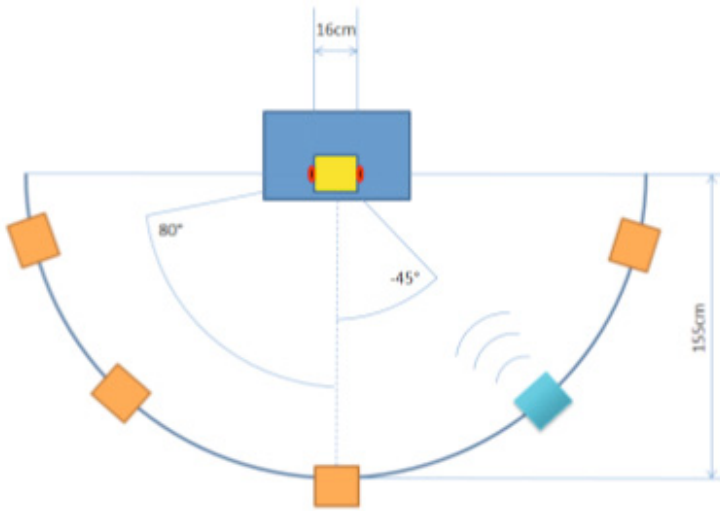
Fig. 5. Testbed configuration
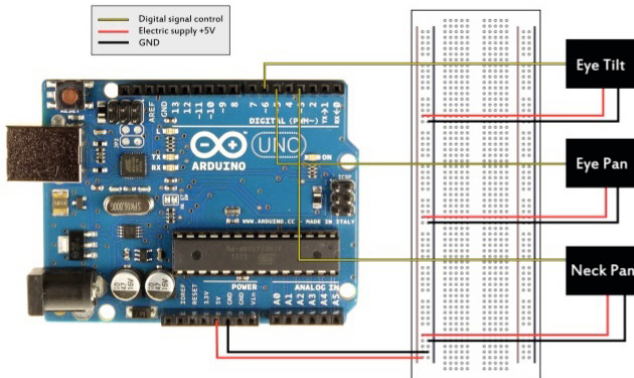
Fig. 6. The Arduino board and its connections
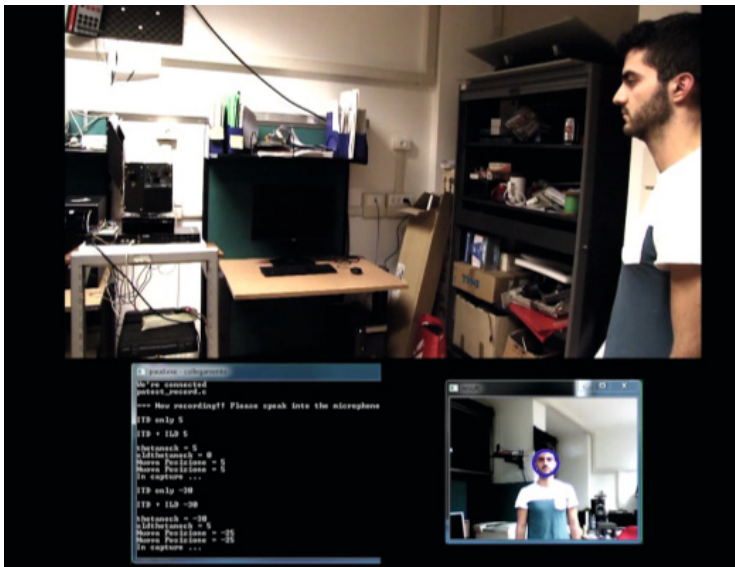
Fig. 7. A single frame of the localization process

image. As an example of the experimental results, fig. 7 shows a single frame of the output video. The artificial head can be seen in the upper part of the image, together with the speaker. In the lower part, it is shown the image as acquired by the camera mounted on the head, after the face of the speaker has been moved to the center.

5   Conclusion

In this paper a possible cooperation between binaural audio and video signals was described. The objective was the localization and tracking of an audio source moving in a closed environment. Some preliminary experiments demonstrated the quality of the proposed solution, also in a real environment. Further research will be devoted to pursue full integration of both audio and visual information.

References

1. Rayleigh, L.: On our perception of sound direction. Phil. Mag. 13, 214–232 (1907)
2. Blauert, J.: Spatial Hearing - The Psychophysics of Human Sound Localization. MIT Press (1996)
3. Raspaud, M., Viste, H., Evangelista, G.: Binaural source localization by joint estimation of ILD and ITD. IEEE Trans. on Audio, Speech and Language Processing 18(1), 68–77 (2010)
4. Monaci, G., Jost, P., Vandergheynst, P., Mailé, B., Lesage, S., Gribonval, R.: Learning multimodal dictionaries. IEEE Trans. on Image Processing 16(9), 2272–2283 (2007)

5. Zhang, C., Yin, P., Rui, Y., Cutler, R., Viola, P., Sun, X., Pinto, N., Zhang, Z.: Boosting-based multimodal speaker detection for distributed meeting videos. IEEE Trans. on Multimedia 10(8), 1541–1552 (2008)

6. Schmalenstroeer, J., Haeb-Umbach, R.: Online diarization of streaming audio-visual data for smart envirnments. IEEE Journ. of Selected Topics in Signal Processing 4(5), 845–856 (2010)

7. Naqvi, S.M., Wang, W., Khan, M.S., Barnard, M., Chambers, J.A.: Multimodal (audio-visual) source separation exploiting multi-speaker tracking, robust beamforming and time-frequency masking. IET Signal Processing 6(5), 466–477 (2012)

8. Minotto, V.P., Jung, C.R., Lee, B.: Simultaneous-speaker voice activity detection and localization using mid-fusion of svm and hmms. IEEE Trans. on Multimedia 16(4), 1032–1044 (2014)

9. Wang, D., Brown, G.J.: Computational Auditory Scene Analysis - Principles, Algorithms, and Applications. IEEE Press, Wiley Interscience (2006)

10. Algazi, V.R., Duda, R.O., Thompson, D.M., Avendano, C.: The CIPIC HRTF database. In: 2001 IEEE Workshop on Applications of Digital Signal Processing to Audio and Acoustics (2001)

11. Kuttruff, H.: Room Acoustics, 4th edn. Taylor & Francis (2000)

12. Stéphenne, A., Champagne, B.: A new cepstral prefiltering technique for estimating time delay under reverberant conditions. Signal Processing 59(3), 253–266 (1997)

13. Parisi, R., Gazzetta, R., Di Claudio, E.: Prefiltering approaches for time delay estimation in reverberant environments. In: Proceedings of ICASSP, vol. 3, pp. III-2997–III-3000 (2002)

14. Zannini, C.M., Parisi, R., Uncini, A.: Binaural sound source localization in the presence of reverberation. In: Proc. of the 17th International Conference on Digital Signal Processing (July 2011)

15. Parisi, R., Camoes, F., Scarpiniti, M., Uncini, A.: Cepstrum prefiltering for binaural source localization in reverberant environments. IEEE Signal Processing Letters 19(2), 99–102 (2012)

16. Viola, P., Jones, M.J.: Robust real-time face detection. Int. J. of Computer Vision 57(2), 137–154 (2004)

17. Freund, Y.Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 55(1), 119–139 (1997)

A Feasibility Study of Using the NeuCube Spiking Neural Network Architecture for Modelling Alzheimer's Disease EEG Data

Elisa Capecci[1,*], Francesco Carlo Morabito[2], Maurizio Campolo[2],
Nadia Mammone[2], Domenico Labate[2], and Nikola Kasabov[1]

[1] Auckland University of Technology - Knowledge Engineering and Discovery
Research Institute, Auckland, New Zealand
{ecapecci,nkasabov}@aut.ac.nz
[2] DICEAM - Mediterranea University of Reggio Calabria, Italy
{morabito,campolo,nadia.mammone,domenico.labate}@unirc.it

Abstract. The paper presents a feasibility analysis of a novel Spiking
Neural Network (SNN) architecture called NeuCube [10] for classifica-
tion and analysis of functional changes in brain activity of Electroen-
cephalography (EEG) data collected amongst two groups: control and
Alzheimer's Disease (AD). Excellent classification results of 100% test
accuracy have been achieved and these have also been compared with tra-
ditional machine learning techniques. Outputs confirmed that the Neu-
Cube is better suited to model, classify, interpret and understand EEG
data and the brain processes involved. Future applications of a NeuCube
model are discussed including its use as an indicator of the early onset of
Mild Cognitive Impairment(MCI) to study degeneration of the pathology
toward AD.

Keywords: Spiking Neural Networks, NeuCube, EEG data classifica-
tion, Alzheimer's Disease.

1   Introdution and Problem Specification

During the past few decades, researchers from all-over the world have been con-
centrating their efforts towards understanding of the human brain. As a conse-
quence of the efforts made, a relevant progress has been achieved and a huge
amount of brain data is becoming available. Neuroinformatics researchers have
been playing a pivotal role in the advancement of these studies and especially
with the use of machine learning techniques. Some of the major contributions
are the improvements in the understanding of the Spatio-Temporal Brain Data
(STBD) available and the development of predictive systems. These are of high
importance for society, as the increase in human lifespan has been followed by
the dramatic rise in the appearance of neurological disorders such as AD [18].

---

* Corresponding author.

We have used spatio-temporal EEG as a type of brain data to study this pathology and its degeneration, as it is one of the most commonly collected data for studying the neural processes and it has been for long used to analyse and stage the decline from MCI to AD (e.g. [11,14,15]). Moreover, it is an affordable technique, easy to manage and it is not considered aggressive for the subjects being studied [19].

In this paper, we analyse and classify the spatio-temporal information available (described in section 2) by use of the brain inspired SNN model called NeuCube [10]. In section 3, we introduce the NeuCube model and the experimental design of the study. Section 4 presents the classification results, which are then compared with traditional approaches. Particularly, in section 4.2, through visualization and analysis of the SNN cube (SNNc) after training, new knowledge is also extracted from the data. Finally, conclusions and future directions based on the proposed methodology are presented in section 5.

2   Data Collection and Description

The EEG data has been collected and made available by the IRCCS Centro Neurolesi of Messina, Italy. For this preliminary analysis, we decided to use just the data recorded from one healthy subjects and one subject diagnosed as having AD. The control was a male subject of 58 years of age and the AD patient was a female subject of 80 years of age. They were both at random selected. Each recording session was carried out using 19 electrodes: Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2 and the G2 electrode was used as reference. Electrodes were placed according to the sites defined by the standard $10-20$ international system. Data was recorded for 65 seconds at $256\,\mathrm{Hz}$, resulting in 16640 data points collected per session. A brain computer interface device was used to collect the EEG data, which was recorded under resting condition. During the experiment, the subjects were sitting with the eyes closed and always under vigilant control.

The data was band-pass filtered between 0.5 and $32\,\mathrm{Hz}$, which includes so the relevant bands for AD diagnosis. No further pre-processing of the data was applied, as the NeuCube model is able to accommodate raw data directly; however, screening and selection of the signals that were visually artefact-free was performed prior to data analysis to avoid misleading results. Then, the original EEG signal concatenated was treated to avoid sub-effects related to the inevitable information loss implied by excluding some components.

For this preliminary study, he EEG data was resized into 3 seconds epochs. Thus, for each of the two classes we had 21 samples of 768 data points recorded for every of the 19 EEG channels. In total, we used 42 samples to run the NeuCube experiments.

3   The NeuCube Spiking Neural Network Architecture

This paper evaluates the ability of the NeuCube SNN framework [10] (Fig. 1) to classify and analyse the functional brain activity produced by the EEG data

recorded from a subject affected by AD and a healthy control. This methodology allows for the creation of different models for STBD based on the following information processing principles as listed in [10]:

– The model has a spatial structure that maps approximately the spatially located areas of the brain where STBD is collected.
– The same information paradigm - spiking information processing, that ultimately generates STBD at a low level of brain information processing. This is used in the model to represent and to process this STBD.
– Brain-like learning rules are used in the model to learn STBD, mapped into designated spatial areas of the model.
– A model is evolving in terms of new STBD patterns being learnt, recognised and added incrementally, which is also a principle of brain cognitive development.
– A model always retains a spatio-temporal memory that can be mined and interpreted for a better understanding of the cognitive processes.
– A visualization of the model evolution during learning can be used as a bio-feedback.

Such models can be used to learn and reveal complex spatio-temporal patterns "hidden" in the STBD, which would not be possible to achieve using other information processing methods. As a result a significantly improved understanding of complex brain processes that generates the data can be gained, along with improved classification and/or prediction accuracy.
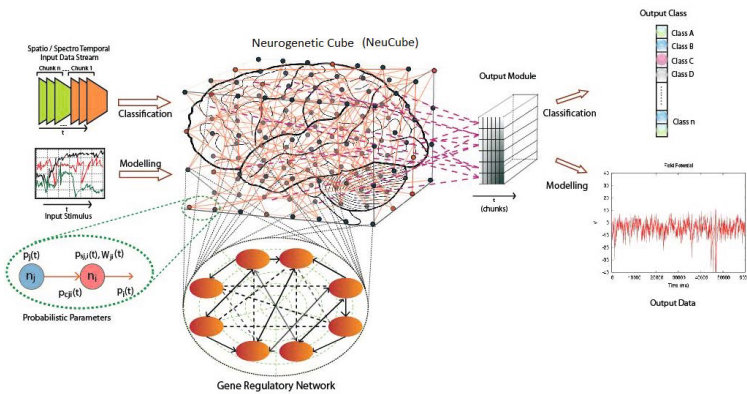
Fig. 1. The NeuCube architecture with its three main modules: input data encoding module; a 3D SNN cube module; an output classification module. Also, an optional Gene Regulatory Network (GRN) module can be incorporated if gene information is available. The spiking neurons can be simple leaky integrate and fire model or probabilistic models (shown in the lower left section).

3.1   Experimental Design and Implementation

The NeuCube-based model used for this study was implemented with a software simulator written in MATLAB [27]. This particular NeuCube consists of three modules:

1. An input information encoding module.
2. The NeuCube 3D SNNc module.
3. An output module for data classification and knowledge extraction.

The process scheme in Fig. 2 summarises the experimental design applied to the study.

I. The raw time series data, obtained from the EEG device, is directly fed into the model as ordered sequence of real-valued data vectors. One of the great advantages of the NeuCube framework is that in many cases there is no need of pre-processing (such as normalization of the data, scaling, smoothing, etc.).

II. Each real value input stream of data is transformed into a spike train using Address Event Representation (AER) method [2]. AER is more convenient when using continuous input data, such as EEG STBD, as this algorithm identify just differences in consecutive values.

III. The spike sequences are then presented to the SNNc, which was implemented using Leaky Integrate and Fire (LIF) neurons [13], as that mimics the information processing of the human brain and it is less computational expensive [20,5]. The SNNc can also evolve according to the number of input variables (i.e. the EEG channels) and the data available.

Due to the size of the data set used for this study, we generated a 3D cube of $13 \times 15 \times 11$ spiking neurons. 1471 of these spiking neurons were mapped according to a brain atlas, the Talairach Atlas [12,24]. Each of these neurons were representing the centre coordinates of a one cubic centimetre area from the 3D Talairach Atlas, including the 19 EEG channels, which also identified the input neurons of the network.

IV. The SNNc is then trained on the input spike trains *via* unsupervised learning method, using Spike Timing Dependant Plasticity (STDP) [23] learning rule. Unsupervised learning is performed to modify the initially set connection weights. The SNNc will learn to activate same groups of spiking neurons when similar input *stimuli* are presented [6]. This makes the NeuCube architecture useful for learning consecutive spatio-temporal patterns and therefore representing a more biologically plausible associative type of memory [10].

V. The output classifier is then trained *via* supervised method. The same STBD used for the unsupervised training is now propagated again through the trained SNNc and output neurons are generated (evolved) and trained to classify the spatio-temporal spiking pattern of the SNNc into pre-defined classes (or output spike sequences). Different SNN methods can be used to learn and classify spiking patterns from the SNNc. For this experimental study, Dynamic Evolving SNN (deSNN) algorithm [9] was used.
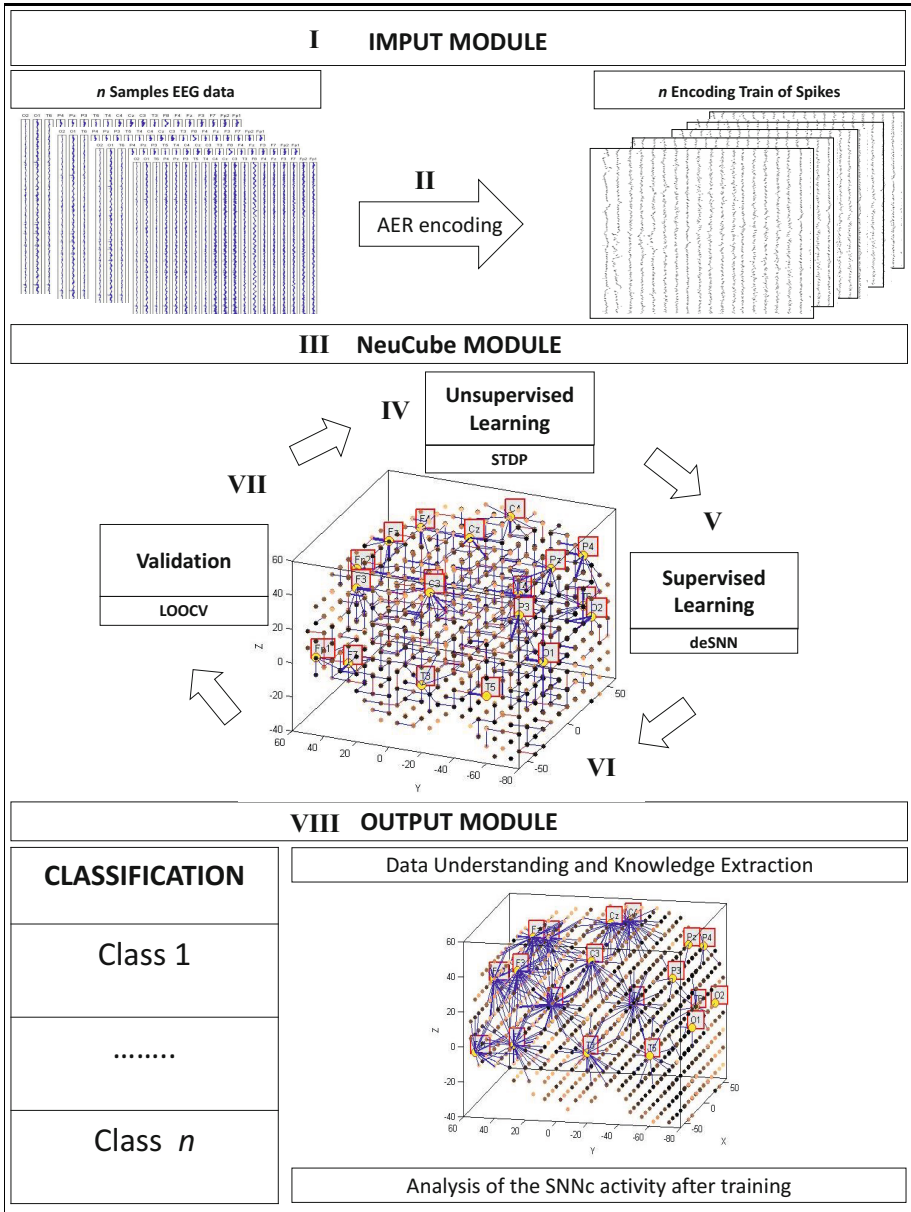
Fig. 2. Process scheme of the NeuCube framework with its three principal modules: the input module, where input data are transformed into trains of spikes that are then presented to the main module, the SNNc; the NeuCube module, where time and space characteristics of the STBD are captured and learned to extract new knowledge from them through the SNNc visualization; and the output module for data classification and understanding. In the scheme are also indicated the VIII processes involved in the NeuCube experiment.

This classification method combines the rank-order learning rule [26] with the STDP [23] temporal learning, for each output neuron to learn a how spatio-temporal pattern using only one pass of data propagation.

VI. The classification results are evaluated using repeated random sub-sampling validation or Leave-One-Out Cross-Validation (LOOCV) respectively.

VII. In order to achieve a desirable classification accuracy, the numerous parameters of the NeuCube needs to be optimized. Therefore, steps (III) to (VI) are repeated changing parameter values. That can be done using a grid search method, a genetic algorithm or the Quantum-Inspired Evolutionary Algorithm [17]. In this study, we have used a grid search, as we will explain in the next section.

VIII. The trained SNNc is visualized and its connectivity and spiking activity analysed for a better understanding of the data and the brain processes that generates it. In fact, it can be observed that new connections are formed between the neurons and this can be further interpreted in the context of different neural activity. Therefore, this represents another key advantage that NeuCube offers: the possibility of knowledge extraction.

4   Results and Discussion

The NeuCube framework has been used and promising results on the analysis of cognitive mental activity [8] and the classification of complex muscular movements for neuro-rehabilitation [25] has been reported. In this paper, we evaluated the feasibility of a NeuCube-based model to correctly classify data with known pattern and extract knowledge from the spatio-temporal EEG signals of a subject affected by AD *versus* a healthy control. Our aim is to develop an analysis and prediction tool to be used by clinician for identifying the appearance of MCI and predict the onset of AD.

To achieve satisfying classification results, the numerous parameters of the NeuCube need to be accurately selected. Based on previous studies that we have conducted (e.g. [8]), we have identified some critical variables requiring careful optimization and we have selected the values that correspond to some of them, making them default. Taking into account that every parameter tuned also involves a considerable amount of processing time, we need to select the proper number of variables to be optimised. The AER threshold was chosen for this study, as it is applied to the entire signal gradient according to the time and therefore the rate of the generated spike trains depend on this threshold. Moreover, since the NeuCube is a stochastic model, altering this value means also altering the initial model configuration each time. Thus, using a grid search, we evaluated the classification accuracy of 10 model configurations adjusting the AER threshold at every new configuration. For that, we have used 50% of the entire time series for training and the other 50% for testing. The parameter's settings which were obtained after optimization are summarised in Tab.1.

Table 1. NeuCube Parameters

NeuCube Parameters			
AER threshold	**Conn. Distance**	**STDP rate**	**Firing threshold**
0,94	0,15	0,01	0,5
Refractory Time	**Training Time**	**deSNN mod**	**deSNN drift**
6	1	0,4	0,25

Table 2. NeuCube's classification results expressed by accuracy percent. Results are obtained using both 50/50%-Trainin/Testing and LOOCV.

NeuCube RESULTS		
CLASS	**(50/50%-Tra/Test.)**	**(LOOCV)**
Control	100%	100%
AD	100%	100%
Average Acuracy	100%	100%

As common practice in machine learning, classification accuracy was calculated by statistical processing of the information obtained from the confusion table. Classifier outputs were evaluated using both random sub-sampling validation and LOOCV, as reported in Tab.2.

4.1 Comparative Analysis

NeuCube results have been also compared with other approaches, such as Multi Layer Perceptron (MLP), Support Vector Machine(SVM), Inductive Evolving Classification Function (IECF)[7] and Evolving Clustering Method for Classification (ECMC)[22].

To process these experiments, the NeuCom platform was used [7], which is a self-programmable, learning and reasoning computer environment freely available on-line (www.theneucom.com).

The LOOCV method was used to evaluate the outputs and datasets were normalised prior to the experiments to ensure the highest classification accuracy result. (i.e. the normalisation protocol applied to each method consisted in a linear standardization of the data's vectors using values between 0 and 1 as a scale).

Classification accuracy was analysed *via* supervised learning method, which is based on classification of data with a known pattern. The results obtained are expressed in the confusion table as number of True Positives (TP) and True Negatives (TN) against False Positive (FP) and False Negative (FN). An analysis of the classification outputs obtained by all different methods was performed based on this information, which was further processed to calculate the following metrics:

– Accuracy percent (A %):

$$A \% = (TP + TN)/(TP + FN + FP + TN) * 100 \qquad (1)$$

– Sensitivity (S):
$$S = TP/(TP + FN) \tag{2}$$

– Specificity (SP):
$$SP = TN/(FP + TN) \tag{3}$$

Results obtained are summarised in Tab. 3 and they were used to plot a Receiver Operating Characteristics (ROC) graph (Fig.3) [3]. The parameter's settings used for each method are reported in Tab. 4.

Table 3. Comparison of the results obtained *via* NeuCube *versus* traditional machine learning methods (MLP, SVM, IECF and ECMC), confusion table and resulted metrics.

Confusion Table						
Control	**AD**	**MLP**		**SVM**		
TP	**FN**	11	14	11	14	
FP	**TN**	10	7	10	7	
IECF		**ECMC**		**NeuCube**		
21	13	21	1	21	0	
0	8	0	20	0	21	

METRICS					
	NeuCube	**MLP**	**SVM**	**IECF**	**ECMC**
A% (1)	100	43	43	69	98
S (2)	1	0.52	0.52	1	1
SP (3)	1	0.33	0.33	0.38	0.95
1-SP	0	0.67	0.67	0.62	0.05

Table 4. Traditional machine learning methods (MLP, SVM, IECF and ECM) parameter's settings

MLP parameters		SVM parameters	
Normalization	yes	**Normalization**	yes
Number of Hidden Units	3	**kernel**	Polynomial
Number of Training Cycles	300	**Degree**,γ	1
Output Value Precision	0.0001		
Output Function Precision	0.0001		
Output Activation Function	linear		
Optimization	scg		
IECF parameters		ECMC parameters	
Normalization	yes	**Normalization**	yes
Max. Influence Field	1	**Max. Influence Field**	1
Min. Influence Field	0.01	**Min. Influence Field**	0.01
M of N	3	**M of N**	3
Membership Function	2		
Epochs	4		
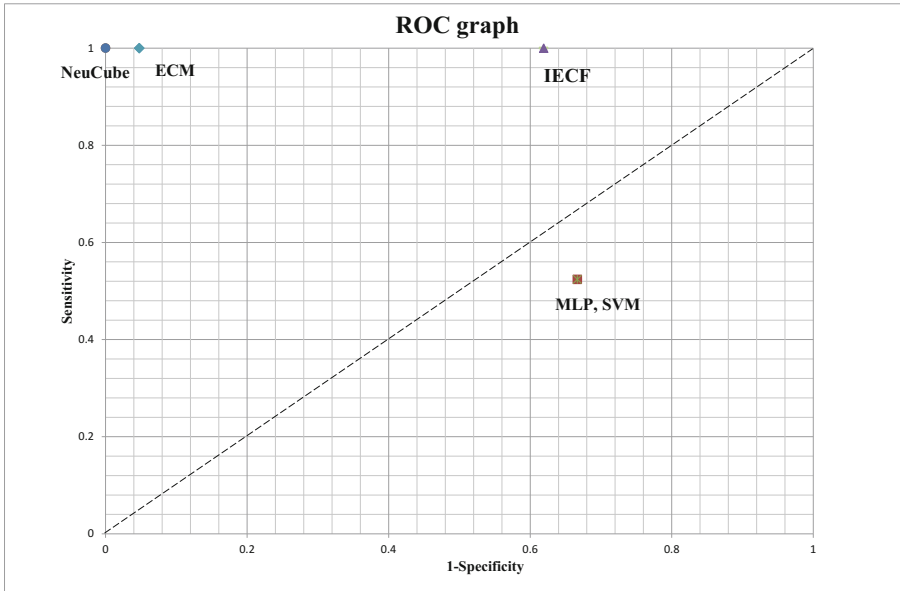
Fig. 3. ROC graph. 1-Specificity is plotted on the X axis and Sensitivity is plotted on the Y axis

As far as the ROC graph is concerned, each classifier produce a sensitivity and a specificity value, which results in a single point on the graph's space. Classifiers falling on the top-left area of the graph are considered achieving desirable results [3]. The NeuCube appears on the top-left hand side of the graph (point 0,1) performing as a perfect classifier. Interesting performance is also reported by ECMC method, while IECF method, even classifying nearly all positives correctly, it reports high false positive, which bring it too far on the right hand side of the graph.

Optimization of the results obtained *via* these techniques is not a trivial process and it requires more sophisticate optimization methods. In fact, the few parameters that influence these methods cannot be tuned independently, one of the reasons is that some of them are discrete values and others are continuous. Thus, the work involved to improve the output results is not viable.

We can conclude that, in terms of the comparison with the other classification methods, NeuCube performed significantly better with the highest accuracy, sensitivity and specificity over all. By means of these metrics, the closest to NeuCube's results was ECMC, whilst the poorest performing were MLP and SVM.

In addition to the above, the NeuCube-based model has other important benefits, such as:

– It requires only one iteration data propagation for learning, while traditional methods as SVM requires numerous iterations.
– The NeuCube-based model is adaptable to new data and new classes, while the other models are fixed and difficult to adapt on new data.
– There is no need of pre-processing of the data (such as normalization, scaling, smoothing, etc.) with the NeuCube model. The raw data can be fed directly into the model as time series transformed into spike trains.
– The NeuCube model demonstrated to be able to achieve a better classification accuracy per class than the other methods.
– The NeuCube model also offers a better understanding of the data and therefore the brain processes that generates it through visualization and analysis of the output SNNc state, as discussed in the following section.

4.2   Model Interpretation and Data Understanding

The NeuCube model constitutes a SNN environment based on some of the most important principles governing the neural activity in the human brain.
Thus, it constitutes a valuable model for on-line learning and recognition of STBD. It also takes into account data features, offering a better understanding of the information and the phenomena of study. In fact, one of the main advantages of the NeuCube model is that after training the SNNc can be visualized and its connectivity and spiking activity observed. This ability of the NeuCube models allows us to trace the development/decline of neurological processes over time and to extract new information and knowledge about them.

Illustrated in Fig. 4 is the SNNc state obtained after it was trained with data from a control subject (top picture) and then after it was trained with data from the subject affected by AD (bottom picture).

We can observe that new connections are formed between the neurons of the network and especially around the input neurons, which were mapped according to the Talairach coordinates of the 19 EEG electrodes. We can depict from Fig. 4 that the neural activity of both the healthy subject and the subject suffering from AD is quiet different. In fact, in the case of the healthy control, the connections evolved are equally distributed in every brain region. On the other hand, in the case of the patient affected by AD, we can observe that this activity decreased in the left hemisphere and so there is a higher activity evolved in the right hemisphere, maybe to compensate the lack of its counterpart and therefore as a consequence of the degeneration of the pathology.
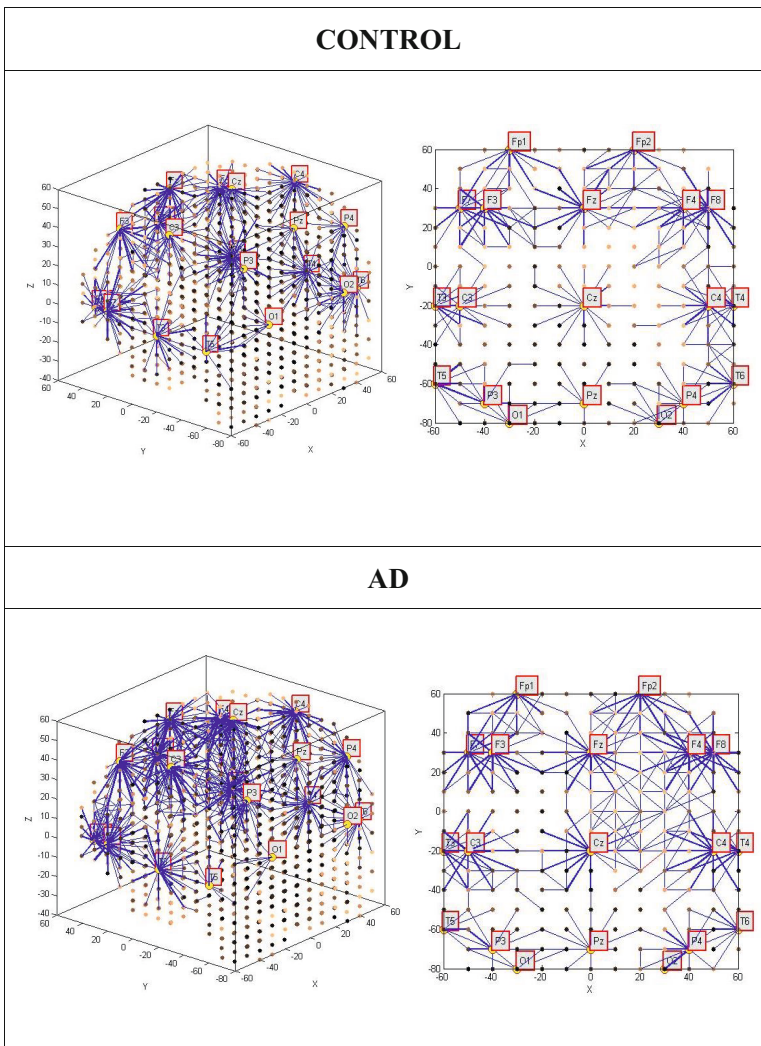
Fig. 4. The SNNc connectivity after training (top control, bottom AD). The figure shows both the 3D cube and the (x,y) plane only of the SNNc. The SNNc can be analysed and interpreted for a better understanding of the EEG data to identify differences between brain states. Blue lines are positive connections, while red lines are negative connections. The brighter the color of a neuron the stronger its activity with a neighbour neuron. Thickness of the lines also identify the neurons enhanced connectivity. In yellow are the input neurons with their labels corresponding to the 19 EEG channels.

5   Conclusions and Future Directions

The goal of the proposed study has been to analyse how the NeuCube model can be used for classifying and analyse AD EEG data. This is important for the creation of new types of BCI and also for early detection of cognitive decline to be used by clinicians in everyday diagnosis. Further improvement of the understanding and use of the model proposed here are believed to significantly contribute to the advancement in machine learning for the prediction and understanding of brain data and more specifically for data related to neurodegenerative pathologies, such as AD.

There are different scenarios and avenue to be taken in the future, some of those includes:

– Extending the proposed methodology using a higher number of data sets from subjects affected by AD and also by MCI, in order to observe the SNNc state and possibly extract degeneration markers.
– Extending the model adding genetic information in terms of gene regulatory networks [1] as an optimization module to help study the impact of genes on cognitive abilities, e.g. how much gene expression levels of neuroreceptors effect certain cognitive tasks.
– STBD modelling and understanding also through visualization and/or virtual reality of the SNNc, which can be used by clinicians to study how patients effectively improve their neurological activity before and after treatments when compared to healthy control.
– Comparison of the model developed with other techniques used for AD classification, such as Random Forest (e.g. [4]), kernel Support Vector Machine Decision Tree (kSVM-DT) (e.g. [28]), Learning Vector Quantization using Support Vector Machine (LVQ-SVM) (e.g. [16]), Hidden Markov Random Field (e.g. [21]).
– Testing the proposed method in a clinical environment for prediction and early diagnosis of cognitive decline.
– Implementation of the proposed method on neuromorphic hardware to explore its potential for a highly parallel computation [5].

References

1. Benuskova, L., Kasabov, N.: Computational Neurogenetic Modelling. Springer, NY (2007)
2. Delbruck, T.: jaer open source project (2007),
   `http://jaer.wiki.sourceforge.net` (April 14, 2014)
3. Fawcett, T.: An introduction to roc analysis. Pattern Recognition Letters 27(8), 861–874 (2006)
4. Gray, K.R., Aljabar, P., Heckemann, R.A., Hammers, A., Rueckert, D.: Random forest-based similarity measures for multi-modal classification of alzheimer's disease. NeuroImage 65, 167–175 (2013)
5. Indiveri, G., Linares-Barranco, B., Hamilton, T.J., Van Schaik, A., Etienne-Cummings, R., Delbruck, T., Liu, S.C., Dudek, P., Häfliger, P., Renaud, S., et al.: Neuromorphic silicon neuron circuits. Frontiers in Neuroscience 5 (2011)
6. Izhikevich, E.M.: Polychronization: Computation with spikes. Neural Computation 18(2), 245–282 (2006)
7. Kasabov, N.: Evolving connectionist systems: The knowledge engineering approach. Springer (2007)
8. Kasabov, N., Capecci, E.: Spiking neural network methodology for modelling, recognition and understanding of eeg spatio-temporal data measuring cognitive processes during mental tasks. Information Sciences (2014)
9. Kasabov, N., Dhoble, K., Nuntalid, N., Indiveri, G.: Dynamic evolving spiking neural networks for on-line spatio- and spectro-temporal pattern recognition. Neural Networks 41, 188–201 (2013)
10. Kasabov, N.K.: Neucube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. Neural Networks 52, 62–76 (2014)
11. Labate, D., Foresta, F., Morabito, G., Palamara, I., Morabito, F.C.: Entropic measures of eeg complexity in alzheimer's disease through a multivariate multiscale approach. IEEE Sensors Journal 13(9), 3284–3292 (2013)
12. Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T.: Automated talairach atlas labels for functional brain mapping. Human Brain Mapping 10(3), 120–131 (2000)
13. Mohemmed, A., Schliebs, S., Matsuda, S., Kasabov, N.: Span: Spike pattern association neuron for learning spatio-temporal sequences. International Journal of Neural Systems (2012)
14. Morabito, F.C., Labate, D., Bramanti, A., La Foresta, F., Morabito, G., Palamara, I., Szu, H.H.: Enhanced compressibility of eeg signal in alzheimer's disease patients. IEEE Sensors Journal 13(9), 3255–3262 (2013)
15. Morabito, F.C., Labate, D., La Foresta, F., Bramanti, A., Morabito, G., Palamara, I.: Multivariate multi-scale permutation entropy for complexity analysis of alzheimer's disease eeg. Entropy 14(7), 1186–1202 (2012)
16. Ortiz, A., Górriz, J.M., Ramírez, J., Martínez-Murcia, F.J.: Lvq-svm based cad tool applied to structural mri for the diagnosis of the alzheimer's disease. Pattern Recognition Letters 34(14), 1725–1733 (2013)
17. Platel, M.D., Schliebs, S., Kasabov, N.: Quantum-inspired evolutionary algorithm: a multimodel eda. IEEE Transactions on Evolutionary Computation 13(6), 1218–1232 (2009)

18. Pritchard, C., Mayers, A., Baldwin, D.: Changing patterns of neurological mortality in the 10 major developed countries - 1979 - 2010. Public Health 127(4), 357–368 (2013)
19. Rodriguez, G., Copello, F., Vitali, P., Perego, G., Nobili, F.: Eeg spectral profile to stage alzheimer's disease. Clinical Neurophysiology 110, 1831–1837 (1999)
20. Schliebs, S., Defoin-Platel, M., Worner, S., Kasabov, N.: Integrated feature and parameter optimization for an evolving spiking neural network: Exploring heterogeneous probabilistic models. Neural Networks 22(5), 623–632 (2009)
21. Shu, H., Nan, B., Koeppe, R., et al.: Multiple testing for neuroimaging via hidden markov random field. arXiv preprint arXiv:1404.1371 (2014)
22. Song, Q., Kasabov, N.: Ecm - a novel on-line, evolving clustering method and its applications. In: Posner, M.I. (ed.) Foundations of Cognitive Science, pp. 631–682. The MIT Press (2001)
23. Song, S., Miller, K., Abbott, L.: Competitive hebbian learning through spike-timing-dependent synaptic plasticity. Nature Neuroscience 3, 919–926 (2000)
24. Talairach, J., Tournoux, P.: Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: an approach to cerebral imaging. Thieme (1988)
25. Taylor, D., Scott, N., Kasabov, N., Capecci, E., Tu, E., Saywell, N., Chen, Y., Hu, J., Hou, Z.G.: Feasibility of neucube snn architecture for detecting motor execution and motor intention for use in bciapplications. In: 2014 International Joint Conference on Neural Networks (IJCNN), pp. 3221–3225 (July 2014)
26. Thorpe, S., Gautrais, J.: Rank order coding. In: Computational Neuroscience, pp. 113–118. Springer (1998)
27. Tu, E., Kasabov, N., Othman, M., Li, Y., Worner, S., Yang, J., Jia, Z.: Neucube(st) for spatio-temporal data predictive modelling with a case study on ecological data. In: 2014 International Joint Conference on Neural Networks (IJCNN), pp. 638–645 (July 2014)
28. Zhang, Y., Wang, S., Dong, Z.: Classification of alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. Progress in Electromagnetics Research 144, 171–184 (2014)

Part V
Applications

Application of Bayesian Techniques to Behavior Analysis in Maritime Environments

Francesco Castaldo[1], Francesco A.N. Palmieri[1], and Carlo Regazzoni[2]

[1] Seconda Università di Napoli (SUN), DIII, via Roma, 29 - 81031 Aversa (CE), Italy
[2] Università di Genova, DITEN, Via all'Opera Pia, 11 - 16145 Genova (GE), Italy

Abstract. The analysis of vessel behaviors and ship-to-ship interactions in port areas is addressed in this paper by means of the probabilistic tool of Dynamic Bayesian Networks (DBNs). The dimensional reduction of the state space is pursued with Topology Representing Networks (TRNs), yielding the partitioning of the port area in zones of different size and shape. In the training phase, the zone changes of interacting moving vessels trigger different events, the occurrence of which is stored in Event-based DBNs. The interactions are modeled as deviation from the common behavior prescribed by a single-ship normality model, in order to reduce the number of conditional probabilities to calculate and store in the DBNs. Inference on the networks is then carried on to analyze the behavior of various ships and vessels maneuvering in the harbor. The results of the algorithm are showed by using simulated data relative to a real port.

Keywords: Interaction Analysis, Ship-to-Ship Interactions, Dynamic Bayesian Networks, Topology Representing Network.

1 Introduction

The sadly famous Costa Concordia accident [1], as other dramatic crashes happened in recent years in port areas or near the coastlines [2], confirm that the design of monitoring systems able to supervise complex and crowded areas as harbors, coastlines, airports, etc., is very far from being considered a closed issue. Nowadays, these areas are monitored by a great number of high-quality sensors, but the lack of robust methodologies able to combine these volumes of data hinders to analyze and comprehend what is really happening in the area under surveillance.

In this paper, we analyze vessels of different kind during the time they reside in generic port areas. The security of maritime environments may be jeopardized by a great number of different threats: ships moving too rapidly or too slowly, pairs of ships sailing too close to each other, small vessels obstructing the passage for larger ships, and so forth. By understanding and labeling the movements in the area we could build an intelligent system, capable of providing alerts or warnings to the human operators (whose presence is obligatory in ports) when the detected situations are not acknowledged as *normal*. The issue is that in

crowded harbors it is likely to find many types of ships (motorboats, tugboats, container ships, etc.) interacting in many different ways with the other moving objects in the scene. In general terms, an interaction in maritime environments [3] occurs when a ship comes too close to another ship, or too close to a river, or to a canal bank. We focus on *ship-to-ship interactions* [3] [4] in ports, i.e. when the presence and the movements of one ship affect the behavior of another, and vice-versa. The ships type, the navigation rules [5] of the country in which the port resides, the wind conditions, are all parameters that concur to define a normal interaction between vessels. We assume the interaction to be over when the two ships reach their destinations (for entering ships) or leave the port area (for exiting ships).

In literature the ship-to-ship interaction problem has been mostly approached by analyzing the hydrodynamic phenomena arising when two or more watercraft are slightly spaced from each another [6] [4]. Bayesian reasoning [7] [8] [9] has been extensively used to study the interaction of objects for different applications and in different scenarios, but little has been done for the behavioral analysis of pairs of ships. The reduction of the state space, necessary to carry out the event-based approach described in the paper, is achieved by means of Topology Representing Networks (TRNs), among which we choose the Instantaneous Topological Map [10].

The paper is structured as follows. In Section 2 we analyze the techniques to reduce the state space and partition the area in zones. Section 3 describes the probabilistic approach based on the event detection and identification. In Section 4 we drawn some results by using data generated in a simulated environment that replicates the port of Salerno, Italy. Section 5 is for conclusions and future developments.

2   Reduction of State Space with Topology Representing Networks

In this paper the actors in play, namely ships and vessels of different kinds and shape, are treated as points moving in the 2D space described by the portion of sea included in the port. For the i-th ship we can define the state vector $s_t^i = (x_t^i, y_t^i)^T$, representing the position of the moving object at time t. The analysis of behaviors and interactions between ships (and in general between moving objects) by evaluating the low-level state space trajectories as they are (without any modification) turns out to be quite a challenge, given the great variability of the state space vectors relative to multiple ships in port areas (even in small ones). However, we can exploit the fact that the "features" of a port (i.e. the position and the shape of the docks, the common routes of the ships, etc.) can be easily known a priori and do not change very often. If we are able to construct a *topological map* of the harbor, it is possible to design a higher-level algorithm where behaviors and interactions are emphasized and emerge with more clarity.

The simplest way to build a map is to partition the area in zones of equal size with a rectangular grid. However, this approach ignores that ships and

vessels take only certain routes to enter or exit the port. The latter information is precious because we expect an "intelligent" map to be more precise in the zones where many ships pass through, and coarser in places rarely touched by ships. The Topology Representing Networks (TRNs) are an important class of algorithms that exploits at best the positional information of the actors in play, by building a map from a dataset of moving objects exploring the scene. The most famous TRN algorithms are the Self Organizing Maps (SOM) [11], but in recent times other approaches have been proposed, as the Growing Neural Gas (GNGs) [12] or the Instantaneous Topological Maps (ITMs) [10]. In this paper the topological maps are built by means of ITMs, and this is motivated by the fact that ITMs are quite good in handling strongly-correlated data, as the one provided by ships and vessels sailing in the port. We do not report the explanation of the algorithm, as it is a straightforward implementation of the procedure detailed in [10]. In order to build the map, we need to set only two parameters, namely the resolution e_{max} and the smoothing parameter ϵ_{itm}.

Therefore we assume to have the map of the environment, i.e. to have a set of N_n nodes, each of which corresponds to a zone. A zone can be defined as the portion of the space whose points are closer (respect to a fixed distance definition, as for example the Euclidean distance) to the generator node (the "center" of the zone). The Bayesian models defined in Section 3 are based on zones changes triggered by the moving vessels in the area. More in detail, when the i-th vessel moves from zone a to zone b, an event $\epsilon_t^{i(a,b)} = l_a \to l_b$ is triggered, where l_a and l_b are the labels identifying two neighboring zones and $t \in \mathbb{N}$ is the time at which the event occurs. The events $\epsilon_t^{i(a,b)}$ can be seen as the outcomes of the discrete random variable $\mathcal{E}_t^i$, that will be the state of the Bayesian networks. If the vessel remains in the same zone a for a T_{max} time, a *still* event $\epsilon_t^{i(a,a)} = l_a \to l_a$ is detected.

3   Bayesian Models

By means of an Event-based Dynamic Bayesian Network (E-DBN) [9] [7], we define a *normality* model Θ^1, relative to a target i-th ship sailing in the port. In the E-DBN we encode the probability of the event ϵ_t^i, given the previous event $\epsilon_{t-\Delta_t^i}^i$ through the following conditional probability (CPD),

$$\Theta^1 = p(\epsilon_t^i | \epsilon_{t-\Delta_t^i}^i). \tag{1}$$

If the target ship is alone in the port (or very far from other ships) and does not behave accordingly to the normality model (zig-zag trajectories, vessels stopping in the middle of the port, etc.), we can infer that the behavior is abnormal and a warning can be send to the operator.

Things are different when other vessels are nearby the target ship, because a deviation from the normality could be due to interactions between the vessels (as for instance a tugboat towing a cargo ship, a motorboat overtaking a sailboat, etc.). For this reason it is useful to define the following parameters: (a) the

Euclidean distances d_j^i (called *influence distances*) between the target i-th ship and the other j-th ships which reside or maneuver in the port area; (b) an *influence threshold* τ_i, that, compared with the d_j^i distances, permits to verify if the target i-th ship and the other j-th ships are close enough to interact ($d_j^i < \tau_i$) or not. Actually, in the experimental part of this paper (Section 4) we will use pairs of ships that, for simplicity, interact with each other during all the time they maneuver in the harbor, but in general the influence distances are very important in a multi-target scenario, where we do not know a priori who interacts with whom.

Given these distances, it is possible to see the interactions between ships as deviation from the normality described by the model defined in Equation (1). More specifically, if the j-th ship is very close to the target ship (i.e. $d_j^i < \tau_i$), the following interaction model can be defined

$$\Theta^m = p(\epsilon_t^i|\epsilon_{t-\Delta_t^i}^i, \epsilon_{t-\Delta_t^j}^j), \tag{2}$$

where $m = 2, .., M$ denotes the type of interaction and $\epsilon_{t-\Delta_t^j}^j$ is the event relative to the j-th ship, with $t < \Delta_t^j \leq \Delta_t^i$. Equation (2) can be written as

$$p(\epsilon_t^i|\epsilon_{t-\Delta_t^i}^i, \epsilon_{t-\Delta_t^j}^j) = \frac{p(\epsilon_t^i, \epsilon_{t-\Delta_t^i}^i, \epsilon_{t-\Delta_t^j}^j)}{p(\epsilon_{t-\Delta_t^i}^i, \epsilon_{t-\Delta_t^j}^j)} = \frac{p(\epsilon_{t-\Delta_t^j}^j|\epsilon_t^i, \epsilon_{t-\Delta_t^i}^i)p(\epsilon_t^i|\epsilon_{t-\Delta_t^i}^i)}{p(\epsilon_{t-\Delta_t^j}^j|\epsilon_{t-\Delta_t^i}^i)}, \tag{3}$$

where $p(\epsilon_t^i|\epsilon_{t-\Delta_t^i}^i)$ is the conditional probability defining the normality model of Equation (1). In other words, we can define the interaction as deviation from the normal model Θ^1, by adding two CPDs, namely $p(\epsilon_{t-\Delta_t^j}^j|\epsilon_t^i, \epsilon_{t-\Delta_t^i}^i)$ and $p(\epsilon_{t-\Delta_t^j}^j|\epsilon_{t-\Delta_t^i}^i)$, and in this way we can reduce the number of CPDs to store and use. In this paper we focus on a very common type of ship-to-ship interaction between two vessels, but the proposed approach can be extended to the (unlikely) case of three and more interacting ships by adding the correspondent events in the model defined in (2). For instance, in the case of three ships we may define the CPD $p(\epsilon_t^i|\epsilon_{t-\Delta_t^i}^i, \epsilon_{t-\Delta_t^j}^j, \epsilon_{t-\Delta_t^n}^n)$, where n denotes the third ship and $t < \Delta_t^n \leq \Delta_t^i$.

The Bayesian networks just introduced can be used to infer the behavior of ships maneuvering in the port, but only after an initial training process, in which the conditional probabilities within the models are estimated and stored. The latter CPDs describe the probability of a cause-effect relation between the events of nearby vessels, and are calculated with a maximum likelihood training algorithm [7], equivalent to counting the number of occurrences of the outcomes of the CPDs in the dataset, normalized to the total number of occurrences.

The training of the network is performed with different datasets, related to behavior and interaction models. We point out that in large port areas different normality models (e.g. relative to different docks of the port) could exist, and the same for the interaction models. In such cases the number of models could

significantly grow, and this is the main reason we decided to model the interactions as deviation from the normal path prescribed by the normality model.

If we assume to have M models Θ^m, $m = 1, .., M$, in order to calculate the probability that a new target ship behaves accordingly with the Θ^m model, the following cumulative normalized measure is proposed

$$\alpha_k^m = \left(\frac{k-1}{k}\right) \alpha_{k-1}^m + \frac{1}{k}\Theta^m, \qquad (4)$$

where $k \in \mathbb{N}$ denotes the number of detected events for the target ship and with $0 < \alpha_k^m < 1$. The normality model Θ^1 can be always evaluated with the data of the target ship, while the interaction models $\Theta^2, .., \Theta^M$ are used only if at least another vessel is nearby the target ship (information provided by the evaluation of the influence distances d_j^i). Given the vessels trajectories, for each couple of events we calculate α_k^m and compare each model Θ^m with a threshold τ_n. If none of the models is compatible with the trajectory (i.e. the α^m values result above the threshold for each m-model), we can infer that the ship behavior is abnormal. We point out that α_k^m is a function that takes into account the past history along with the probability of the current events, and its trend can be analyzed in real time to infer the behavior of the ship during the time period it resides in the harbor.

4   Preliminary Results

In the following we report results of behavior analysis of ships in the Port of Salerno, Italy. The data are provided by a realistic simulator of trajectories, which reproduces the real structure of the port and generates the movements of ships entering the port (for simplicity we assume only entering ships, but the same reasoning can be applied when we have at the same time exiting ships). Figure 1 left depicts an image of the harbor of Salerno, and indicates in black the dock on which we focus our behavioral analysis. Figure 1 right depicts a frame of the simulator.

As explained in Section 3, the first step is to build the normality model Θ^1. This is accomplished with $N_{itm} = 150$ noisy trajectories of vessels heading to the dock. These trajectories are used at first to build the Instantaneous Topological Map defined in Section 2, with $e_{max} = 5$ and $\epsilon_{itm} = 0.1$, and then to store the CPDs of Equation (1) for different consecutive events. In Figure 1 left the ITM is superimposed to the port image.

Given the normality model, it is possible to construct ship-to-ship interaction models Θ^m, $m = 2, .., M$, that are allowed in the portion of the port under surveillance. For simplicity, we build a single interaction model, and show how interactions not compatible with that model are robustly recognized. Given the European maritime rules in harbors [13] [14], we define an interaction model Θ^2 relative to a sailboat and a motorboat trying to enter the port area at the same time. The navigation rules prescribe that the ship with the highest level of maneuverability (in this case, the motorboat) stops its engine, lets the other

Fig. 1. Left: Satellite photo of the port of Salerno. We focus our attention to the right dock (indicated with black lines) where small vessels as motorboats and sailboats are allowed to land or depart (the other two major docks on the left are only for container or cargo ships). On the same image the ITM on which the events are gathered is superimposed. The green lines connect the neighbor nodes, and for each node we define the correspondent zone as the locus of points that are closer to that node, with respect to the others. Right: A picture taken from the simulator used in this paper, that accurately reproduces the shape of the port and generates realistic trajectories of ships in the area.

ship pass through the entrance and only after enters the port. We call this a *motorboat-sailboat* interaction, because the target ship is always the motorboat and the other ship is always the sailboat, and we generate the model by using $N_{ms} = 300$ noisy trajectories extracted from the simulator. We remark again that other interactions (for instance two motorboats entering the port in the same moment, a tugboat towing a container ship, etc.) are possible and can be easily built in different Θ^m models. We have chosen empirically the value of $\tau_i = 0.4$ and $T_{max} = 3$.

Once the ITM is created and models are assembled, inference on the data can be carried on. More in detail, a high-quality system has to guarantee two features: (a) low false alarm rates; (b) robust recognition of uncommon and abnormal behaviors or interactions. In order to assess the first feature, in the first experiment we test the Bayesian models with $N_{t_1} = 200$ noisy trajectories of two nearby ships that act as motorboat and sailboat of the interaction model Θ^2. The single trajectories of these ships are compared with the normality model Θ^1, while at the same time the data from the two vessels are combined and compared with the Θ^2 model. In Figure 2 we depict the trend over the events of the cumulative measure defined in Equation (4). The analysis of the figure permits to draw the following conclusions: (a) the two trajectories singularly are almost always recognized as belonging to the normality model (their trends in very few cases and for little time are below the recognition threshold τ_i). This is true because in the model there is no indication of the time spent during
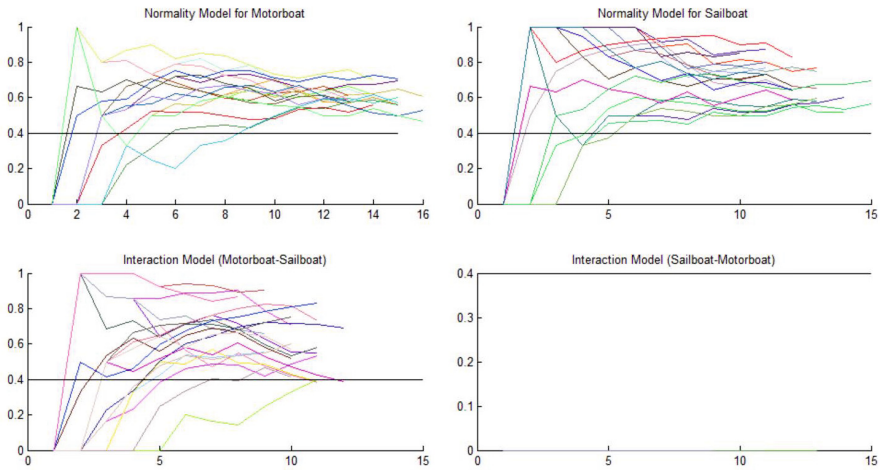
Fig. 2. This figure depicts the cumulative trends over the trajectories of interacting motorboats and sailboats. The two plots on top are relative to the single trajectories compared with the normality model Θ^1, while the bottom plots are for the interaction model Θ^2.

the transition between events, therefore the fact that the motorboat stops at the entrance is not captured by the normality model; (b) the interaction model recognizes in most cases the motorboat-sailboat coupled behavior. Of course this is true when the first ship is the motorboat and the other is the sailboat (bottom left of Figure 2), and not when the ship roles are switched (bottom right of Figure 2).

In the second experiment we assess the ability of the system to alert the operator of strange or dangerous behaviors. We generate $N_{t_2} = 100$ trajectories relative to an interaction named *tugboat-cargo*, representative of the situations in which a large cargo ship is towed in the port by a tugboat. This type of interaction is not allowed in the dock we are monitoring, therefore it is a dangerous situation that should be recognized. Even if the two ships are not a cargo and a tugboat but two motorboats or sailboats traveling together, this can be considered still a noteworthy situation because two different ships so close in the port area could collide and cause relevant damages to the harbor structures. In Figure 3 are depicted the results, that are quite good and can be interpreted as follows: the two trajectory, taken singularly, are compatible with the Θ^1 model, but their interaction is not recognized by the Θ^2 model, except for very few cases and only for a few number of events. Such situation (two ships behaving in a normal way singularly but not interacting in a known way) can be easily reported to the operator, that can decide to intervene or not.
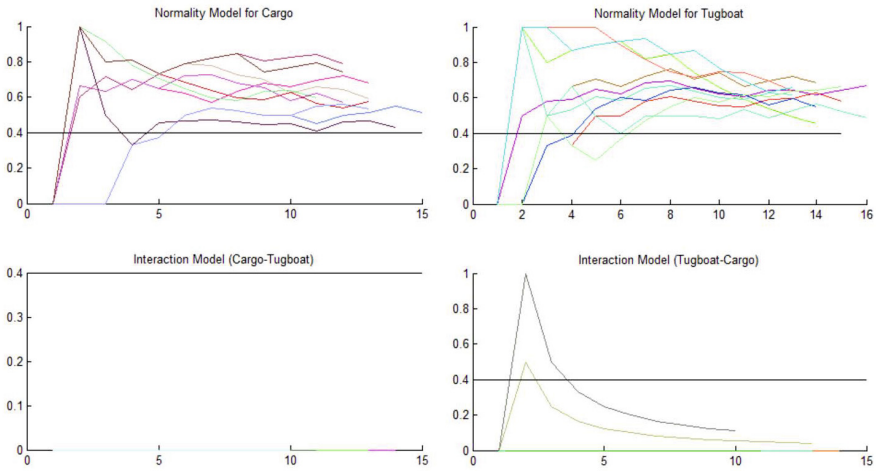
Fig. 3. In this figure the models are compared with the data of ships interacting according to the tugboat-cargo model, in which one ship (the tug) tows the other (the cargo) into the port. While singularly the two ships are behaving correctly, this interaction is not allowed in the small dock we are observing, and the trends of the various cumulative measures permit to automatically evaluate such situation and to report it to the operator.

5 Conclusion

This paper has presented an application of Bayesian networks for behavioral analysis of multiple ships in port areas. The idea is to preserve the port safety by classifying the movements of the different actors in the scene. The analysis is complicated by the fact that multiple ships can interact in many ways, with a number of interaction models that could become very large: the idea pursued in this paper is to relate the interactions to normality models, i.e. by modeling the interaction as deviation from the normal path taken by a ship maneuvering without other vessels in the port area. In this way we construct interactions starting from the normality model, reducing in this way the probabilistic data we have to gather and use for inference. The computational load of the algorithm is quite low, because after the training step the inference is carried on by simply updating the cumulative measure for the normality and interaction models.

Other information can be gathered from moving ships and used to enhance the probabilistic model. For instance, the travel time of the ships into the zones can be saved along with the zone changes, and this information could be precious to recognize abnormal behaviors strictly connected with the vessel speed (i.e. ships that are too fast or slow, that stop into the middle of the port, etc.). Another useful information could be the initial position of the ship entering in a zone (i.e. from which part of the zone the ships usually enter), that could be used to construct, within the zone, a low-level tracking model by which follow the ship.

The latter information could be used to anticipate the behavioral analysis at the level of the tracker instead of waiting for consecutive events (that for large zones could be triggered after quite long times).

References

1. Costa concordia: What happened,
   `http://www.bbc.com/news/world-europe-16563562`
2. Cargo ship crashes into port control tower in genoa killing three,
   `http://www.theguardian.com/world/2013/`
   `may/08/italian-cargo-ship-crashes-genoa`
3. Barrass, B.: Ship Design and Performance for Masters and Mates. Elsevier Science (2004), `http://books.google.it/books?id=2KaLDCpZgbQC`
4. Eloot, K., Vantorre, M.: Ship behaviour in shallow and confined water: an overview of hydrodynamic effects through efd. In: Assessment of Stability and Control Prediction Methods for NATO Air and Sea Vehicles. NATO. Research and Technology Organisation (RTO), p. 20 (2011)
5. U. N. E. C. for Europe. Working Party on Inland Water Transport. CEVNI:, ser. TRANS/SC. 3/115/Rev. 2. UN (2002),
   `http://books.google.it/books?id=RVVsQticMUgC`
6. The Second International Conference on Manoeuvring in Shallow & Confined Waters: Ship to Ship Interaction (May 2011)
7. Barber, D.: Bayesian Reasoning and Machine Learning. Cambridge University Press (2012)
8. Koller, D., Friedman, N.: Probabilistic Graphical Models: Principles and Techniques. MIT Press (2009)
9. Murphy, K.: Dynamic bayesian networks: Representation, inference and learning. Ph.D. dissertation, UC Berkeley, Computer Science Division (July 2002)
10. Jockusch, J., Ritter, H.: An instantaneous topological mapping model for correlated stimuli. In: International Joint Conference on Neural Networks, IJCNN 1999, vol. 1, pp. 529–534 (1999)
11. Kohonen, T.: Self-organized Formation of Topologically Correct Feature Maps. In: Anderson, J.A., Rosenfeld, E. (eds.) Neurocomputing: Foundations of research, pp. 509–521. MIT Press, Cambridge (1988),
    `http://dl.acm.org/citation.cfm?id=65669.104428`
12. Fritzke, B.: A growing neural gas network learns topologies. In: Advances in Neural Information Processing Systems 7, pp. 625–632. MIT Press (1995)
13. I. M. Organization, International Convention for the Safety of Life at Sea: consolidated text of the 1974 SOLAS Convention, the 1978 SOLAS Protocol, the 1981 and 1983 SOLAS Amendments, ser. IMO Publication. IMO (1986),
    `http://books.google.it/books?id=_5oTAAAAYAAJ`
14. I. M. Organization, ISPS Code: International Ship and Port Facility Security Code and SOLAS Amendments 2002 Adopted 12 December 2002, ser. IMO publication. International Maritime Organization (2003),
    `http://books.google.it/books?id=MdUQAQAAIAAJ`

Domestic Water and Natural Gas Demand Forecasting by Using Heterogeneous Data: A Preliminary Study

Marco Fagiani, Stefano Squartini, Leonardo Gabrielli,
Susanna Spinsante, and Francesco Piazza

Department of Information Engineering
Università Politecnica delle Marche, Ancona, Italy
{m.fagiani,s.squartini,l.gabrielli,s.spinsante,f.piazza}@univpm.it

Abstract. In this paper a preliminary study concerning prediction of domestic consumptions of water and natural gas based on genetic programming (GP) and its combination with extended Kalman filter (EKF) is presented. The used database (AMPds) are composed of power, water, natural gas consumptions and temperatures. The study aims to investigate novel solutions and adopts state-of-the-art approaches to forecast resource demands using heterogeneous data of an household scenario. In order to have a better insight of the prediction performance and properly evaluate possible correlation between the various data types, the GP approach has been applied varying the combination of input data, the time resolution, the number of previous data used for the prediction (lags) and the maximum depth of the tree. The best performance for both water and natural gas prediction have been achieved using the results obtained by the GP model created for a time resolution of $24\,h$, and using a set of input data composed of both water and natural gas consumptions. The results confirm the presence of a strong correlation between natural gas and water consumptions. Additional experiments have been executed in order to evaluate the effect of the prediction performance using long period heterogeneous data, obtained from the U.S. Energy Information Administration (E.I.A.).

Keywords: domestic consumption forecasting, heterogeneous data, computational intelligence, genetic programming.

1 Introduction

Nowadays, unlike the electrical energy scenario, where several databases and computational intelligence approaches exist, water and natural gas fields, as highlighted in Fagiani *et al.* [5], are affected by a severe lack of databases and studies. The only publicly available databases have been presented in Makonin *et al.* [9] and in Nasseri *et al.* [10]. Additional data are available on the U.S. Energy Information Administration (E.I.A.) site[1]. For both the Almanac of

---

[1] http://www.eia.gov/

Minutely Power dataset (AMPds) [9] and the U.S. Natural Gas Consumption
the temperature data are also available on Climate[2] and National Climatic Data
Center[3], respectively.

Concerning the load forecasting techniques, many approaches have been evalu-
ated for the water scenario and only a few for the natural gas case. A comparative
study between grey forecast model and RBF neural network model for annual
water demand prediction has been presented in Liu and Chang [8]. In order
to forecast the urban water consumption, the combination of quantum particle
swarm optimization (QPSO) algorithm with RBF neural network has been in-
troduced by Zhu and Xu [17]. Nasseri *et al.* [10] have introduced the application
of genetic programming to forecast the monthly water demand of Tehran, and
have discussed its application in combination with the extended Kalman filter.

Specifically, in the state of the art only Azari *et al.* [1], Tabesh and Dini [16],
and Pang [11] have recently discussed the effects of the application of multiple
data type, like weather information or temperature, on the forecasting perfor-
mance. Unfortunately, due to the presence of non-standard evaluation criteria
and since each new method has been tested with a different database [5], it is
extremely difficult to perform a comparison between different approaches. In ad-
dition, none of them have approached the study of correlation effects between
the usage of different resources: water, natural gas and power.

Moving from such a state of the art analysis, the main issues characterizing
the methodological approach followed in this preliminary study are concern the
identification of the most relevant and publicly available databases for the water
and gas case studies and their use for experimentation (data heterogeneity is
one of the key feature), the development of suitable Computational Intelligence
algorithms for load forecasting in different operative contexts and, finally, a
performance evaluation in according to the most used criteria in the field and
comparison with similar techniques.

The authors are confident that the spread of innovative monitoring systems,
which are more and more often based on low-power wireless devices [13–15], will
ensure a facilitation for collecting and making publicly available large amount
of data containing multi-utilities information. Therefore, the AMPds [9] rep-
resents the most complete database presently available that fulfils the author
purposes. It is composed of power, water and natural gas meter data of a single
house, recorded at one minute intervals for an entire year. Additional new data
will be added in April of each year, starting 2013. Although the dataset refers
to measurement of a single house for just one year, the diversity of resource
types and the large amount of data allows to lead a study about correlation
aspects between various resources, as well as correlation between resources and
meteorological conditions.

Concerning the water and natural gas consumption forecast, in this paper pre-
dictions based on genetic programming (GP) and extended Kalman filter (EKF)
are presented. The study aims to investigate novel solutions to forecast domestic

---

[2] `http://climate.weather.gc.ca/climateData/dailydata_e.html?StationID=889`
[3] `http://www.ncdc.noaa.gov/cag/`

demands using heterogeneous data composed of different resource types, in order to research mutual correlation effects. In the future, such information can be use to improve both prediction and leakage detection techniques. Up to the authors' knowledge, there are no others studies that have used the publicly available AM-Pds dataset to execute experiments concerning the short term load forecasting of water and natural gas consumption in a domestic scenario.

This is the paper outline. In Section 2 the GP and EKF-GP algorithms are briefly described. The experimental tests are described and related results are commented in Section 3, whereas in Section 3.1 results comparison and further tests on monthly forecasting are described. Section 4 concludes the paper.

2   Computational Intelligence Algorithms

Genetic Programming

The genetic programming is inspired by the evolution process and, as this, it is based on three main criteria [2]: heredity, variability and fecundity. In this way, an individual of a population will be able to adapt to an environment.

The algorithm consists of the following steps:

1. **Initialization:** the parameters are set and the first generation is randomly created.
2. **Selection:** the best individual is selected using the sum of the absolute errors (SAE) and it is opportunely used for the creation of the new generation.
3. **Control:** the second step is repeated for the new generation until either a stop condition or the maximum generation is reached.

At the end, the algorithm returns the best individual found since the beginning of the simulation.

In order to have a better insight of the prediction performance and properly evaluate possible correlation between different input data, the authors have applied the algorithm to each possible combination of input data and forecast resolution. In each test the best GP model has been sought varying both the number of previous data used for the prediction (lags) and the maximum depth of the node.

Extended Kalman Filter

The extended Kalman filter model produces a suboptimal solution using a first order linear approximation of the filter proposed by Kalman [7]. The algorithm consists of successive uses of *predict* and *update* equations, for each input/output relation.

The main steps of the algorithm are:

– **Initialization**: the first estimate is inferred from the only available information of the first input/output relation.
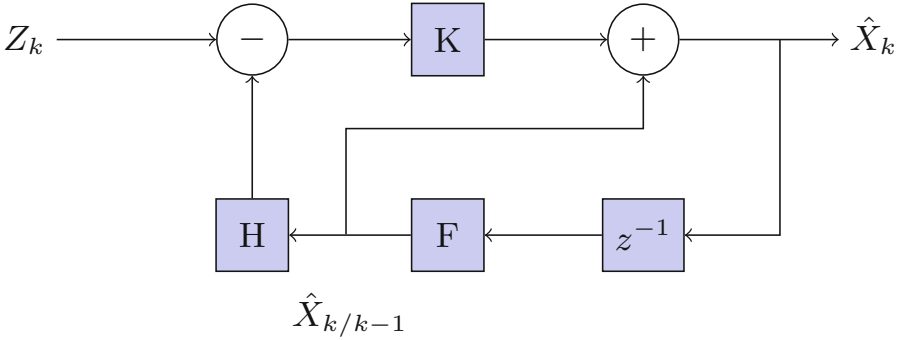
Fig. 1. Extended Kalman filter block diagram

- **Iteration**: the information of the previous estimate are used to correct the current estimate, minimizing the prediction mean-square error.

In Fig. 1 the block diagram of the extended Kalman filter is shown. Let F be the transition matrix and A the Jacobian matrix of partial derivatives of F, the *predict* equations at time k are defined as:

$$\hat{X}_{k/k-1} = F_{k/k-1}\,\hat{X}_{k-1/k-1}\,, \tag{1}$$

$$P_{k/k-1} = A_{k-1}\,P_{k-1/k-1}\,A_{k-1}^{T} + Q_{k-1}\,. \tag{2}$$

The *update* equations, assuming that H is the observation model and Z_k is the measurement vector, are given as:

$$K_k = P_{k/k+1}\,H_k^{T}\,(H_k P_{k/k-1} H_k^{T} + R_k)^{-1}\,, \tag{3}$$

$$\hat{X}_{k/k} = \hat{X}_{k/k-1} + K_k(Z_k - H_k\,\hat{X}_{k/k-1})\,, \tag{4}$$

$$P_{k/k} = P_{k/k-1} - K_k\,H_k\,P_{k/k-1}\,. \tag{5}$$

EKF-GP

The combination of extended Kalman filter and genetic programming, as indicated by Nasseri *et al.* [10], consists of the following steps:

1. The GP algorithm is performed and the best model is selected.
2. The EKF is applied at the dataset with the following settings:
   - the selected GP model is used as transition matrix F;
   - the noise covariances Q_k and R_k are set in accordance with the *NMSE* of the GP model;

      – the observation model H and the state variable X are vectors of dimension $m \times 1$, with m equal to the lags number.

3. The EKF predicted states are re-computed using the GP model.

Therefore, the k-th EKF predicted states, whose number depends on the lags assumed, are used as input variables for the selected GP model. The k-th predicted value is given as output of the model equation.

    More precisely, keeping the same selected GP function, new values of input lags are estimated, through the EKF, in order to improve the predictions. The original lags, that are used to compute the k-th prediction, are assigned as values of the predicted state estimate vector, $\hat{X}_{k/k-1}$, and, setting it as input, the result achieved by the GP model is used as measurement vector, Z_k, [10]. Furthermore, assuming that F is the GP function, in compliance with the definition of transition matrix, entails that the matrix A is diagonal and it is composed of partial derivatives of the GP function, F, with respect to the m input variables (lags) $x_1, x_2, \ldots, x_m$, as depicted below.

$$A_{m,m} = \begin{pmatrix} \frac{\partial F}{\partial x_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial F}{\partial x_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial F}{\partial x_m} \end{pmatrix} \tag{6}$$

    In addition, also the covariance estimate, P, is diagonal with dimension $m \times m$, and its initial values are all set equal to the noise covariance, R_k.

    The selected operators for the GP are: plus ($+$), minus ($-$), product ($\times$), division ($\div$), power (x^n), sine (sin) and cosine (cos). The MATLAB toolboxes EKF/UKF and GPLAB [6,12] have been used to execute the simulations.

3   Experimental Results

The experiments have been conducted using the AMPds database, presented by Makonin *te al.* [9] and described in Section 1. Six combination of input data, for both natural gas and water prediction, composed by merging the available input information as reported in Table 1, have been evaluated. The 70% of each set has been randomly selected and used for the training process; the remaining data (30%) have been used for testing the model and generate the reported results.

    The forecasts for $1\,h$, $6\,h$, $12\,h$ and $24\,h$ have been evaluated. For each set, different models have been trained and tested for 5, 3 and 2 lags and for a maximum tree depth of 20, 15 and 10. The population size, the maximum number of generation, and the cross-over and mutation probability have been set, and left untouched for all the simulation, to 100, 1000 and 0.1, respectively.

The results have been mainly evaluated in terms of normalized mean square error (*NMSE*) and coefficient of determination (R^2) [4], defined as:

$$\text{NMSE} = \frac{\sum_{}^{N}(\tilde{y}_i - y_i)^2}{\sigma_y \cdot N} , \tag{7}$$

$$R^2 = 1 - \frac{\frac{1}{N}\sum_{}^{N}(y_i - \tilde{y}_i)^2}{\frac{1}{N}\sum_{}^{N}(y_i - \bar{y})^2} . \tag{8}$$

where y_i indicates the i-th observed value, $\tilde{y}_i$ the corresponding i-th forecast value, the average $\bar{y}$ and the variance σ_y of the N observed values. Additional evaluation criteria taken into account are the mean square error (*MSE*), the mean absolute percentage error (*MAPE*) and the relative root mean square error (*RRMSE*).

Table 1. Summary of the best performance obtained for each time resolution. The "Input data" column indicates the different combinations of available input resources.

Input data	1 h		6 h		12 h		24 h	
	$NMSE$	R^2	$NMSE$	R^2	$NMSE$	R^2	$NMSE$	R^2
Natural Gas Prediction								
G	0.85	0.15	0.37	0.62	0.31	0.68	0.25	0.75
GT	0.46	0.53	0.39	0.60	0.30	0.69	0.28	0.71
WG	0.84	0.16	0.39	0.60	0.33	0.67	**0.24**	**0.76**
WGT	0.79	0.20	0.39	0.61	0.29	0.71	0.28	0.71
WGE	0.88	0.11	0.40	0.60	0.31	0.69	0.27	0.72
WGET	0.86	0.14	0.39	0.60	0.33	0.67	0.28	0.72
Water Prediction								
W	0.94	0.06	0.42	0.57	0.45	0.55	0.43	0.56
WT	0.95	0.05	0.44	0.56	0.44	0.56	0.43	0.57
WG	0.92	0.08	0.45	0.55	0.48	0.51	**0.41**	**0.58**
WGT	0.92	0.08	0.46	0.54	0.53	0.47	0.45	0.54
WGE	0.90	0.10	0.45	0.55	0.54	0.46	0.44	0.55
WGET	0.91	0.09	0.46	0.53	0.50	0.50	0.43	0.57
W = water	G = natural gas		E = electric power		T = temperature			

The best results obtained for each time resolution and input set are reported in Table 1. For the natural gas, For the natural gas, decreasing the time resolution of the prediction produces a clear performance improvement. The best performance has been obtained for a 24 h time interval using as input 5 lags of both natural gas and water consumption, and a tree depth of 15. For this condition the obtained *MSE* is 0.099, the *RRMSE* is 39.24%, and the *MAPE* is 25.23%. On the contrary, the water prediction performance show a good improvement decreasing the time

Table 2. Comparison of the best results obtained with the GP and EKF-GP for each time resolution.

	$1\,h$		$6\,h$		$12\,h$		$24\,h$	
	$NMSE$	R^2	$NMSE$	R^2	$NMSE$	R^2	$NMSE$	R^2
	Gas Prediction							
	WGT		G		WGT		WG	
GP	0.79	0.20	0.37	0.62	0.29	0.71	0.24	0.76
EKF-GP	0.79	0.20	−	−	0.32	0.68	**0.23**	**0.77**
	Water Prediction							
	WGE		W		WT		WG	
GP	0.90	0.10	0.42	0.57	0.44	0.56	**0.41**	**0.58**
EKF-GP	0.90	0.10	0.44	0.56	0.44	0.56	0.58	0.42

−: no results, derivatives differs too much.

Table 3. Results comparison with the state-of-the-art studies.

	Tech.	MSE	NMSE	R^2
Gas Prediction (WG)	GP	0.007	0.412	0.584
Bakker *et al.* [3]	Adaptive	-	-	$0.658 - 0.802$
Tabesh and Dini [16]	Fuzzy	0.042	0.465	0.760
Tabesh and Dini [16]	Neuro-fuzzy	0.007	0.064	0.936

resolution from 1 hour to 6 hours, and a smooth improvement from 6 to 24 hours. However, the best performance is reached for a $24\,h$ time interval using as input 2 lags of both water and natural gas consumption, and a tree depth of 20. For this condition the obtained *MSE* is 0.0070, the *RRMSE* is 27.37%, and the *MAPE* is 21.15%. In both natural gas and water prediction the correlation between the resources is evident.

The results obtained applying the EKF-GP approach at the best performing GP model of each resolution time are reported in Table 2. The prediction performance remains almost the same introducing the EKF, except for the $24\,h$ water prediction, for which its performance decrease drastically.

3.1   Further Remarks

In Table 3 a comparison with results achieved in the state-of-the-art experiments, concerning the $24\,h$ resolution water forecasting, is reported. The authors are aware that the results achieved by GP approach are lower than those of the comparable studies. However, it should be noted that, differently from the AMPds, which data have been collected in 1 year, the database used in Tabesh and Dini [16] is composed of about $4,300$ samples, over 12 years of recording, with both temperature and humidity information. Similarly, Bakker *et al.* [3] performed the forecasting with a database of $210,336$ samples collected over 6 years.

Table 4. Comparison of the best results obtained for the GP approach and the results achieved applying the EKF-GP, for each data combination.

	GP		EKF-GP	
	NMSE	R^2	*NMSE*	R^2
Maine	0.166	0.822	0.231	0.768
Maine + Temp.	0.103	0.896	**0.032**	**0.968**
Illinois	0.069	0.930	0.104	0.896
Illinois + Temp.	0.068	0.931	**0.061**	**0.938**
Louisiana	0.199	0.798	0.156	0.844
Louisiana + Temp.	0.153	0.845	**0.153**	**0.845**

Therefore, in order to have a better insights on the prediction capability of the GP and EKF-GP using heterogeneous data collected over long period, the approaches have been applied to forecast the monthly natural gas demands of three U.S. states, using the information available on U.S. Energy Information Administration (E.I.A.) site from 1989 to 2013, as described in Section 1.

The monthly data of 3 states have been chosen: Maine, Illinois and Louisiana, representing a cold, a mild and a hot climate, respectively. For each state two different data sets have been evaluated: one composed of both gas consumptions and temperature data, and the other composed of gas consumption data only. As for the previous approach, 70% of each set has been randomly selected and used for training the corresponding GP model, the remaining data (30%) has been used for the testing and generating the reported results. The results of both GP and EKF-GP combination are shown Table 4.

As seen for the domestic prediction, the introduction of heterogeneous data results in a performance improvement, remarkable in the case of Maine data. The application of the EKF-GP approach generates an additional enhancement for the combined data. Moreover, although the obtained results cannot be compared with state-of-the-art studies, the performance reached in Table 4 seem to be good and representing a valuable starting point for future developments.

4   Conclusion

In this paper the authors present a preliminary study concerning the use of heterogeneous data to forecast the domestic consumptions of natural gas and water. Unlike heterogeneous data applications that have been shown in the state-of-the art studies [1,11,16], which experiments have been executed using combinations of a single resource type with weather conditions or temperatures, the data heterogeneity expressed in this paper is extended to the combination of multiple resource types. Using the AMPds [9] database the authors have performed, and up to their knowledge no other study has done it, forecasting experiments of water and natural gas consumption in a household scenario.

The prediction experiments have been computed using the genetic programming paradigm, and its combination with the extended Kalman filter has been

also evaluated. In order to analyse the suitability of the GP and EKF-GP algorithms for prediction purposes, experiments have been conducted using the monthly gas consumption of U.S. States, in combination with temperature information.

In the domestic prediction experiments, the best performance for both water and natural gas prediction are achieved using only the GP model created for a time resolution of $24\,h$, with a set of input data composed of both water and natural gas consumptions. This results point out the evident correlation between water and natural gas consumptions. This information leads to prove the use of natural gas for the production of hot water for the evaluated household scenario. In confirmation of this, in the AMPds FAQ[4] within the appliances that use natural gas is reported the instant hot water unit.

The results obtained with the long period data have shown the effectiveness of the EKF-GP approach. The forecasting performance obtained with the GP model exhibited a general improvement applying the EKF method. As for the domestic forecasting, the combination of consumptions and temperatures has shown better results than the forecast computed using the consumptions data alone, denoting the advantage of heterogeneous data utilization as well as the strong correlation among them.

For this reasons, the authors will go beyond the preliminary status of the present work and develop further computational intelligence and machine learning techniques for comparative purposes. Moreover, as soon as the new year data of the AMPds will be released, this is allow a proper evaluation of the seasonality issue.

References

1. Azari, A., Shariaty-Niassar, M., Alborzi, M.: Short-term and Medium-term Gas Demand Load Forecasting by Neural Networks. Iranian Journal of Chemistry and Chemical Engineering (4), 77–84 (2012)
2. Babovic, V., Abbott, M.B.: The Evolution of Equations from Hydraulic Data Part I: Theory. Journal of Hydraulic Research 35(3), 397–410 (1997)
3. Bakker, M., Vreeburg, J., van Schagen, K., Rietveld, L.: A Fully Adaptive Forecasting Model for Short-term Drinking Water Demand. Environmental Modelling & Software 48, 141–151 (2013)
4. Bennett, N.D., Croke, B.F., Guariso, G., Guillaume, J.H., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V.: Characterising Performance of Environmental Models. Environmental Modelling & Software 40, 1–20 (2013)
5. Fagiani, M., Squartini, S., Gabrielli, L., Pizzichini, M., Spinsante, S.: Computational Intelligence in Smart Water and Gas Grids: An Up-to-date Overview. In: 2014 International Joint Conference on Neural Networks (IJCNN), pp. 921–926 (July 2014)

---

[4] http://ampds.org/

6. Hartikainen, J., Solin, A., Särkkä, S.: Optimal Filtering with Kalman Filters and Smoothers - A Manual for MATLAB Toolbox EKF/UKF Version 1.3. Department of Biomedical Engineering and Computational Science, Aalto University School of Science (2011),
   http://becs.aalto.fi/en/research/bayes/ekfukf/documentation.pdf

7. Kalman, R.E.: A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME – Journal of Basic Engineering 82(Series D), 35–45 (1960)

8. Liu, J., Chang, M.: Application of the Grey Theory and the Neural Network in Water Demand Forecast. In: 2010 Sixth International Conference on Natural Computation (ICNC), vol. 2, pp. 1070–1073 (2010)

9. Makonin, S., Popowich, F., Bartram, L., Gill, B., Bajic, I.V.: AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research. In: IEEE Electrical Power and Energy Conference, pp. 1–6 (2013)

10. Nasseri, M., Moeini, A., Tabesh, M.: Forecasting monthly urban water demand using extended kalman filter and genetic programming. Expert Systems with Applications 8(6), 7387–7395 (2011)

11. Pang, B.: The Impact of Additional Weather Inputs on Gas Load Forecasting. Ph.D. thesis, Marquette University (2012)

12. Silva, S., Almeida, J.: Gplab - A Genetic Programming Toolbox for MATLAB. In: Proc. of the Nordic MATLAB Conference (NMC-2003), pp. 273–278 (2005)

13. Spinsante, S., Pizzichini, M., Mencarelli, M., Squartini, S., Gambi, E.: Evaluation of the Wireless M-Bus Standard for Future Smart Water Grids. In: 9th International Wireless Communications and Mobile Computing Conference, pp. 1382–1387 (2013)

14. Spinsante, S., Pizzichini, M., Mencarelli, M., Squartini, S., Gambi, E., Piazza, F.: Wireless M-Bus Sensor Networks for Smart Water Grids: Analysis and Results. International Journal of Distributed Sensor Networks (2014) (to appear)

15. Squartini, S., Gabrielli, L., Mencarelli, M., Pizzichini, M., Spinsante, S., Piazza, F.: Wireless M-Bus Sensor Nodes in Smart Water Grids: The Energy Issue. In: Fourth International Conference on Intelligent Control and Information Processing, pp. 614–619 (2013)

16. Tabesh, M., Dini, M.: Fuzzy and Neuro-fuzzy Models for Short-term Water Demand Forecasting in Tehran. Iranian Journal of Science & Technology, Transaction B, Engineering 33(B1), 61–77 (2009)

17. Zhu, X., Xu, B.: Urban Water Consumption Forecast Based on QPSO-RBF Neural Network. In: Eighth International Conference on Computational Intelligence and Security, pp. 233–236 (2012)

Radial Basis Function Interpolation for Referenceless Thermometry Enhancement

Luca Agnello[1], Carmelo Militello[2], Cesare Gagliardo[3], and Salvatore Vitabile[3,4]

[1] Department of Chemical, Management, Computer, and Mechanical Engineering,
University of Palermo, Palermo, Italy
[2] Institute for Molecular Bioimaging and Physiology,
National Research Council (IBFM CNR-LATO), Cefalù, Italy
[3] Department of Biopathology, Medical and Forensic Biotechnologies,
University of Palermo, Palermo, Italy
[4] MIRC srl, Academic spin-off of the University of Palermo, Catania, Italy
luca.agnello@gmail.com, carmelo.militello@ibfm.cnr.it,
{fcesare.gagliardo,salvatore.vitabileg}@unipa.it

Abstract. MRgFUS (Magnetic Resonance guided Focused UltraSound) is a new and non-invasive technique to treat different diseases in the oncology field, which uses Focused Ultrasound (FUS) to induce necrosis in the lesion. Temperature change measurements during ultrasound thermo-therapies can be performed through magnetic resonance monitoring by using Proton Resonance Frequency (PRF) thermometry. It measures the phase variation resulting from the temperature-dependent changes in resonance frequency by subtracting one phase baseline image from actual phase. Referenceless thermometry aims to reduce artefacts caused by tissue motion and frequency drift, fitting the background phase outside the heated region. The aim of this contribution is to propose a novel background phase reconstruction method using Radial Basis Function (RBF) interpolation. The effectiveness of the method has been demonstrated by comparing it against the classical PRF shift and polynomial referenceless approach. The comparison evaluates temperature rises in uterine fibroids during MRgFUS treatments on a set of 10 patients.

Keywords: Radial Basis Function, Interpolation, Referenceless Thermometry, Artificial Neural Network, MRgFUS.

1   Introduction

Hyperthermia is a type of clinical treatment in which body tissues are exposed to high temperatures that can kill pathological lesion, like uterine fibroids [2]. In MRgFUS treatments [7][8], high temperatures are applied on local and small areas by using ultrasound beams that deliver energy to heat the tumour. MRgFUS treatment is performed using the ExAblate 2100 equipment (InSightec, Haifa, Israel), integrated with a Signa HTxt MR scanner (GE Medical Systems, Milwaukee, WI). Thermal ablation of fibroids tissue is done using sonication process: the tissue is heated with Focused

UltraSound concentrating a high-energy beam on a specific point. This allows to reach a temperature higher than 50°C causing proteins coagulation and consequently inducing the fibroid tissue necrosis.

The planning, treatment and evaluation processes are possible thanks to MR Imaging (MRI) guidance, which can also be used to reconstruct maps of tissues temperature. This makes it particularly enabling for guiding and monitoring thermal therapies. Temperature monitoring is feasible with MRI thanks to temperature sensitivity to MR parameters such as Proton Resonance Frequency, T1 and T2 relaxation times, Proton Density, Magnetization Transfer and so on. In a MRgFUS treatment, a little area of the patient's organ is heated by a focused ultrasound beam, increasing temperature in that point. The process is repeated several times until the whole lesion area is treated. In Fig. 1 it is possible to see a sequence of temperature maps: the sonication burst takes about 30 second for each sonication, and the number of sonications are related to position, type, and size of uterus fibroid.
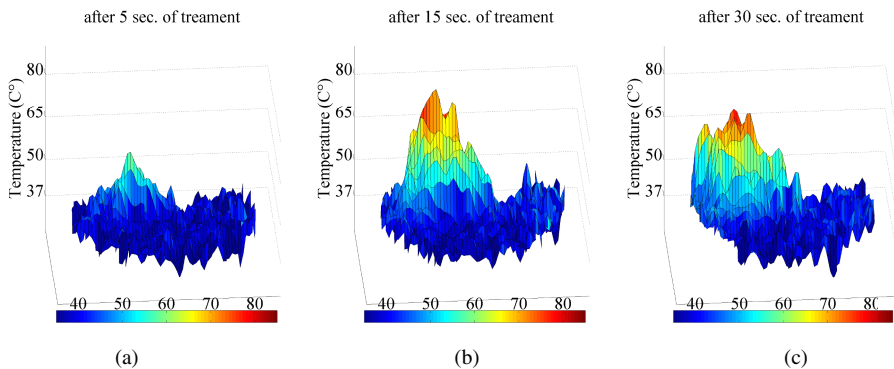
Fig. 1. (a) 3D plot of temperature map after 5 seconds treatment; the maximum temperature reached in the peak is about 52°C. (b) The temperature map after 15 seconds of sonication; (c) The temperature map after 30 seconds of sonication. The temperature peak is 87°C.

An interesting overview of MRI temperature methods is shown in [1]. Proton spectroscopic imaging, like PRF shift thermometry, uses phase mapping created from temperature-induced water proton chemical shift. MRI-derived temperature maps can be constructed using Gradient Recalled Echo (GRE) imaging sequences [5], by measuring the phase change resulting from the temperature-dependent change in resonance frequency. In order to eliminate temperature-independent artefacts such as magnetic field in-homogeneities, one or more baseline images are acquired before thermal therapy and subtracted from images during heating. The phase differences are proportional to temperature-dependent PRF shifts and the assessment of temperature rises ΔT is possible according to the following equation:

$$\Delta T = \frac{\varphi(T) - \varphi(T_o)}{\gamma \alpha B_o TE} \tag{1}$$

where $\varphi(T)$ is the phase map in the current temporal instant, $\varphi(T_0)$ is the baseline phase at a known temperature (i.e. 37°C), γ is the gyromagnetic ratio, $\alpha=-0.01ppm/C°$ is the PRF change coefficient, B_0 is the strength of the magnetic field, and TE is the echo time of the MR acquisition protocol.

Motion of anatomical region undergone to the MRgFUS treatment is one of the most prevalent problems for temperature monitoring with PRF phase mapping. Intra-scan motion is caused by movement of an object during MR image acquisition, resulting in a poor quality image with typical blurring and ghosting artefacts. Inter-scan motion is due to motion or displacement of an object between the acquisitions of consecutive images. Methods for temperature estimation in presence of motion can be divided into two categories: (i) methods based on a multi-baseline strategy and (ii) methods based on a referenceless strategy.

Multi-baseline methods take background phase information before heating at various position of the organs during the respiratory and/or cardiac cycle. The selection of the corresponding baseline image is performed by determining the organ's position [9][10][11].

Referenceless methods estimate heating from a treatment image itself, without a baseline image used as temperature reference. Supposing that the phase image has a smooth tendency under the heated area, this kind of methods fit polynomial functions [3] or uses a weighted least-squares fit [4] to the surrounding phase. The extrapolation of the reconstructed piece of baseline image is useful for background phase estimation, which is then subtracted from the actual phase to evaluate the phase difference after the heating caused by ultrasound sonication.

Considering that in classical PRF shift thermometry there are obvious problems of artefacts, most prevalently due to motion, and in referenceless thermometry the accuracy of the interpolation lacks in precision, a novel interpolation method is applied to the issue of the referenceless thermometry. This method has been successfully tested [16] using MRgFUS ablation on a ex-vivo animal muscle and reconstructing temperature maps using RBF interpolation methods. In this paper method has been applied to real in-vivo treatments of uterine fibroids, evaluating the baseline phase maps with great results.

The paper is organized as follows: in Section 2 theoretical background on Radial Basis Function (RBF) is introduced; in Section 3 the proposed interpolation model is presented; Section 5 illustrates the obtained experimental results; finally, in Section 6, some conclusions are reported.

2   Radial Basis Functions

The idea of RBF Networks derives from the theory of function approximation. One of the most used approaches in literature to address the interpolation problem is to fit data using a polynomial function. However, an invertible system that defines the interpolator is not guaranteed for all the interpolation points, and often shows spurious bumps. The main features of RBF interpolators are:

- they are two-layer feed-forward neural networks;
- each hidden nodes implement a radial basis function;

- the network training finds the weights from the input to hidden layer and then the weights from the hidden to output layer are calculated;
- the geometry of the input points is not restricted to a regular grid;
- the networks are very good for interpolation purposes, in particular if there are large areas of missing data.

The interpolation of N data points requires that each of the D dimensional input vectors $x^p = \{x_i^p : i = 1, 2, ..., D\}$ is mapped onto the corresponding output target t^p, finding a function $f(x^p) = t^p$, $\forall p = 1, 2, ..., N$. The RBF function approach uses a set of basis functions that are combined linearly:

$$f(x) = \sum_{p=1}^{N} w_p \phi(\|x - x^p\|) \tag{2}$$

The idea is to find the weights w_p so the function fits through the data points:

$$f(x^q) = \sum_{p=1}^{N} w_p \phi(\|x^q - x^p\|) = t^p \tag{3}$$

The non-linear function $\phi(\cdot)$ is the interpolating radial basis function kernel. Some of the most commonly used basis functions are:

Table 1. Common kernels for radial basis functions

Radial Basis Function	Expression	Constrain
Linear	$\phi(r) = r$	
Multi-Quadric	$\phi(r) = (r^2 + \sigma^2)^{1/2}$	width $\sigma > 0$
Thin-Plate-Spline	$\phi(r) = r^2 \ln(r)$	
Gaussian	$\phi(r) = \exp\left(-\dfrac{r^2}{2\sigma^2}\right)$	width $\sigma > 0$

The RBF network is composed of two layers, and the N training patterns $\{x_i^p, t^p\}$ determine the weights directly. The hidden layer multiplies the activation units as shown in Fig. 2.
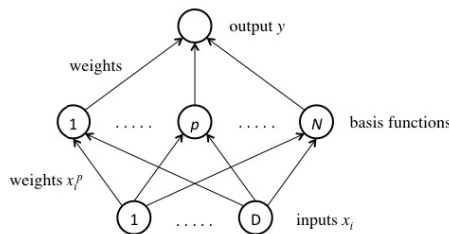
Fig. 2. The structure of a neural network implementing the RBFs as hidden layer

In this paper we are particularly concerned with 2D (depth-map) [12][13] data and we will consider Linear, Thin-Plate Spline and Multi-Quadratic interpolators.

The surveys of Powell and Light [14][15] are excellent references for the properties of radial basis functions. The σ value (in Multi-Quadric function) is responsible for the sensitivity of the interpolator. In the experiments, a very little value (of order 10^{-3}) is good for the interpolation purpose.

Moreover, often the data to interpolate are noisy. In presence of noise, one may consider to relax the exact interpolation requirement by means of regularization. This is possible by modifying the equation (2) as follows:

$$f(x) = \sum_{p=1}^{N} w_p \phi(\|x - x^p\|) + \lambda I \qquad (4)$$

adding a relaxation parameter λ that controls the amount of smoothing of the interpolation, and I is the identity matrix. In the $\lambda=0$ case, the equation is reduced to exact interpolation; in case that the parameter is highly regularized, the TPS model degenerates to the least-squares affine model [18].

3   The Workflow of the Proposed Interpolation Method

The MRgFUS treatment uses ultrasound beams that hit the interested organ. In this paper we have evaluated temperature reconstructions of treatments regarding women affected by single/multiple uterine fibroid. In Fig. 3, an uterine fibroid has been highlighted before the thermal treatment.

(a)	(b)	(c)
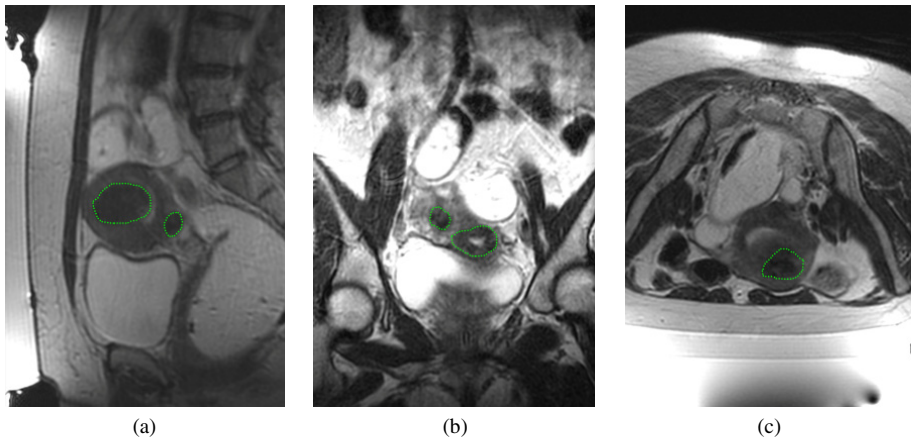
Fig. 3. An uterine fibroid in sagittal (a), coronal (b) and axial (c) view. The fibroid will be ablated with MRgFUS treatment

Referenceless thermometry estimates baseline phase from each acquired image phase and subtracts it from current image as in classical PRF thermometry [3].

In this work we have focused our attention to the effectiveness of the referenceless thermometry that uses a Polynomial for the interpolation [3], and we compared our proposed interpolation, that uses Radial Basis Functions, against the Polynomial one.

(a)   (b)   (c)
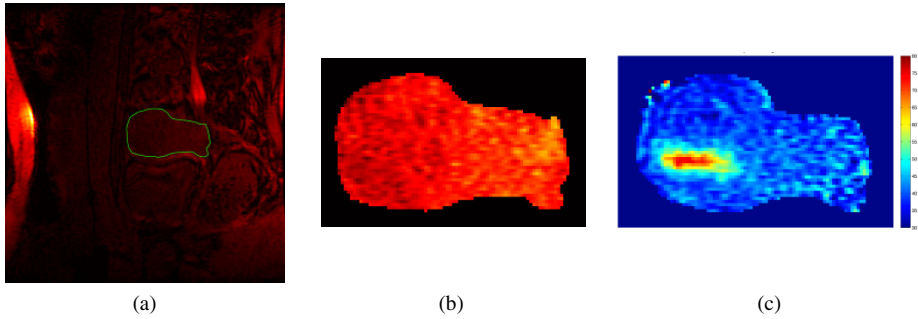
Fig. 4. (a) The area interested by the treatment; (b) Magnification of interested area; (c) Thermal map of the MRgFUS treatment. The blue area will be used for interpolation.

Here, our gold standard is the temperature extrapolated using classic PRF shift method [1]. As shown in Fig. 4, estimation is possible because in thermal therapy only a small region of the organ is affected by temperature change, and the phase outside the heated region can be used to determine the background phase.

For each temporal instant, the baseline phase below the heated area is evaluated interpolating the surrounding area using the Radial Basis Functions. The selection of a Region Of Interest (ROI) inside the treated organ (the female uterus in this case) makes possible to extract the heated area from the surrounding (not treated) area. The heated area that contains phase variations due to thermal treatments is removed, and the remaining area is used as input data to train the artificial neural network, as shown in Fig. 5. In correspondence of the sonication spot (Fig. 5a) the phase map shows a negative peak that represents a positive temperature variation (because the α coefficient is negative).

The proposed method (Fig. 6) for enhancing the referenceless thermometry by using RBF interpolation has been implemented as follows:

1. once the series of images is acquired, we recover the original phase from the 2π-wrapped phase images by using the Goldstein, Zebker and Werner's algorithm [6];
2. the RBF artificial neural network takes input data from the region between the sonicated area and the uterine contour;
3. the area to be reconstructed is iteratively interpolated by using RBF, which represents a practical solution for the problem of interpolating incomplete three-dimensional surfaces. The implementation of the reconstruction algorithm invokes iterative refinement to improve the accuracy of the solution;
4. for each temporal instant the extrapolated baseline phase that is used together with the global (currently heated) image.

This follows the PRF principles used for temperature rise assessment. The RBF network interpolates the masked area according to the specific radial basis functions: Linear (or Euclidean), Multi-Quadratic, Thin-Plate Spline, etc., solving it by using the double precision diagonal pivoting method from Lapack [17].

(a)

(b)

(c)

(d)
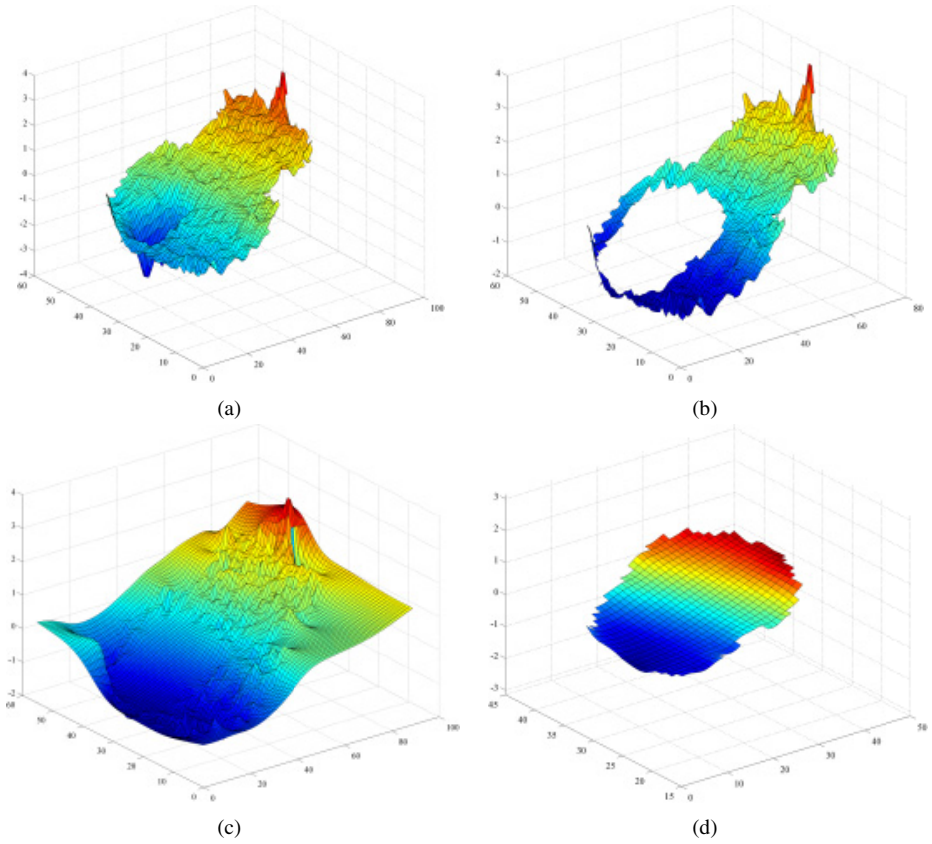
Fig. 5. (a) A 3D plot of the whole phase information of the organ: with a depression corresponding to the sonication shot (blue area); (b) the region affected by the sonication is removed, and the surrounding area is used to train the radial basis function network; (c) the reconstructed baseline obtained through the proposed interpolation algorithm (in this case it is used the Thin-Plate Spline kernel); (d) magnification of the interpolated area
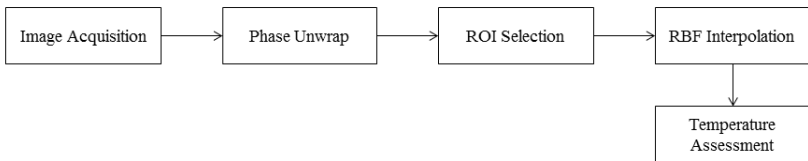
Fig. 6. The workflow of the proposed method

We found experimentally that the better interpolator for this particular kind of noisy data is the Linear RBF, but often during experimentation we found that also TPS and Multi-Quadratic have shown good results. The baseline phase is now reconstructed for each temporal instant and using it in conjunction of the phase image, according to PRF equation (1), we obtain the temperature rise of the whole sonication.

4   Experimental Results

Ten MR datasets related to ten female patients undergone to MRgFUS treatments for ablation of intra-uterine fibroid have been processed and evaluated. All the MR images are acquired by a GE Signa HDxt 1.5 Tesla scanner, and the ultrasound sonications are performed by an Insightec ExAblate 2100 system. Each hyperthermia sonication takes several seconds to focus high power ultrasounds in the chosen focal point. During each sonication the MR scanner records about 8-12 temporal instants, and each of them is composed of a tern of morphological-real-imaginary images. The real and imaginary parts are combined together to reconstruct phase maps. The evaluation of our approach was performed by calculating Root Mean Square (RMS) errors between the original baseline and each reconstructed (Polynomial and our Radial Basis Functions) interpolation, and calculating the differences (in C°) of the mean temperature value between the original PRF temperature and those provided by polynomial and our RBF approach. The kernels here used are the Euclidean, Thin-Plate Spline and the Multi-Quadratic one (Fig. 7).
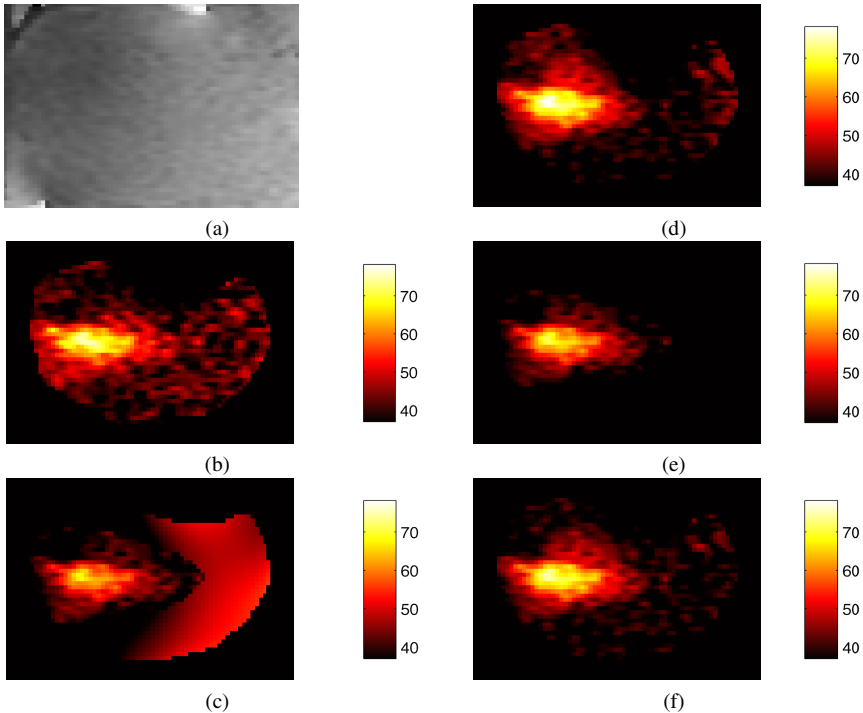
Fig. 7. Temperature reconstruction for a temporal instant during MRgFUS treatment: (a) the morphologic MR image; (b) temperature assessment using the classical PRF shift method; (c) temperature assessment using the Polynomial method; (d) temperature assessment using the Linear RBF method; (e) temperature assessment using the Multi-Quadratic RBF method; (f) temperature assessment using the Thin-Plate Spline RBF method. The depicted values are in °C.

The natural criterion for assess a reconstructed phase image is how closely it matches the baseline surface prior to the removal of the heated area. The interpolator fitted to the incomplete phase-map is then compared with the original baseline surface. Obtained temperature assessments in a MRgFUS treatment for the ablation of a uterine fibroid are shown in Fig. 8a. In this figure the RMS error shows that RBF reconstructions (Linear and Multi-Quadratic) has better results with respect to Polynomial reconstruction, assuming that the PRF temperature is the gold standard. Results show a huge increase of precision on the whole reconstructed area. These results are confirmed in Fig. 8b, where all the mean temperatures of the treated areas related to thermal treatments of all patients have been compared to PRF temperature.
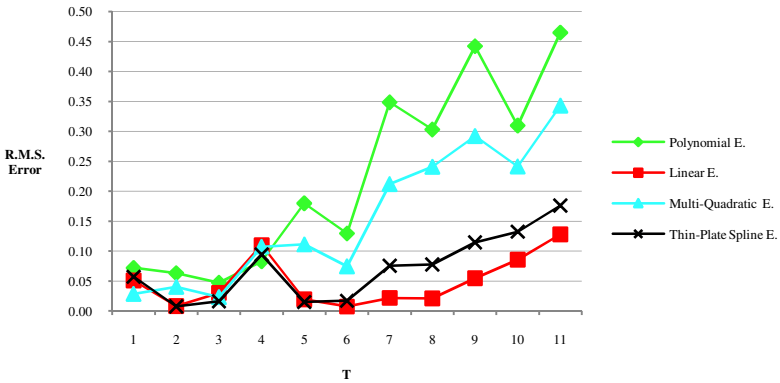
(a)
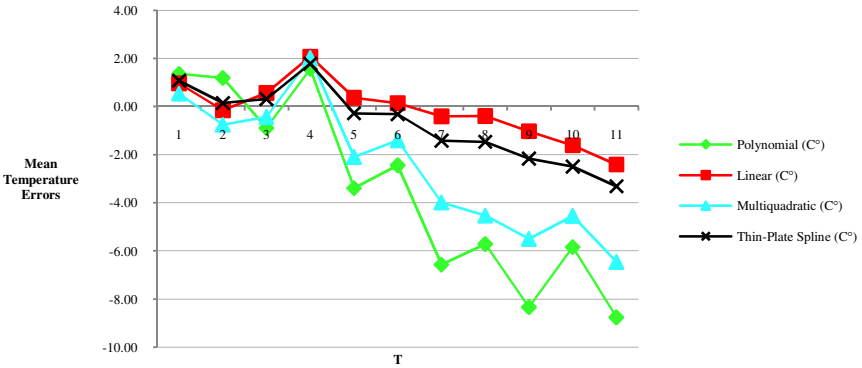
(b)

Fig. 8. (a) RMS errors for different kind of reconstruction methods compared to classical PRF Shift thermometry; (b) Mean temperature errors (°C) of the whole area hit by thermal treatment.

In Fig. 9a is depicted to see the temperatures evaluation in a random chosen point of the treatment area. All the RBF-based reconstructed temperatures (the blue, cyan, and black lines) runs very close to the gold standard PRF temperature (red line); we cannot say the same for the polynomial interpolation (green line). This demonstrates that radial basis functions are a very good kind of interpolator for this type of noisy data, even if there are large regions with missing data.

(a)

(b)
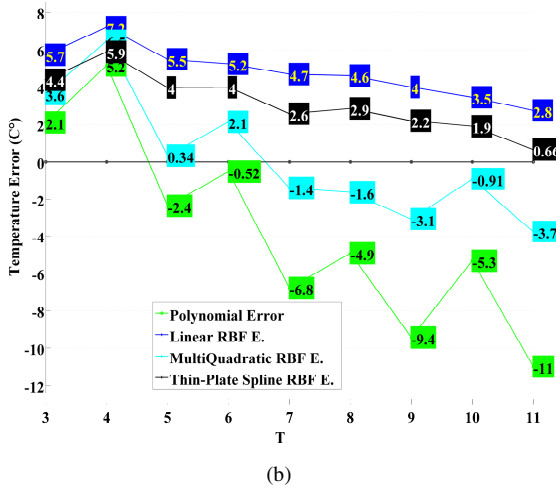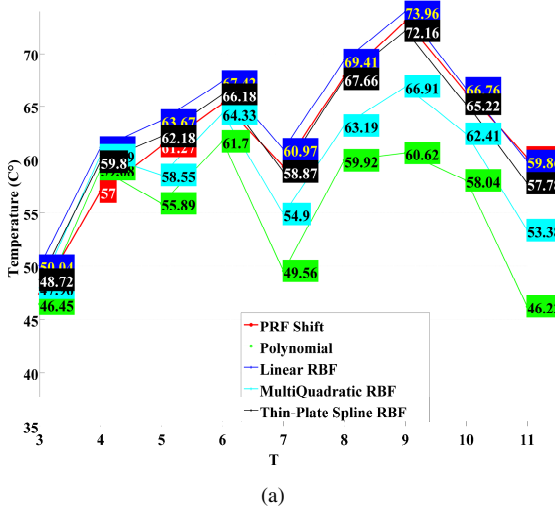
Fig. 9. Temperature behaviour in a point of the treatment area: (a) temperature rise (in °C) for a treatment of about 32 seconds. The red line is the reference *PRF* temperature, the green line is the *Polynomial* reconstructed temperature; the black *(Thin-Plate Spline)*, blue *(Linear)* and cyan *(Multi-Quadratic)* lines are the RBF-based reconstructed temperatures. (b) The variation (error) of reconstructed temperatures compared with the *PRF* temperature.

Fig. 9b confirms the goodness of the RBF reconstruction: for example, in the ninth temporal instant, the PRF temperature is 73.04°C. The RBFs temperatures differs of 3-4°C, while the Polynomial temperature is about 10°C. less. In a MRgFUS treatment, this can lead to continue the sonication process even if it is not necessary, surely causing pain to the patient and possible damages in surrounding tissues. In conclusion, the RBF reconstruction method gains all the advantages of referenceless thermometry avoiding lacks of precision of the Polynomial interpolation temperature reconstruction.

5   Conclusion

RBF neural networks are a good and flexible tools that allow for the reconstruction of unknown data. The effectiveness of the proposed approach has been demonstrated using 10 MR dataset of 10 female patients undergone to uterine fibroids ablation MRgFUS treatments. Polynomial reconstruction can over/under estimate the temperatures: this can lead to break the sonication before reaching the temperature established. The risk is the missing proteins denaturation, pain inducted in patients, and damage to surrounding tissues. RMS errors and temperature differences show a huge increase of precision in comparison with other kind of interpolators.

Future works will investigate the real precision of the PRF method, by measuring real temperature rises in MRgFUS treatments using thermocouples or optical fibres inserted in a phantom and acquiring the phase variations induced by the heating process. Since the reconstruction method is heavily dependent from ROIs selection, we are also investigating automatic methods for organ and sonication spot segmentation[8]. The integration of this RBF-based interpolation method with automatic segmentation approaches could reduce the operator-dependence of the algorithm and, consequently, the final error in the temperature reconstruction.

References

1. Rieke, V., Butts, K.: MR Thermometry. Journal of Magnetic Resonance Imaging 27, 376–390 (2008)
2. Kim, J.H., Hahn, E.W.: Clinical and biological studies of localized hyperthermia. Cancer Res. 39, 2258–2261 (1979)
3. Rieke, V., Vigen, K.K., Sommer, G., Daniel, B.L., Pauly, J.M., Butts, K.: Referenceless PRF shift thermometry. Magn. Reson. Med. 51, 1223–1231 (2004)
4. Kuroda, K., Kokuryo, D., Kumamoto, E., Suzuki, K., Matsuoka, Y., Keserci, B.: Optimization of self-reference thermometry using complex field estimation. Magn. Reson. Med. 56, 835–843 (2006)
5. Ishihara, Y., Calderon, A., Watanabe, H., Okamoto, K., Suzuki, Y., Kuroda, K., Suzuki, Y.: A precise and fast temperature mapping using water proton chemical shift. Magn. Reson. Med. 34, 814–823 (1995)
6. Goldstein, R.M., Zebker, H.A., Werner, C.L.: Satellite radar interferometry: two-dimensional phase unwrapping. Radio Sci. 23, 713–720 (1988)

7. Hurwitz, M., Machtinger, R., Fennessy, F.: Magnetic resonance-guided focused ultrasound surgery for treatment of painful osseous metastases. In: Progress in Biomedical Optics and Imaging - Proceedings of SPIE, vol. 7901, art. no. 79010M (2011)

8. Militello, C., Vitabile, S., Russo, G., Candiano, G., Gagliardo, C., Midiri, M., Gilardi, M.C.: A semi-automatic multi-seed region-growing approach for uterine fibroids segmentation in MRgFUS treatment. In: Proceedings of 7th International Conference on Complex, Intelligent, and Software Intensive Systems, pp. 176–182 (2013), doi:10.1109/CISIS.2013.36

9. Vigen, K.K., Daniel, B.L., Pauly, J.M., Butts, K.: Triggered, navigated, multi-baseline method for proton resonance frequency temperature mapping with respiratory motion. Magn. Reson. Med. 50, 1003–1010 (2003)

10. Shmatukha, A.V., Bakker, C.J.G.: Correction of proton resonance frequency shift temperature maps for magnetic field disturbances caused by breathing. Phys. Med. Biol. 51, 4689–4705 (2006)

11. Roujol, S., Ries, M., Quesson, B., Moonen, C., de Senneville, B.D.: Real-time MR-thermometry and dosimetry for interventional guidance on abdominal organs. Magn. Reson. Med. 63, 1080–1087 (2010)

12. Beatson, R., Newsam, G.: Fast evaluation of radial basis functions: I. Comput. and Math. with Applicat. 24(12), 7–19 (1992)

13. Carr, J.C., Fright, W.R., Beatson, R.K.: Surface Interpolation with Radial Basis Functions for Medical Imaging. IEEE Transactions on Medical Imaging 16(1) (February 1997)

14. Powell, M.J.D.: The theory of radial basis function approximation in 1990. In: Light, W.A. (ed.) Advances in Numerical Analysis II: Wavelets, Subdivision Algorithms and Radial Functions, pp. 105–210. Oxford Univ. Press, Oxford (1992)

15. Light, W.A.: Some aspects of radial basis function approximation. In: Singh, S.P. (ed.) Approximation Theory, Spline Functions and Applications, pp. 163–190. Kluwer, Dortrecht (1992)

16. Agnello, L., Militello, C., Gagliardo, C., Vitabile, S.: Referenceless Thermometry using Radial Basis Function Interpolation. In: World Symposium on Computer Applications & Research, WSCAR (2014), doi:10.1109/WSCAR.2014.6916834

17. Lapack Users Guide, Society for Industrial and Applied Mathematics, Philadelphia, PA (1992)

18. Belongie, S., Malik, J., Puzicha, J.: Shape Matching and Object Recognition Using Shape Contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(4), 509–522 (2002)

A Grid-Based Optimization Algorithm for Parameters Elicitation in WOWA Operators: An Application to Risk Assesment

Marta Cardin and Silvio Giove

Department of Economics, Ca' Foscari University of Venice, Venice, Italy
`sgiove@unive.it`

Abstract. In this paper we propose a joint grid-based and stochastic search for parameters elicitation in the case of WOWA aggregation functions. The method uses a grid search approach to determine the parameter of a monotonic quantifier, and for each of the values, a stochastic search in the space of the criteria weights minimizes the sum of the quadratic error between the computed and the real output of a learning set. A simulated application is proposed in the case of transportation risk assessment, using an *ad hoc* questionnaire applied to risk matrices.

Keywords: Risk assessment, OWA, WOWA.

1 Introduction

Criteria aggregation is widely used in applied sciences, as the process by which different data have to be aggregated into a single positive number, like signals, sensors, but also human judgements - each of them normalized into a common scale [6]. The specialized literature provides many different approaches for aggregation, starting from the simplest ones, the Weighted Averaging (WA) up to other more sophisticated, like Generalized Means, Ordered Weighted Averaging (OWA), Weighted Ordered Weighted Averaging (WOWA), fuzzy measures, and other ones [2], [21].

Every method requires the assessment of some parameters, for instance, in the case of WA the *weights* of the criteria need to be elicited, representing the trade-off between the criteria themselves. Given a collection of input-output data, this is usually done considering a mathematical optimization problem. WA is unable to represent *interactions* among the criteria, as *synergies* and *conflicts*, because WA is a complete *compensative* method. For this reason, other more general approaches were proposed, as the one based on *fuzzy measures*, see Section 2 below. We note that such methods require many parameters to be elicited, increasing the computational complexity. Thus the literature proposed other approaches, which include interactions among the criteria but, at the same time, reduce the numerical complexity. Among them, we focus the attention on OWA introduced by Yager, [15], [16], [17] , [18], [19] and on the extension of OWA, the WOWA aggregation operators, introduced by Torra, [12], [13]. As for the WA case, OWA

weights can be directly elicited from an input-output data collection, using mathematical programming, but the same approach cannot be applied for the WOWA, see Section 3. For this reason, we propose an innovative elicitation approach, based on two phases, a grid-based approach and a stochastic search method. In the first phase the characterizing parameter of a Monotone Quantifier moves along a predetermined grid, while in the second phase, for each value of this parameter the criteria weights are sampled from an uniform distribution. At each iteration, we store the quadratic error between the computed WOWA and the real output value. The best parameter values are selected as the ones which minimizes the quadratic error. The algorithm is applied to an environmental risk assessment problem, where different types of risk need to be aggregated, as described in [10]. In this case, we proposed a suitable questionnaire to be fulfilled by a (simulated) Decision Maker, which has to assign a score to every scenario in the questionnaire, as described in Section 4. The optimization problem is solved using the two phases approach, and the numerical results showed good approximation results between the computed WOWA and the real output (questionnaire scores). The same approach can be applied to any other WOWA elicitation parameters if a data collection is available.

The paper is structured as follows. Section 2 introduces Aggregation Operators, mainly focusing on OWA and WOWA operators. Section 3 reports the parameters identification algorithm, based on a grid search and on Monte-Carlo simulation. In Section 4 is reported an application of the algorithm to a risk assessment problem. The last Section concludes and presents possible extensions and future work.

2   Aggregation Functions; OWA and WOWA

Aggregation Operators (AO) are widely used in many problems where different numerical values coming from different sources need to be aggregated together in order to obtain an overall, single output, see [1] [2] [6], [21]. An AO is a monotone and idempotent function: $F : [0,1]^n \to [0,1]$, such that $F(0,0,...,0) = 0, F(1,1,...,1) = 1$ (border conditions). Every AO requires some parameters to be elicited, that is not an easy task, with exception of few simple cases. Using a notation similar to [7], the WA operator is defined as follows:

$$WA_p(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} p_i x_i \tag{1}$$

being $\{x_i\}$ the normalized value of the criteria, $i = 1, 2, ..., n$, and $\{p_i\}$ a *weights*, $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$. The weights represent the *relative* importance of the criteria, and need to be determined by direct assignment or using a learning procedure. The WA is *compensative* and *homogeneous* of degree 1, [7].

In many applications *compensativeness* is not a desirable property, given that a low value of a criteria cannot be compensated by an high value of an other one; for instance in the field of *sustainability* evaluation, an economic development cannot be paid by a critical value of toxicity [11].

While WA is a complete *compensative* method, it cannot represents *inter-actions* among the criteria - *synergies* and *conflicts*. On the other side, other methods can overcame this limit. Among them, let us recall the Non Additive Measures (*Fuzzy Measures*) and the *Choquet* integral, see [2],[4], [8], a very general approach but requiring an huge number of parameters to be elicited. An other class of AO is the family of OWA operators, introduced by Yager [15], [16], [17], [18], [19]. For a recent review of OWA operators see [20].

2.1   The OWA Operator

An OWA operator is defined as follows [15], [16]:

$$OWA^w(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} w_i x_{\sigma(i)} \tag{2}$$

being σ a permutation of the index set $\{1, 2, ..., n\}$ such that $x_{\sigma(1)} \geq x_{\sigma(2)} \geq ... \geq x_{\sigma(n)}$, thus it is linear w.r.t. the *ordered* values of the criteria. Modulating the values of the weights, different AO can be obtained; if $w_1 = 1, w_i = 0, \forall i = 2, ..., n$, the MAX operator is obtained, while if $w_n = 1, w_i = 0, \forall i = 1, ..., n - 1$, we obtain the MIN operator. If $w_i = \frac{1}{n}, \forall i = 1, ..., n$ the OWA collapses to the simple averaging. Again, the median, the k-th order statistic, the Hurwicz operator and many others AO can be obtained with different weight choice. OWA operators are continuous, symmetric and homogeneous of degree 1, and represent a special case of the Choquet integral w.r.t. a symmetric non additive measure [1], [18]. An OWA operator can be *andness*-type (*orness*-type) if the aggregated value is more or less close to the minimum (maximum) of its arguments. The following *orness* index is defined as:

$$Orness(OWA^w) = \sum_{i=1}^{n} w_i \frac{n-1}{n-1} \tag{3}$$

We obtain *Orness*= 1 in the *optimistic* case, corresponding to the MAX case, the logic quantifier *At least one*; while *Orness*= 0, in the *pessimistic* case) corresponding to the MIN case, the quantifier *All*.

Different numerical procedure were proposed to elicit the OWA weights [1]. If a collection of input-output data is available, linear or quadratic programming can be applied, as for the case of WA; apart the re-ordering of the input vector, the aggregation function is linear w.r.t. the weights.

Alternatively, the parameters can be elicited using a *Monotonic Quantifier*, that is a a *fuzzy* linguistic quantifier, which can be associated to the term sets "it exists", "most", "at least half", "for all" and so on [17], [19]. More precisely, a Regular Increasing Monotonic Quantifier (RIM) is a continuous and monotone function $Q : [0, 1] \to [0, 1]$ such that $Q(0) = 0, Q(1) = 1$, from which the OWA weights can be generated as follows:

$$w_i = Q(\frac{i}{n}) - Q(\frac{i-1}{n}) \tag{4}$$

The following parameterized family of RIM is widely used in applications, with $\alpha \geq 0$ [16] [19] [2]:

$$Q(z) = z^{\alpha} \tag{5}$$

An *andness* behavior is obtained for $\alpha \simeq 0$, while if $\alpha \to +\infty$ there is a tendency to *orness*. Tuning the value of α we move from the extreme *non compensative* tendency to to the opposite one, a *complete compensative* behavior. Thus a possible procedure to numerically optimize the α value consists into the computation of OWA with many different values of α, selecting the value which minimizes the quadratic average error; this phase will be better detailed in the next Section.

2.2 Extending OWA to WOWA

An extension of OWA operator was proposed in the recent literature, in such a way to include even the *importance* criterion weight, given that an OWA operator cannot implement the WA but only *simple* averaging. The result is a new operator, the Weighted OWA (WOWA), which uses two different sets of weights, the first referring to the criteria weights (*importance* weights), while the second refers to the *order* of the criteria (*OWA* weights), [12] [13]. We limit to note that given a Monotonic Quantifier $Q(z)$ and a set of weights $\{p_i\}$, the corresponding WOWA operator can be computed as follows:

$$WOWA_p^Q(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} w_i x_{\sigma(i)} \tag{6}$$

being σ a suitable permutation of the index set $(1, 2, ..., n)$ such that $x_{\sigma(1)} \geq x_{\sigma(2)} \geq ... \geq x_{\sigma(n)}$ and:

$$w_i = Q(\sum_{j=1}^{i} p_{\sigma(j)}) - Q(\sum_{j=1}^{i-1} p_{\sigma(j)}) \tag{7}$$

This way, both the relative importance of the criteria and the tendency to *optimism* can be taken into account. Finally let us observe that while OWA operators correspond to *symmetric* non additive measures, this is not necessarily true for WOWA [12], [13].

3   The GBS Approach for WOWA Parameters Elicitation

As above pointed to, the OWA weights can be elicited from an input-output data collection using quadratic or linear programming [15], [16], [3], [20]. This is not possible for WOWA operator, since by definition, the problem of weights

elicitation is NP-hard in function of the number of criteria, splitting the problem into 2^n sub-problems, each of them subjected to one of all the possible order of the criteria. Then the the numerical complexity can be unacceptable, and for this reason we propose a two-phases approach based on a grid search and a stochastic optimization, the GBS algorithm (Grid-based and Stochastic optimization), partially following what suggested for a CES function[1] [5]. Using a suitable questionnaire, a learning set is defined, formed by a collection of K input-output data (*scenario*); the cardinality of each scenario is $n+1$. After having defined a *grid* for the parameter α, we run many sub-iterations, randomly extracting the values of the n weights $\{p_i\}$ from an uniform distribution in $[0, 1]$. In each sub-interaction, the WOWA aggregated values are computed for each scenario in the learning set.

The average quadratic error, obtained averaging the square difference between the WOWA computed value and the corresponding output in the learning set, is then minimized, storing at each interaction the best value of the parameters (α and $\{p_i\}$). The procedure is repeated until all the grid values are considered. The algorithm is structured as follows:

INPUT DATA
1) $LS = \{x_1(k), x_2(k), ..., x_n(k); y(k)\}, k = 1, ..., K$: learning set ($x_i(k)$: input values, $i = 1, .., n$, $y(k)$: output value);
2) $Gr = \{\alpha_j\}, j = 1, ...Ng$: vector containing the values of α, the grid elements[2];
iii) N_{it}: number of iteractions for each value of α.

OUTPUT DATA
1) $\alpha^*, weight^*$: optimal values of parameters α and criteria weight 2) *Error*: absolute average error between computed and learning data

GBS ALGORITHM

a) Set $Error = \infty$
b) $\forall \alpha \in GR$:
c) $\forall i = 1, .., N_{it}$
d) $p_i \leftarrow U[0, 1], j = 1, .., n$
e) Compute $Q^\alpha(j), j = 1, .., n$
f) Compute $Agg(k) = WOWA_p^Q(x_1(k), x_2(k), ..., x_n(k)), \forall k = 1, .., K$:
g) Compute $\triangle = \frac{1}{K} \sum_{k=1}^{K} |Agg(k) - y(k)|$
h) If $\triangle < Error$ then $Error \leftarrow \triangle, \alpha^* \leftarrow alpha, weight^*(i) \leftarrow p_i, i = 1, .., N_{it}$

---

[1] A CES is a Constant Elasticity of Substitution function.
[2] In the numerical simulations a constant-step grid has been used, that is $\alpha_j = \frac{j}{Ng}$.

4   Transport Risk Evaluation; Simulated Questionnaire and Parameters Elicitation

Environmental risk assessment requires the aggregation of different sources and types of risk. Among other, let us quote a recent case study in transport risk as reported by [10], where four different types of risk have to be aggregated into a single Risk Index. In the quoted papers the Authors applied different methods, from simple averaging up to Generalized Mean, comparing the obtained results, but no parameter optimization was carried on. Anywise risk aggregation, is multidimensional, and a *preference structure* is required, normally elicited from one or a set of Expert(s). In this contribution we develop the elicitation of the preference structure using a suitable questionnaire to be fulfilled by one Expert[3], including the personal (subjective) attitude to risk, that is the *way of thinking* of the representative Expert. We suppose that the Expert's preference structure can be represented by an WOWA operator, a very general aggregation tool, see above. Basing on the quoted paper, the available information is formed by four *risk matrices*, [10] and the references therein. The risk matrices refers to four types 9of risk, *Assets, Reputation, People, Environment* (**A,R,P,E** for brevity). For each of the four types, a two-entries matrix is assigned, whose dimensions are *Likelihood* and *Severity* of the damage. The Likelihood is partitioned into five linguistic labels, *Improbable, Remote, Occasional, Probable,Frequent*, while the Severity is defined by *None, Negligible, Minor, Moderate, Significant, Severe*. Each cell of the matrix contains a numerical evaluation of the risk, a number associated to the risk evaluation, in a common scale, say $1, 10$ where 1 corresponds to the null risk, and 10 to the maximum risk, the two extreme situations, the best and the worst.

After having assigned these values for the four risk matrices, every spatial region is characterized by a four-dimensional vector whose elements are the four risk assessment values. This way, given a risk scenario, thgat is a couple Likelihood-Severity for each of the four risk types, a single cell is activated for each of the four matrix. For instance the vector $(2, 10, 1, 5)$ corresponds to the case where there is a Low risk for *Assets*, a Very High risk for *Reputation*, a Null Risk for *People*, and a Medium risk for *Environment*.

Given a scenario, the four risk values are to be aggregated into a single number. In [10] this problem is approached using different types of Generalized Mean, but in the quoted paper no suggestion is given to the parameters optimization. A part the case of dominance, there is no an *objective* way to put the single risk values together, thus the aggregation operator (WOWA) needs to reflects the *preference structure* of the Expert. For this reason, we propose a method that uses a *preference structure* implicitly obtained by a questionnaire designed *ad hoc*, submitted to the Expert. Using the obtained answers, the GBS algorithm was applied to optimize the parameters of the WOWA operator. The structure of

---

[3] Normally a set of Expert should be considered applying Multi-Experts Decision Making and consensus analysis. Nevertheless this is not the focus of our proposal; namely we consider a *representative Expert*.

the questionnaire is described below. The following Section reports the structure of the questionnaire, together with some numerical text showing the satisfactory properties of the GBS algorithm in the case study.

5   Numerical Tests

This Section reports the results of some simulated scenarios, formed by a set of questions, each of them referring to the four risk types. The last column is the DM answer in the same scale $(1, 10)$ with 1 corresponding to Null (global) risk, 3 to Low Risk, and so on. The structure of the questionnaire is presented in the Table 1. Each row in the Table represents a scenario, that is a simulated question, thus 9 scenarios are considered at all. The first four columns report the instances of the four considered risk types $(\mathbf{A,R,P,E})$ in the normalized scale $(1, 10)$, and the three last columns a couple of numbers in parenthesis, in order the values fulfilled by the Expert, and the values obtained by the algorithm. Each of the last three columns refers to three different simulated DMs, each of them more or less *andness*-oriented. In the first case we simulated a *linear* Expert with equal weights. In the second case we maintain equal importance weights, but simulated an *andness*-oriented Expert, while in the last text we modified the weights but with same *andness* as for the second DM. The results obtained by the GBS algorithm showed good correspondence with the data fulfilled by each Expert. The obtained parameters reflected the preference structure with satisfactory precision. Moreover, the WOWA operator exhibits good generalization capabilities, able to represent the non linear relationship of simulated DM, at least for the simulated cases. Three situations are considered: a *Linear* DM with equal importance weights (column n. 5), a *Conjunctive* DM with equal importance weight (column n. 6) and finally a *Conjunctive* DM with unequal importance weight (column n. 7); in this case, the weights were $(0.27, 0.13, 0.07, 0.53)$. Thus, for instance, the couple $(1.45, 1.43)$ in the last column and row n. 5 means that the DM assigned to the $5 - th$ scenario the value 1.45 and the algorithm result is 1.43, for the *Conjunctive* DM with unequal weights. In the three cases, the value of α were $(1, 0.05, 0.5)$ respectively.

Table 1. Questionnaire structure, simulated and computed output in three cases

(R1)	(R2)	(R3)	(R4)	Linear	Conjunctive E.W.	Conjunctive NOT E.W.
1	1	5	10	(4.25,4.23)	(1.21,1.21)	(1.37,1.35)
10	2	1	2	(3.75,3.83)	(1.18,1.18)	(1.25,1.26)
7	3	2	8	(5.00,4.98)	(2.22,2.22)	(2.47,2.47)
5	4	6	7	(5.50,5.52)	(4.12,4.12)	(4.18,4.18)
1	8	10	8	(6.75,6.70)	(1.50,1.51)	(1.45,1.43)
8	7	1	1	(4.25,4.22)	(1.22,1.21)	(1.17,1.18)
2	1	3	1	(1.75,1.80)	(1.05,1.05)	(1.02,1.03)
2	1	3	1	(4.75,4.73)	(4.05,4.05)	(4.02,4.02)
4	6	5	4	(7.00,7.04)	(6.08,6.08)	(6.10,6.10)

The relative absolute errors (between the real and the computed output) are respectively $(0.007, 0.001, 0.003)$, clearly negligible in all the considered cases.

6   Conclusion

The WOWA operator extends the OWA tool considering also the *importance* of the criteria, thus combining both the cardinal properties of an averaging operator with the ordinal properties of OWA. For this reason, WOWA can be well tailored for many application problems in the field of Decision Theory. Nevertheless, few efforts were devoted, at our knowledge, in the parameters elicitation. In this contribution we proposed a combined grid-search and stochastic optimization algorithm (GBS algorithm), based on implicit knowledge elicitation, with the aim to obtain both the importance weights and the characterizing parameter of a Monotone Quantifier, inferring the preference structure of the Decision Maker using an *ad hoc* questionnaire. This approach has been applied to a transportation risk assessment, but it could be applied to other type of risk. Some preliminary texts confirmed the goodness of the method. In a future work, the algorithm will be extended to the case where *a priori* information will be available, like the distribution of the importance weights, which could be preliminarily inferred from the Decision Maker. Again, we intend to apply the GBS algorithm to Group Decision Theory, combining judgements of several Experts, partially following what propose by [9], [22] in the field of OWA operator.

References

1. Beliakov, G., Pradera, A., Calvo, T.: Aggregation Functions: a guide for practitioners. STUDFUZZ, vol. 221. Springer, Heidelberg (2007)
2. Calvo, T., Mayor, G., Mesiar, R.: Aggregation Operators: new trends and applications. STUDFUZZ, vol. 91. Springer, Heidelberg (2002)
3. Filev, D., Yager, R.R.: On the issue of obtaining OWA operator weights. Fuzzy Sets and Systems 94, 157–169 (1998)
4. Grabisch, M., Roubens, M.: Application of the Choquet integral in multicriteria decision making. In: Grabisch, M., Murofushi, T., Sugeno, M. (eds.) Fuzzy Measures and Integrals Theory and Applications, pp. 348–374. Physica Verlag (2000)
5. Henningsen, A., Henningsen, G.: Om estimation of the CES production function - Revisited. Economic Letters 115, 67–69 (2012)
6. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers (2000)
7. Llamazares, B.: On generalization of weighet means and OWA operators. In: Proc. of EUSLAT-LFA 2011 (2011)
8. Marichal, J.: An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. The IEEE Transactions on Fuzzy Systems 8(6), 800–807 (2000)
9. Palomares, I.: Applying OWA Operators and RIM Quantifiers to measure Consensus with Large Group of Experts. In: Proc. 6th International Summer School on Aggregation Operators - AGOP 2011 (2011)

10. Pérez, R., Alonso, P., Daz, I., Montes, S.: Aggregation of Risk Assesment Matrices for Umnao Reliability in Transportation Systems. In: De Baets, B., Fodor, J., Montes, S. (eds.) Proc. of EUROFUSE20- Uncertainty and Imprecision Modelling in Decision Making, pp. 225–232. Ediciones de la Universidad de Oviedo (2013)
11. Pinar, M., Cruciani, C., Giove, S., Sostero, M.: Constructing the FEEM sustainability index: A Choquet integral application. Ecological Indicator 39, 189–202 (2014)
12. Torra, V., Lv, Z.: On the WOWA operator and its interpolation function. International Journal of Intelligent Systems 24, 1039–1056 (2009)
13. Torra, V.: The weighted OWA operator. International Journal of Intelligent Systems 12, 153–166 (1997)
14. Torra, V., Narukawa, Y.: Modeling Decisions. Springer (2007)
15. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. on Systems, Man and Cybernetics 18, 183–190 (1988)
16. Yager, R.R.: Applications and extensions of OWA aggregation, Internat. Journal Man Machine Studies 37, 103–132 (1992)
17. Yager, R.R.: Families of OWA operators. Fuzzy Sets and Systems 59, 125–148 (1993)
18. Yager, R.R., Kacprzyk, J. (eds.): The ordered Weighted Averaging Operators, Theory and Applications. Kluwer Academic Publisher (1997)
19. Yager, R.R.: Quantifier guided aggregation using OWA operators. International Journal of Intelligent Systems 11, 49–73 (1996)
20. Yager, R.R., Kacprzyk, J., Beliakov, G. (eds.): Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice. STUDFUZZ, vol. 265. Springer, Heidelberg (2011)
21. Xu, Z.S., Da An, Q.L.: overview of operators for aggregation informations. International Journal of Intelligent Systems 18, 953–969 (2003)
22. Wang, Y.M., Parkan, C.: A minimax disparity approach for obtaining OWA operator weights. Inform. Sci. 175, 20–29 (2005)

An Heuristic Approach for the Training Dataset Selection in Fingerprint Classification Tasks

Giuseppe Vitello[1], Vincenzo Conti[1], Salvatore Vitabile[2], and Filippo Sorbello[3]

[1] Faculty of Engineering and Architecture
University of Enna Kore, Enna, Italy
`{giuseppe.vitello,vincenzo.conti}@unikore.it`
[2] Department of Biopathology, Medical and Forensic Biotechnologies
University of Palermo, Palermo, Italy
`salvatore.vitabile@unipa.it`
[3] Department of Chemical, Management, Computer and Mechanic Engineering
University of Palermo, Palermo, Italy
`filippo.sorbello@unipa.it`

Abstract. Fingerprint classification is a key issue in automatic fingerprint identification systems. It aims to reduce the item search time within the fingerprint database without affecting the accuracy rate. In this paper an heuristic approach using only the directional image information for the training dataset selection in fingerprint classification tasks is described. The method combines a Fuzzy C-Means clustering method and a Naive Bayes Classifier and it is composed of three modules: the first module builds the working datasets, the second module extracts the training images dataset and, finally, the third module classifies fingerprint images in four classes. Unlike literature approaches using a lot of training examples, the proposed approach requires only 18 directional images per class. Experimental results, conducted on a consistent subset of the free downloadable PolyU database, show a classification rate of 87.59%.

Keywords: Fingerprint Classification, Directional Images, Fuzzy C-Means, Naive Bayes Classifier, Training Dataset Optimization.

1   Introduction

The new emerging market of mobile users rapidly grows, influencing several scenarios such as commercial, banking, and government applications. As result, secure access system design [1] and high response speed are currently the main issues. In this field, fingerprint authentication and classification systems represent a valid solution. The identification process performed in a database divided in classes is fast, since the number of the needed comparisons can be reduced by searching the fingerprint only in the same class of the database [2]. Many fingerprint classification approaches are reported in literature based on macro features [3] [4], structural information [5], neural networks [6] [7] [8], fuzzy-neural networks [9], probabilistic model [10] [11], and so on. Unfortunately, macro features are not always present in fingerprint images,

for example due to image acquisitions not correctly performed such as in partial fingerprint [12]; in this case, may be useful the approach proposed in [13] using pseudo-singularity points.

In this paper, an heuristic approach to optimize the training phase in fingerprint classification tasks is proposed. It is inspired by the works described in [14] and [15], and it uses only the directional image information [3]. Every element in the directional image represents the local orientation of the fingerprint ridges in the original gray-scale image (Figure 1). With more details, the proposed approach combines a Fuzzy C-Means clustering method and a Naive Bayes Classifier, and it is composed of three modules. The first module, Datasets Building, builds the working datasets; the second module, Training Dataset Extraction, extracts the dataset of training images; the third module, Fingerprint Classification, classifies a fingerprint image into one of the following NIST standard classes: Left Loop, Right Loop, Tented Arch and Whorl (Plain Loop and Central Pocket Loop) [16]. Unlike literature approaches using a lot of training examples, the proposed one requires only the use of 18 directional images per class. Experimental results, conducted on a consistent subset of the free PolyU (Hong Kong Polytechnic University) database [17], show a classification rate of 87.59%.
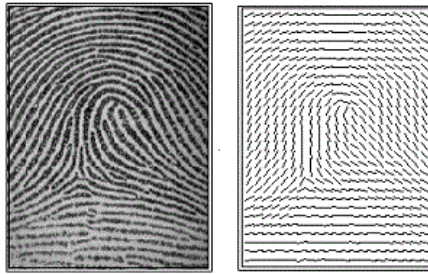
Fig. 1. Example of a fingerprint image with the related directional image

The paper is structured as follow. Section 2 reports the main literature works on fingerprint image classification. Section 3 describes the proposed approach. Section 4 outlines the experimental results. Finally, conclusions and future works are reported in section 5.

2   Related Works

Many classification approaches have been proposed in literature and many researchers are still working in this field. Below, some of these works are briefly reported. Ballan et al. in [3] reduce image distortion and contrast, before computing the fingerprint directional image on NIST database [16]. Successively, they extract singular points and classify the fingerprint using topological and numerical considerations about these points. Maio and Maltoni in [5] compute a relational graph, summarizing the fingerprint macro-structure, from the segmentation of the directional image. The obtained graph is compared with model graphs in order to classify the fingerprint. They don't say which database they used. Patil and Suralkar in [6] use a neural

network as decision stage. The network is ready to perform matching process and is successfully developed to identify and classify the fingerprint using back propagation algorithm. The presented methodology has been validated on a standard database of 800 images (they don't say the database name) classified into six classes obtaining a classification rate of 80.2%. Mohamed and Nyongesa in [9] present a classification scheme based on the encoding of singular points (Core and Delta) together with their relative positions and directions. The image analysis is carried out in four stages: segmentation, directional image estimation, singular point extraction and feature encoding. Successively, a fuzzy-neural network is used to implement the classification of input feature codes obtaining an average classification accuracy of 98.5% on NIST-4 database. Jung and Lee in [10] use a probabilistic approach (Markov model) based on the ridge characteristics of fingerprint classes on FVC2000 DB1 [18] and FVC2002 DB1 [19] databases, while Senior in [20] describes a novel method of classification based on hidden Markov models and decision trees to recognize the ridge structure on NIST-4 database. Pattichis et al. in [21] focus their work primarily on the image and feature enhancement and on finding improved classifiers, and less on the development of novel fingerprint representations. Using an AM–FM (Amplitude Modulated-Frequency Modulated) representation for each fingerprint of NIST database, they obtain significant gains in classification performance.

3   The Proposed Approach

The proposed heuristic approach, inspired by the works described in [14] and [15], is an efficient and effective method to optimize the training phase in fingerprint classification tasks, using only the directional image information. It combines a Fuzzy C-Means clustering method and a Naive Bayes Classifier, and it is composed of three modules: Datasets Building, Training Dataset Extraction and Fingerprint Classification (Figure 2). Unlike literature approaches using a lot of training examples (e.g., in [7] authors use 30 images per class, while in [10] half of each class in the whole database), the proposed one requires only the use of 18 directional images per class.

The following subsections describe the three modules.
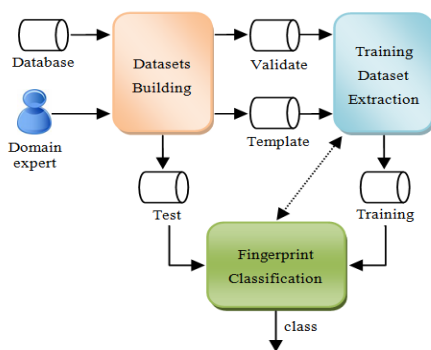
Fig. 2. The architecture of the proposed system

3.1   Datasets Building Module

This module plays an important role because it requires the contribution of a domain expert to choose the Template dataset (from which extracting the best training set). It is composed of three sub-modules following described (Figure 3).
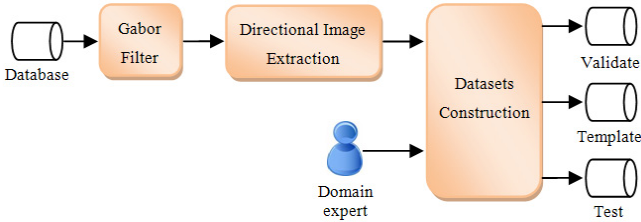
Fig. 3. The proposed Datasets Building Module

Gabor Filter Sub-module
It is applied to all images of the used database to enhance their quality [22].

Directional Image Extraction Sub-module
It extracts the directional image in three steps: extraction of the direction for each pixel; processing of the previous step output assembling the pixels in 8x8 blocks; computing of the predominant direction for each block (in every 8x8 block, the direction with greater frequency is attributed to the considered block). In this work, 8 directions have been chosen, from 0° to 180° as shown in Figure 4, and codified as a number in [0, 7].
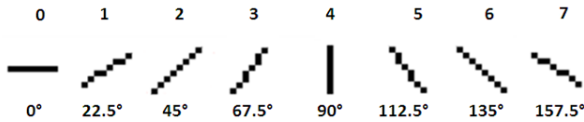
Fig. 4. The 8 directions used to build the directional image

In order to extract the direction $D(i,j)$ of the point (i,j) a vector v is calculated by the following equation (1):

$$v[k] = \sum_{d=1}^{q}[C(i_d, j_d) - C(i,j)] \qquad with\ k = 0..7 \qquad (1)$$

where $C(i,j)$ and $C(i_d, j_d)$ indicate the grey level of points (i,j) and (i_d, j_d), respectively, while q=16 is the number of selected pixels along a considered direction. The direction is finally obtained as the position of the minimum value in the vector v. However, acquisitions not correctly performed can affect the calculation of predominant directions inside spoiled zones; therefore, a smoothing algorithm is applied. This is achieved by calculating the directional histogram, comparing the directions in areas

of 3x3 blocks: the direction of the central block is replaced by the higher frequency direction of the neighboring blocks.

Datasets Construction Sub-module
It builds three different datasets: the 150 images Template dataset requires a domain expert, since it is hand selected; the 100 images Validate dataset is randomly selected, following the common distribution in nature of fingerprint classes, and then it is hand divided into the considered four classes by the domain expert; the Test dataset consists of the remaining images of the original database.

3.2   Training Dataset Extraction Module

It is composed of two sub-modules (Figure 5), following described, and works in cooperative way with the Fingerprint Classification Module.
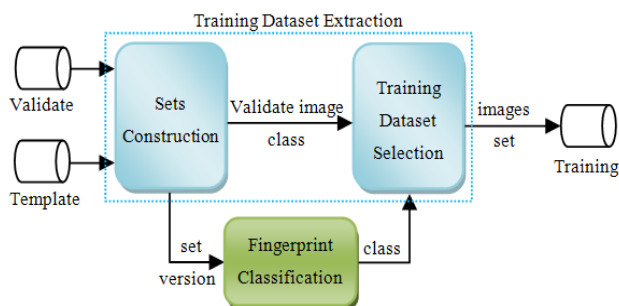
Fig. 5. The proposed Training Dataset Extraction Module

Sets Construction Sub-module
From the four clusters obtained applying the Fuzzy C-Means clustering method to the Template dataset, it builds 250 collections of 18 randomly selected images per cluster: 12 images near the cluster center and 6 images near the boundary. With more details, every boundary is identified calculating the Euclidean semi-distance among each cluster centers pair. Successively, for each collection, it builds 200 different sets, each of one composed of 3 groups of 6 randomly selected images per cluster (Figure 6). Finally, for each set, it creates 100 different set versions, adding one Validate image per group.

Training Dataset Selection Sub-module
It stores the accuracy rate of each set. Successively, it selects the one with the highest value over a threshold, experimentally fixed to 80%. The threshold is used to fix the number of items of the Validate sets and collections.

3.3   Fingerprint Classification Module

It is composed of five sub-modules (Figure 7), following described, and works in cooperative way with the Fingerprint Classification Module only for the training dataset extraction.
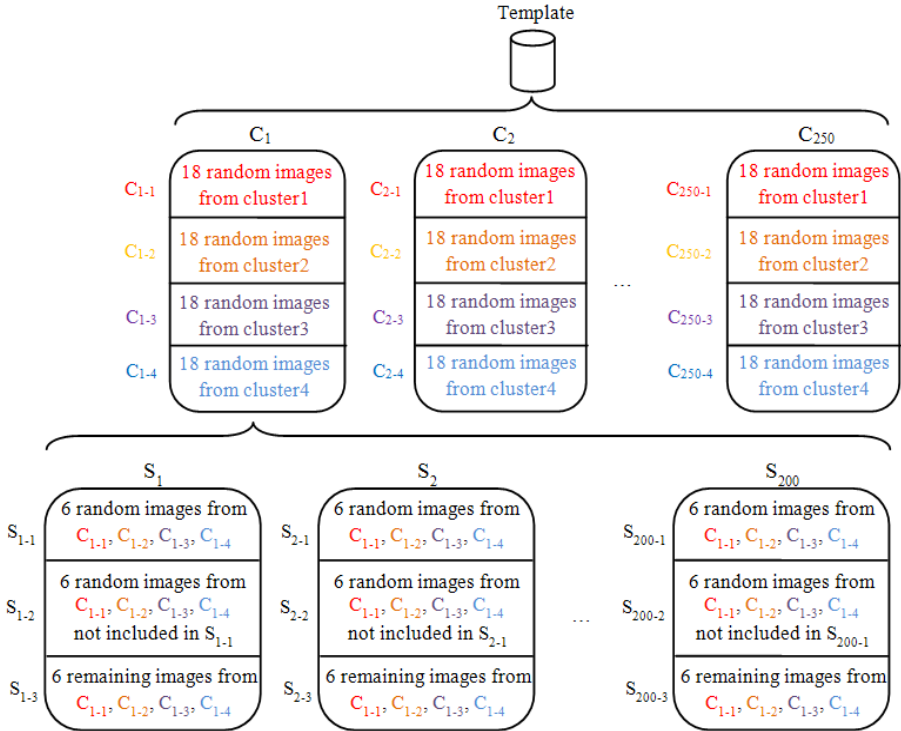
Fig. 6. The proposed set construction approach (the different colors represent the 4 clusters)
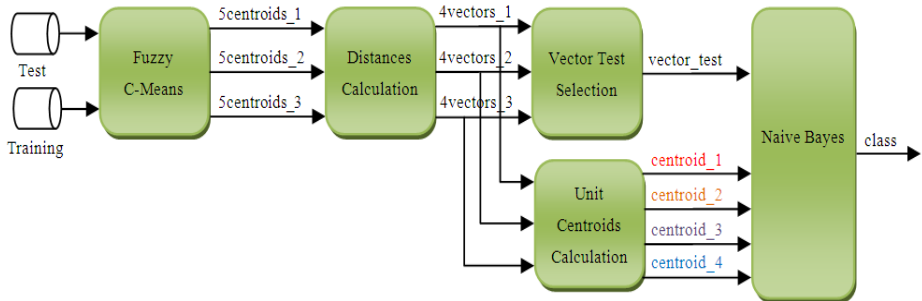
Fig. 7. The proposed Fingerprint Classification Module

Fuzzy C-Means Sub-module

It is composed of three components, each processing one group. Each component calculates five centroids: one centroid for the Test image and four centroids for the Training images (applying the average function on the elements of the same cluster).

Fuzzy C-Means [23] is an iterative algorithm and its purpose is to find cluster centers to minimize the objective function described by the formula (2):

$$J_{FCM} = \sum_{i=1}^{C} \sum_{k=1}^{N} u_{ik}^{p} \|x_k - v_i\|^2 \tag{2}$$

where $p = [1, \infty)$, the constant that determines the fuzziness degree of the classification process, has been experimentally fixed to 6. The algorithm stop condition is described by the relation (3), where ϵ has been experimentally fixed to 0.1.

$$|J_{FCM}[i^{th} \ iteration] - J_{FCM}[(i-1)^{th} \ iteration]| < \epsilon \tag{3}$$

Distances Calculation Sub-module

It calculates for each group the distances, element by element, between the centroid of the Test image and the four centroids of the Training images. By comparing such distances with a threshold, experimentally fixed to 0.1, it creates 4 binary vectors, for each group (Figure 8).
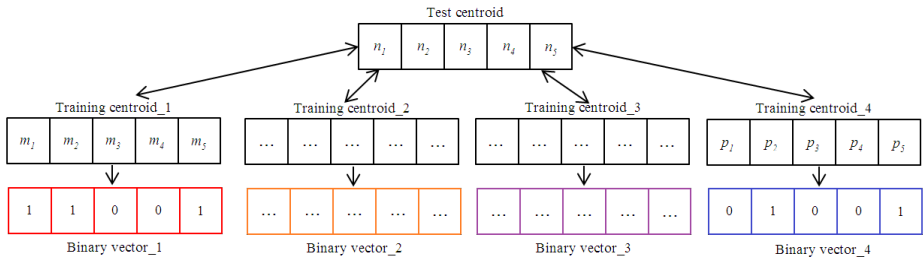
Fig. 8. The cells of each black vector (Test centroid and 4 Training centroid vectors) contain the centroid coordinates; the other 4 vectors represent the Binary vectors

Vector Test Selection Sub-module

It analyzes the 12 binary vectors and identifies the best one representing the Test image. It chooses the first among those vectors containing more 1 than 0 values.

Unit Centroids Calculation Sub-module

It reorganizes the 12 binary vectors in 4 units, so that each unit is composed of the 3 vectors of the same cluster. Successively, it calculates 4 centroids, one per unit, computing the average of the respective elements (Figure 9).
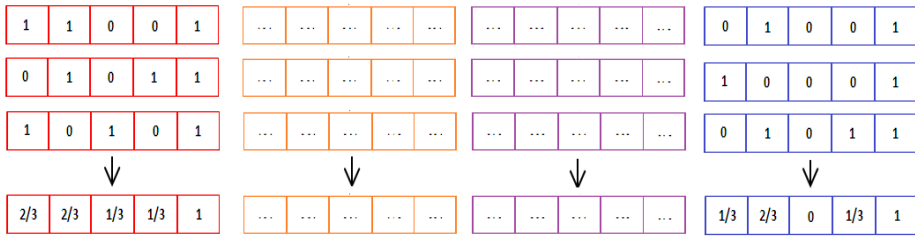
Fig. 9. The 4 unit centroids obtained applying the average function to the 3 vectors of the same cluster (a different color is used for each cluster)

Naive Bayes Sub-module

It classifies the Test image using the 5 vectors, obtained by the two previous sub-modules. A Naive Bayes Classifier is a simple probabilistic classifier based on the Bayes theorem with strong independence assumptions [14]. It assumes that the domain variables are independent, given the class, and each variable has a finite number of values. Usually, the model parameters (e.g., prior class probabilities and feature probability distributions) are approximated with the relative frequencies from the training database. In the proposed work, the class probabilities of the used NIST database have been experimentally fixed as in Table 1, following the natural distribution of the fingerprint classes.

Table 1. Class probabilities in the used NIST database

NIST class	Value
Tented Arch	0.07
Left Loop	0.20
Right Loop	0.25
Plain Loop/Central Pocket Loop	0.48

4 Experimental Results

To test the effectiveness of the proposed approach the free downloadable database PolyU [17] has been used. It contains 1480 fingerprint images belonging to the following NIST classes: Left Loop, Right Loop, Arches (Plain and Tented) and Whorl (Plain Loop, Central Pocket Loop, Accidental Loop and Double Loop) [16]. However, in the proposed work, a consistent PolyU subset of 1185 images, containing the Left Loop, Right Loop, Tented Arch, Plain Loop and Central Pocket Loop images, has been used. Since in the literature no classification systems has been tested using the PolyU database, we have performed a comparison (reported in Table 2) between the proposed approach and a standard Multilayer Perceptron (MLP) approach (50 hidden neurons), in terms of classification rate. The used Training and Test set sizes are described in Table 3.

Table 2. The proposed approach vs the standard MLP based approach

Class	Proposed Approach		MLP Approach 50% split		MLP Approach 70% split	
	Test dataset (correct/total)	Classification rate	Test dataset (correct/total)	Classification rate	Test dataset (correct/total)	Classification rate
Tented Arch	27/34	79.41%	11/25	44.00%	4/5	80.00%
Left Loop	142/160	88.75%	70/147	47.62%	87/97	89.69%
Right Loop	243/282	86.17%	82/193	42.49%	92/105	87.62%
Plain Loop/ Central Pocket Loop	407/459	88.67%	98/227	43.17%	136/148	91.89%
TOTAL	819/935	87.59%	261/592	44.09%	319/355	89.86%

Table 3. The used Training and Test set sizes

	Proposed Approach	MLP Approach 50% split	MLP Approach 70% split
Training set	250 images (21%) [1]	593 images	830 images
Test set	935 images (79%)	592 images	355 images

5   Conclusions and Future Works

In this paper an heuristic approach to optimize the training phase in fingerprint classification tasks is described. The approach combines the classification properties of a Fuzzy C-Means clustering method and a Naive Bayes Classifier on directional image information. Unlike literature approaches using a lot of training examples, the proposed approach requires only 18 directional images per class. Experimental results, conducted on a consistent subset of the free PolyU database, show a classification rate of 87.59%. Similar results can be obtained with a neural network based approach if and only if a very large training set is used.

Future works will be aimed to the study and implementation of an automatic technique for the training dataset extraction as well as to allow system employment in a different application domain.

References

1. Conti, V., Vitabile, S., Vitello, G., Sorbello, F.: An embedded biometric sensor for ubiquitous authentication. In: Proc. of AEIT Annual IEEE Conference, pp. 1–6 (2013), doi:10.1109/AEIT.2013.6666815, ISBN: 978-8-8872-3734-4
2. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2009) ISBN: 978-1-84882-253-5

---

[1] 150 and 100 images are used as Template dataset and Validate dataset, respectively.

3. Ballan, M., Sakarya, F.A., Evans, B.L.: A Fingerprint Classification Technique Using Directional Images. In: Proc. of the 31st Asilomar Conference on Signals, Systems and Computers, vol. 1, pp. 101–104 (1997), doi:10.1109/ACSSC.1997.680037, ISSN: 1058-6393

4. Awad, A.I., Baba, K.: Singular Point Detection for Efficient Fingerprint Classification. International Journal on New Computer Architectures and their Applications (IJNCAA) 2(1), 1–7 (2012) ISSN: 2220-9085

5. Maio, D., Maltoni, D.: A Structural Approach to Fingerprint Classification. In: Proc. of the 13th International Conference on Pattern Recognition (ICPR), vol. 3, pp. 1051–4651 (1996), doi:10.1109/ICPR.1996.547013, ISSN: 1051-4651

6. Patil, S.R., Suralkar, S.R.: Fingerprint Classification using Artificial Neural Network. International Journal of Emerging Technology and Advanced Engineering 2(10), 513–517 (2012) ISSN: 2250-2459

7. Conti, V., Militello, C., Vitabile, S., Sorbello, F.: An Embedded Fingerprints Classification System based on Weightless Neural Networks. In: I O S Press (ed.) Frontiers in Artificial Intelligence and Applications. New Directions in Neural Networks, vol. 193, pp. 67–75 (2009), doi:10.3233/978-1-58603-984-4-67, ISSN: 0922-6389

8. Kamijo, M.: Classifying Fingerprint Images using Neural Network: Deriving the Classification State. In: Proc. of the IEEE International Conference on Neural Networks, vol. 3, pp. 1932–1937 (1993), doi:10.1109/ICNN.1993.298852 ISBN: 0-7803-0999-5

9. Mohamed, S.M., Nyongesa, H.O.: Automatic Fingerprint Classification System Using Fuzzy Neural Techniques. In: Proc. of the IEEE International Conference on Fuzzy Systems, vol. 1, pp. 358–362 (2002), doi:10.1109/FUZZ.2002.1005016, ISBN: 0-7803-7280-8

10. Jung, H.-W., Lee, J.-H.: Fingerprint Classification Using the Stochastic Approach of Ridge Direction Information. In: Proc. of the IEEE International Conference on Fuzzy Systems, pp. 169–174 (2009), doi:10.1109/FUZZY.2009.5277309, ISSN: 1098-7584

11. Qing, S., Xisheng, L., Hui, Y., Chen, Q.: Naive Bayes Classifier applied in Droplet Fingerprint Recognition. In: Proc. of the 3rd Global Congress on Intelligent Systems (GCIS), pp. 152–155 (2012), doi:10.1109/GCIS.2012.68, ISBN: 978-1-4673-3072-5

12. Conti, V., Vitello, G., Sorbello, F., Vitabile, S.: An Advanced Technique for User Identification using Partial Fingerprint. In: Proc. of the 7th International IEEE Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 236–242 (2013), doi:10.1109/CISIS.2013.46, ISBN: 978-0-7695-4992-7

13. Conti, V., Militello, C., Vitabile, S., Sorbello, F.: Introducing Pseudo-Singularity Points for Efficient Fingerprints Classification and Recognition. In: Proc. of the 4th IEEE International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 368–375 (2010), doi:10.1109/CISIS.2010.134, ISBN: 978-0-7695-3967-6

14. Tang, Y., Pan, W., Li, H., Xu, Y.: Fuzzy Naive Bayes Classifier Based on Fuzzy Clustering. In: Proc. of IEEE International Conference on Systems, Man and Cybernetics, vol. 5 (2002), doi:10.1109/ICSMC.2002.1176401, ISSN: 1062-922X

15. Vitello, G., Conti, V., Migliore, G.I.M., Sorbello, F., Vitabile, S.: A Novel Technique for Fingerprint Classification based on Fuzzy C-Means and Naive Bayes Classifier. In: Proc. of the 8th International IEEE Conference on Complex, Intelligent and Software Intensive Systems (CISIS), pp. 155–161 (2014), doi:10.1109/CISIS.2014.23, ISBN: 978-1-4799-4325-8

16. National Institute of Standards and Technology, `http://www.nist.gov`

17. PolyU database,
`http://www.comp.polyu.edu.hk/~biometrics/HRF/HRF.htm`

18. FVC Databases, `http://bias.csr.unibo.it/fvc2000/databases.asp`
19. FVC Databases, `http://bias.csr.unibo.it/fvc2002/databases.asp`
20. Senior, A.: A Combination Fingerprint Classifier. IEEE Transaction on Pattern Analysis and Machine Intelligence 23(10), 1165–1174 (2001), doi:10.1109/34.954606, ISSN: 0162-8828
21. Pattichis, M.S., Panayi, G., Bovik, A.C., Hsu, S.-P.: Fingerprint Classification Using an AM–FM Model. IEEE Transactions on Image Processing 10(6), 951–954 (2001), doi:10.1109/83.923291, ISSN: 1057-7149
22. Batra, D., Singhal, G., Chaudhury, S.: Gabor filter based fingerprint classification using support vector machines. In: Proc. of the 1st IEEE India Annual Conference (INDICON), pp. 256–261 (2004), doi:10.1109/INDICO.2004.1497751, ISBN: 0-7803-8909-3
23. Zhang, J.-S., Leung, Y.-W.: Improved Possibilistic C-Means Clustering Algorithms. IEEE Transaction on Fuzzy Systems 12(2), 209–217 (2004), doi:10.1109/TFUZZ.2004.825079, ISSN: 1063-6706

Fuzzy Measures and Experts' Opinion Elicitation

An Application to the FEEM Sustainable Composite Indicator

Luca Farnia[1,2] and Silvio Giove[2,3]

[1] Mediterranean Center for Climate Change (CMCC), Bologna, Italy
[2] Fondazione Eni Enrico Mattei, Venice, Italy
[3] University Cà Foscari, Venice, Italy
luca.farnia@feem.it, sgiove@unive.it

Abstract. To over pass the limits inherent in liner models, suitable aggregation operators are required, taking into account interactions among the criteria. This becomes more and more crucial in Decision Theory, where all the information can be inferred by one or more Experts, using an *ad hoc* questionnaire. This is the case of the FEEM SI sustainability index, a composite geo-referenced index which aggregates several economic, social and environmental dimensions-structured in a decision tree- into a single number between zero (the worst sustainable country) and one (the best one). Fuzzy measure (*non additive measures*) are here proposed for the aggregation phase. To this purpose, each intermediate node of the structure combines the values of the sub-nodes using a model based on *second-order* non additive measure. To infer the value of the measure for each node, a suitable questionnaire has been fulfilled by a set of Experts, and the obtained answers were processed using an optimization algorithm. To guarantee the strict convexity of the algorithm, the questionnaire needs to be carefully designed. The individual measures are subsequently aggregated and the numerical results permitted to compare the sustainability of all the considered territorial units.

Keywords: Fuzzy measures, non-additive measures, Choquet integral, preference structure, sustainability, aggregation operators.

1   Introduction

In Multi Attribute problems the Weighted Averaging operator (WA) is widely used to aggregate the normalized values of the criteria. Anywise, the linear WA method is unable to include interactions (synergies or redundancies) among the criteria. For this reason, more specialized algorithms need to be considered. One of the most commonly used is based on fuzzy measures (or capacity, or non-additive measures, NAMs for brevity), which assigns a weight to every possible coalition of criteria, and not to a singleton only. A suitable aggregation operator, the Choquet integral [Beliakov, 2009], extends the Weighted Averaging (WA) to the computation of an aggregated results using NAM. The interested reader can refer to [Grabisch, 2000], [Marichal, 2000-2] for a methodological analysis of fuzzy measures. In the case of WA only n

parameters need to be elicited, if n is the number of the criteria to be aggregated. But in NAM the number of the required parameters – the value of the fuzzy measures - exponentially increases with n, and the numerical complexity becomes crucial, especially in the case of Decision Theory, where the information can be inferred only by one or more Experts. The *reduced order* model, which considers interactions only between subset of limited cardinality, can be applied to reduce the numerical complexity. In particular, the second order model admits interactions only between couple of criteria. In this paper we propose an elicitation method based on a suitable questionnaire. The questionnaire is formed by a set of alternatives, i.e. *what ... if...* questions, and the answers can be used to elicit the non-additive measure by means of the *Least Square* optimization algorithm that minimizes the sum of squared distances between the answers of the Expert(s) and the solutions of the problem. The procedure was applied to a real world case study, the project FEEM SI, a composite Sustainability Index including 23 indicators grouped in a tree structure. A team of Experts were asked to fulfil a questionnaire, by which the NAMs were implicitly elicited, one for each node of the tree. To avoid a too strong mental effort, the number of the questions needs to be limited, then a second order model was selected as aggregation operator. The paper is organized as follows. Section 2 introduces the concept of fuzzy measures, included the reduced order model, and the Choquet integral as aggregation operator. Section 3 describes the proposed elicitation approach, while Section 4 reports its application to the FEEM SI sustainability indicator. Some conclusive comments and suggestions are reported in Section 5; technical results are briefly reported in the Appendix.

2   Fuzzy Measures and the Choquet Integral

A fuzzy measures or NAM, defined over the set of criteria $N = \{1, 2, ..., n\}$, is a set function $\mu: 2^N \to [0,1]$ satisfying the following boundary and monotonicity conditions:

$$\begin{cases} \mu(\emptyset) = 0 \\ \mu(N) = 1 \\ S \subseteq T \subseteq N \ \Rightarrow \ \mu(S) \leq \mu(T) \leq 1 \ \forall \, S, T \subseteq N \end{cases} \tag{1}$$

A NAM assigns to every subset (coalition) of criteria a measure that is not necessarily the sum of the measures of their singletons. Namely, if the measure of a coalition is greater (smaller) than the sum of the measures of their singletons, that measure represents a synergic (redundant) interaction among the criteria belonging to the coalition. Instead when the measure of a coalition equals the sum of the measures of the singletons belonging to it, the NAM collapses to the linear aggregation (WA) and no interaction exists among the criteria. Given a NAM μ, its Möbius representation is the following set function, see [Marichal, 2000-2]:

$$m(S) = \sum_{T \subseteq S} (-1)^{s-t} \mu(T) \, , \forall \, S, T \subseteq N \tag{2}$$

where $s = card(S)$, $t = card(T)$. Moreover the following boundary and the monotonicity conditions are required, see [Marichal, 2000-2]:

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{T \subseteq N} m(T) = 1 \\ \sum_{\substack{T \subseteq S \\ T \ni i}} m(T) \geq 0, \ \forall S \subseteq N, \ \forall i \in S \end{cases} \tag{3}$$

Let μ be a NAM defined on N and $X = \{x_1, x_2, \dots, x_n\}$ the normalized values of the criteria belonging to N; the (discrete) Choquet integral with respect to μ is given by [Grabisch, 2000]:

$$C_\mu(x_1, \dots, x_n) = \sum_{i=1}^n (x_{(i)} - x_{(i-1)}) \mu(A_{(i)}) \tag{4}$$

where (i) means that the indices have been permutated in such a way that $x_{(1)} \leq \cdots \leq x_{(n)}$, while $A_{(i)} = \{x_{(1)}, \dots, x_{(n)}\}$ and $x_{(0)} = 0$. Using the Möbius representation the Choquet integral can be written as:

$$C_m(x_1, \dots, x_n) = \sum_{T \subseteq N} m(T) \bigwedge_{i \in T} x_i \tag{5}$$

being $\wedge$ the minimum operator. To define a capacity μ on N, 2^{N-1} parameters are required. If the capacity represents the preference structure of an Expert, it needs to be directly or implicitly elicited using a suitable questionnaire. In the case of a complete model, the number of parameters exponentially increases with n, usually a prohibitive task, thus a reduced order model is proposed [Grabisch, 1997], satisfying the compromise between criteria interaction and numerical complexity. A capacity μ on N is said to be k-additive if its Möbius representation satisfies $m(T) = 0 \ \forall T \subseteq N$ such that $t > k$, and there exists at least one subset T with $card(T) = k$ such that $m(T) \neq 0$. In this way a k-additive capacity with $k \in \{1, 2, \dots, n\}$ is completely defined by the identification of $\sum_{j=1}^k \binom{n}{j}$ parameters. If $k = 2$ we have a *second order* model (2-order model for brevity) and only $\frac{n(n+1)}{2}$ parameters are required.

2.1   Behavioural Analysis

We limit to mention the two most popular indices developed in order to have a direct interpretation of non additive measures: the *Shapley* value [Shapley, 1953] and the *Interaction* index [Grabisch, 1997].

The Shapley value is a measure of the relative importance of a criterion, computed averaging all the marginal gains between any coalition not including the criterion, and the one which includes it. In terms of Möbius representation the Shapley value of criterion i is defined as following:

$$\varphi(m; i) = \sum_{T \ni i} \frac{1}{t} m(T) \ \forall \ T \subseteq N \tag{6}$$

When fuzzy measures are not additive, interaction among criteria exists; in terms of Möbius representation the Interaction index among a combination S of criteria, can be formulated as:

$$I_S(m; S) = \sum_{T \supseteq S} \frac{1}{t-s+1} m(T) \quad \forall \ S \subseteq N \tag{7}$$

In a 2-order model the interaction index of a coalition coincides with its Möbius representation.

3   The Least Square Elicitation Approach

The specialized literature reports several approaches to elicit fuzzy measures. Among them, we limit to recall the Least Square (LS) and the Heuristic Least Square (HLS) [Grabisch, 1995], the approach of Marichal and Roubens (MR) [Marichal, 2000-1], the Minimum Variance (MV) [Kojadinovic, 2007] and Minimum Distance (MD) [Kojadinovic, 2000]. Such methods differ together by the required information (for instance, cardinal or ordinal type), the objective function, and the applied algorithm. The LS and the HLS require cardinal information about the alternatives, that is, a global evaluation (sometimes defined as *utility*), assigned by an Expert. For this reason, we define these methods as *cardinal-based* models. Conversely, the MR, MV and MD method require only a preference order of the alternatives; they can be defined *ordinal-based* models. In this contribution we applied the LS, after having designed carefully the questionnaire, in such a way that the optimization problem is strictly convex and thus the LS can return a unique solution[1].

To this purpose, let be:

1. $B = \{c_1, c_2, \dots, c_n\}$ the set of the criteria;

2. $A = \{a_1, a_2, \dots, a_v\}$ a scenario, that is a set of v alternatives;

3. $\mathbf{X} = \begin{pmatrix} x_1(1), \dots, x_n(1) \\ \dots \dots \dots \dots \dots \dots \\ x_1(v), \dots, x_n(v) \end{pmatrix}$ the normalized values of the criteria for each alternative

    (or the criteria utility for each alternative);

4. $y(a)$: the utility assigned by the Expert to alternative a

5. $C_m(j)$, $j = 1,,. v$: the value of the j-th alternative computed through the 2-order

    model, with the elicited parameters, $C_m(a) \equiv C_m(x_1(a), \dots, x_1(a))$;

We now briefly formalize the LS optimization algorithm which minimizes the sum squared differences between the evaluation of the Expert and the ones computed by the *k*-order model.

---

[1] If the optimization problem is not strictly convex, the solution is not unique [P. Miranda, M. Grabisch, 1999]. Nerveless the strict convexity depends on the questionnaire structure, see Subsection 3.1.

LS: $MIN_{m(T)} \sum_{a=1}^{v} \left(C_m(a) - y(a) \right)^2$

$s.t.$

$$\left. \begin{array}{l} \sum_{\substack{T \subseteq S \\ T \ni i}} m(T) \geq 0 \ \forall S \subseteq N, \forall i \in S \\ \sum_{T \subseteq N} m(T) = 1 \end{array} \right\} \quad Boundary \ and \ monotonicity \ conditions$$ (8)

If an algebraic condition is satisfied about the scenario's values proposed in the questionnaire -see Subsection 3.1- the solution to problem (8) is unique.

3.1   Optimization Issue

We underline that the LS approach returns a unique vector of Möbius representations *iff* the number of alternatives to be judged and the utilities associated to the criteria satisfy the mathematical condition stated in the Property below.

The property needs the following preliminary definition: let define the *full* scenario matrix **A** whose elements in the first v rows are partitioned by the matrix **X** of utilities associated to each criteria for each alternative, and the matrix **Λ** containing the minimum utilities values of each coalition and for each alternative; all the elements in the last row are equal to one (representing the boundary condition):

$$\mathbf{A} = \begin{bmatrix} \mathbf{X} & \mathbf{\Lambda} \\ \mathbf{1'} & \mathbf{1'} \end{bmatrix} = \begin{bmatrix} x_1(1) & \dots & x_n(1) & \bigwedge_{i \in T} x_i(1) \\ x_1(2) & \dots & x_n(2) & \bigwedge_{i \in T} x_i(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(v) & \dots & x_n(v) & \bigwedge_{i \in T} x_i(v) \\ 1 & \dots & 1 & \mathbf{1'} \end{bmatrix}$$ (9)

with $card(T) \leq k$.

Property 1

The LS approach returns a unique vector of Möbius representations *iff* the questionnaire is designed in such a way that the *full* scenario matrix has rank equal to the number of Möbius representations to be elicited in a *k*-order model.

Consider for instance a *2*-additive model with two criteria and two alternatives; following the formulation in Section 3, the *full* scenario matrix has the following structure:

$$\mathbf{A} = \begin{bmatrix} x_1(1) & x_2(1) & Min(x_1(1), x_2(1)) \\ x_1(2) & x_2(2) & Min(x_1(2), x_2(2)) \\ 1 & 1 & 1 \end{bmatrix}$$

A   necessary   and   sufficient   condition   to   have $rank(\mathbf{A}) = 3$ is $x_i(1) > x_j(1) \ and \ x_i(2) < x_j(2)$.

4   The FEEM SI Composite Index

The project FEEM Sustainability Index (FEEM SI) started some years ago [FEEM Sustainability Index Methodological Report, 2011] with the aim to construct a global sustainability index on a national scale. The data generating model is based on a Computable General Equilibrium (CGE) dynamic economic model which is able to forecast macro-economic variables of interest over a time window of about 20 years. The CGE furnishes for each nation and for each year inside the time window, the forecasted values of the macro variable. We do not discuss here the CGE methodology, see the quoted references for a detailed explanation, but focus the attention on the aggregation methodology that, year by year and for every nation, computes a single Sustainability Index for any considered nation in the world. Doing so, it permits an immediate ranking and comparison of nations putting in evidence possible variation of nation sustainability, strength and weakness points. Figure 1 shows the FEEM SI composite index, which is split into the three main pillars of sustainability (*economic*, *social* and *environmental*). Each pillar is split again into sub-dimension, and so on until the 23 leaves that are the sampled indicators. For each node and sub-node, the aggregation follows a bottom-up procedure using the 2-order model, from the leaves up to the root, the FEEM SI composite index.

4.1   Methodology

FEEM SI is based on two main and complementary blocks of operations: the CGE model, which processes the macro variables used as input data, and the criteria weighting scheme which aggregates the normalized data for each block of criteria used. Data input have been normalized by means of suitable functions that transform the input criteria data on a scale between zero and one, given suitable thresholds imposed by Experts specialized in that field. This is hence a data-insensitive normalization approach, necessary to neutralize the risk of potential *rank reversal* problem, as can appear, for example, in the *min-max* normalization approach.

The criteria weighting scheme is based on Experts' opinion elicitation by means of an *ad-hoc* questionnaire satisfying the solution uniqueness condition as explained in Subsection 3.1. Subsections below explain how the questionnaire has been developed, together with the methodology used to aggregate Experts' preferences into a single "representative" one.

4.1.1   Ad-Hoc Questionnaire for Fuzzy Experts 'Opinion Elicitation'

Each Expert has been asked to evaluate some hypothetical nations on the base of the joint performance of some criteria considered. Given the structure of the decision tree whose nodes are formed by different set of criteria, this process has been performed for all nodes. Table 1. shows the discrete qualitative scale used in the questionnaire to describe the criteria performance (first column) and the Expert choices to evaluate alternatives (second column); the third column describes the equivalent numerical scale used in the elicitation algorithm.

Table 1. Evaluation scheme

Qualitative Scale		Numerical Scale
Criteria Performance	*Expert Evaluation*	
Very bad	Very Dissatisfied	0
Bad	Dissatisfied	0.25
Fair	Nor Diss./ Sat.	0.5
Good	Satisfied	0.75
Excellent	Very Satisfied	1

The criteria performance of each alternative has been set according to **Property 1** (see Subsection 3.1). Table 2 is an example of the FEEM SI main node, where 5 hypothetical nations with different performances in the economic, social and environmental dimension have to been evaluated by each interviewed Expert.

Table 2. FEEM SI (main node) questionnaire example

Nation	Criteria			Expert Overall Evaluation
	Economic	*Social*	*Environment*	
1	Excellent	Good	Bad	-
2	Excellent	Bad	Good	-
3	Good	Excellent	Bad	-
4	Bad	Excellent	Good	-
5	Bad	Good	Excellent	-

4.1.2   Experts' Opinion Elicitation and Their Aggregation

Given that NAM approach is sufficiently general to cover many preference structures, Expert's preference has been weighted according to his/her overall consistency in judging the alternatives proposed. This is indeed an important step, especially when a survey is conducted without having a direct and immediate control on Expert's evaluation. We measure Expert's consistency as a function of the sum of squared distances in problem 8), in such a way that the greater (smaller) this sum, the smaller (greater) the contribution from the relative Expert. The above conditions can be formalized as following. Given v alternatives to be judged, let define the vector $\boldsymbol{\varepsilon}_j$ $(1 \times v)$ whose elements represent the differences between the overall utilities values set by the j-th Expert and the respective Choquet values (solution of the problem 8)).

Let g_j be the sum of squared distances for the j-th Expert:

$$g_j = \boldsymbol{\varepsilon}_j' \boldsymbol{\varepsilon}_j. \tag{10}$$

The sum of squared distances for the j-th Expert is filtered using an exponential model:

$$h_j = e^{-\rho g_j} \quad \text{with } \rho > 0, \tag{11}$$

In such a way that $h_j \in (0, 1]$, being ρ a suitable positive constant.

The relative contribution from the *j*-th Expert (on a total of *D* interviewed) is defined as the following *importance* weight:

$$w_j = \frac{h_j}{\Sigma_{j=1}^{D} h_j}. \tag{12}$$

The final Möbius representation as the result of the weighted average of each Expert's preference, can be defined in the following:

$$m^*\{T\} = \Sigma_{j=1}^{D} w_j m_j\{T\} \quad \forall\, T \subseteq N \ and\ t \leq k \tag{13}$$

4.2   Some Results

We limit to show the results of the elicitation process in the main node of the FEEM Si index where the three pillars of sustainability have been jointly considered. Figure 2 illustrates the results of the weighting scheme with $\rho = 3$; from the left to the right are shown in decreasing order the weights to be associated to each of the 23 Expert interviewed in accordance to their overall coherence in this node. The data row of Figure 2 represent the numerical values of equations 10), 11) and 12) respectively. Figure 3 illustrates the Shapley values derived by the elicited Möbius representation for each Expert. Figure 4 illustrates the Interaction indices derived by the elicited Möbius representation. Table 3 shows the results of Experts' preferences aggregation in terms of Shapley values, from which the social dimension appears to be the most important pillar (38.6%), followed by the environmental pillar (35.70%), while the economic pillar is the least (25.70%). Table 4 shows the relative importance of the criteria belonging to the second level of the decision tree; *wellbeing* is considered the most important factor of sustainability and *GDP p.c.* the least one.

5   Conclusions

In this paper we proposed a multi-criteria approach based on the second order Choquet Integral, with the aim to take interaction among the criteria into account and, at the same time, to maintain the numerical complexity as low as possible. The fuzzy measures are elicited using an *ad hoc* questionnaire to be fulfilled by a panel of Experts, and a suitable optimization algorithm based on Least Square approach. In particular we showed that, when the questionnaire is structured in a particular way, the solution of the optimization problem is unique. The method was applied to the FEEM SI project, which is a computable geo-referenced sustainability index, organized into an hierarchical tree. For each node of the tree, the second order Choquet algorithm aggregates the values of the sub-node referring to it, bottom-up moving from the leaves up to the root. The aggregated sustainability index is computed for every considered territorial unit, permitting an immediate comparison. We are going to develop other optimization techniques based on alternative algorithms, like Goal Programming, testing the efficiency in comparison with the Least Square method.

References

1. Athanasoglou, S., Weziak-Bialowolska, D., Saisana, M.: Environmental Performance Index, – JRC Analysis and Recommendations, EPI-JRC, pp. 1–33 (2014)
2. Beliakov, G.: Construction of aggregation functions from data using linear programming. Fuzzy Sets and Systems 160, 65–75 (2009)
3. Cardin, M., Giove, S.: Approximation of fuzzy measures using second order measures: Estimation of andness bounds. In: Masulli, F. (ed.) WILF 2013. LNCS (LNAI), vol. 8256, pp. 150–160. Springer, Heidelberg (2013)
4. Decanq, K., Lugo, M.A.: Weights in environmental indices of wellbeing: an overview. Econometric Review 32(1), 7–34 (2013)
5. Despic, O., Simonovic, S.P.: Aggregation operators for decision making in water resources. Fuzzy Sets and Systems 115(1), 11–33 (2000)
6. FEEM Sustainability Index Methodological Report 2011, Fondazione Eni Enrico Mattei (2011), http://www.feemsi.org/documents/
7. Jones, D., Mehrdad, T.: Practical Goal Programming, vol. 141. Springer (2010)
8. Ishii, K., Sugeno, M.: A model of human evaluation process using fuzzy measure. International Journal of Man-Machine Studies 67, 242–257 (1996)
9. Grabisch, M., Nguyen, H.T., Walker, E.A.: Fundamentals of uncertainty calculi with applications to fuzzy inference. Kluwer Academic, Dordrecht (1995)
10. Grabisch, M.: A new algorithm for identifying fuzzy measures and its application to pattern recognition. In: Proceedings of international 4th IEEE Conference on Fuzzy Systems, Yokohama, pp. 145–150 (1995)
11. Grabisch, M.: k-order additive discrete fuzzy measures and their representation. Fuzzy Sets and Systems 92, 167–189 (1997)
12. Grabisch, M., Roubens, M.: Application of the Choquet integral in multicriteria deci-sion making. In: Grabisch, M., Murofushi, T., Sugeno, M. (eds.) Fuzzy Measures and Integrals – Theory and Applications, pp. 348–374. Physica Verlag (2000)
13. Ishii, K., Sugeno, M.: A model of human evaluation process using fuzzy measure. International Journal of Man-Machine Studies 67
14. Kojadinovic, I.: Quadratic distances for capacity and bi-capacity approximation and identification. A Quarterly Journal of Operations Research (in press), doi: 10.1007/s10288-006-0014-4; Marichal, J.-L., Roubens, M.: Determination of weights of interacting criteria from a reference set. European Journal of Operational Research 124, 641–650 (2000)
15. Kojadinovic, I.: Minimum variance capacity identification. European Journal of Operational Research 177, 498–514 (2007)
16. Marichal, J.: An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. The IEEE Transactions on Fuzzy Systems 8(6), 800–807 (2000)
17. Marichal, J.-L.: Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. European Journal of Operational Research 155, 771–791 (2004)
18. Marichal, J.-L.: K-intolerant capacities and Choquet integrals. European Journal of Operational Research 177, 1453–1468 (2007)
19. Meyer, P., Roubens, M.: Choice, ranking and sorting in fuzzy Multiple Criteria Decision Aid. In: Figueira, J., Greco, S., Ehrgott, M. (eds.) Multiple Criteria Decision Analysis: State of the Art Surveys, pp. 471–506. Springer, New York (2005)
20. Miranda, P., Grabisch, M.: Optimization issues for fuzzy measures. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 7(6), 545–560 (1999)

21. Mori, T., Murofushi, T.: An analysis of evaluation model using fuzzy measure and the Choquet integral. In: Proceedings of 5th Fuzzy System Symposium, Kobe, Japan, pp. 207–212 (1989) (in Japanese)
22. Murofushi, T.: A technique for reading fuzzy measures (I): the Shapley value with respect to a fuzzy measure. In: 2nd Fuzzy Workshop, Nagaoka, pp. 39–48 (1992) (in Japanese)
23. OECD/EC JRC, Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD, Paris (2008)
24. Pinar, M., Cruciani, C., Giove, S., Sostero, M.: Constructing the FEEM sustainability index: A Choquet integral application. Ecological Indicator 39, 189–202 (2014)
25. Shapley, L.S.: A value for n-person games. In: Kuhn, H.W., Tucker, A.W. (eds.) Contributions to the Theory of Games, Vol. II. Annals of Mathematics Studies, vol. 28, pp. 307–317. Princeton University Press, Princeton (1953)

Appendix
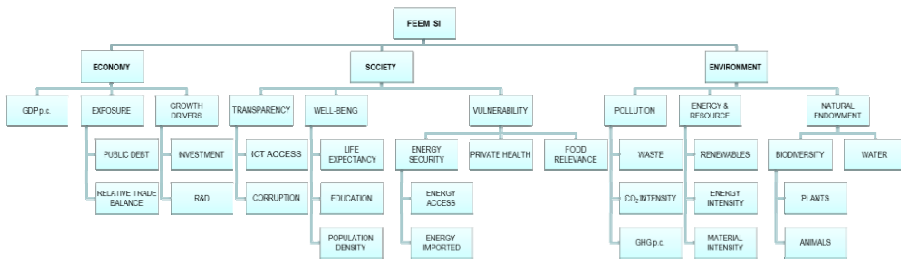
Fig. 1. FEEM SI decision tree

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
g(j)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.3	0.3	0.3	0.4	0.5	0.5	0.5	0.6	0.6
h(j)	1.0	1.0	1.0	1.0	0.9	0.9	0.9	0.8	0.7	0.7	0.6	0.6	0.5	0.5	0.4	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1
w(j)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
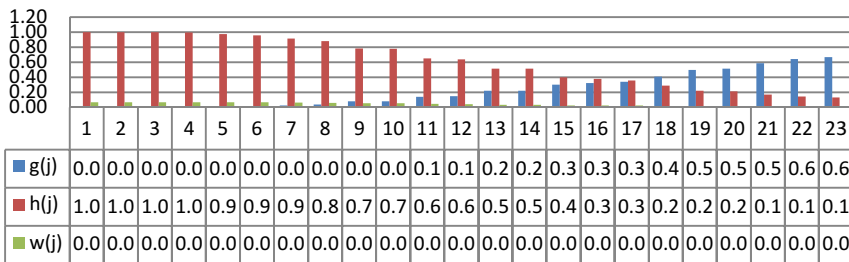
Fig. 2. Weighting scheme - FEEM SI node

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
■ Eco	0.2	0.3	0.3	0.1	0.1	0.2	0.2	0.3	0.2	0.2	0.3	0.0	0.0	0.2	0.3	0.3	0.3	0.4	0.0	0.2	0.3	0.3	0.3
■ Soc	0.1	0.1	0.4	0.4	0.5	0.4	0.3	0.3	0.3	0.4	0.3	0.5	0.5	0.3	0.3	0.4	0.3	0.3	0.6	0.5	0.3	0.3	0.3
■ Env	0.6	0.4	0.2	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.3	0.3	0.1	0.2	0.2	0.3	0.3	0.2	0.3	0.3
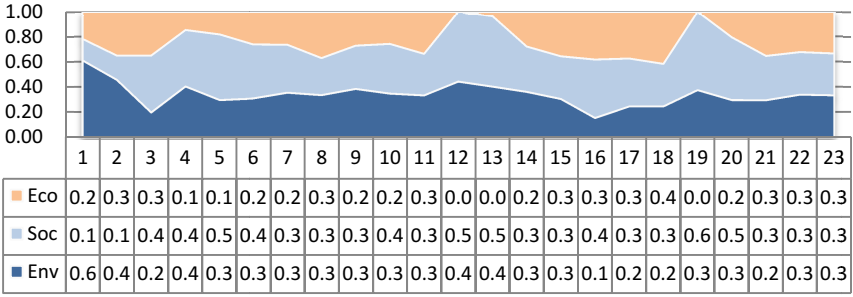
Fig. 3. Shapley values for each respondent - FEEM SI node

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
■ Env - Soc	0.2	0.2	-0.	0.4	0.5	0.1	0.6	-0.	0.2	0.6	0.3	0.2	0.4	0.4	-0.	0.0	0.2	0.4	0.7	0.1	0.2	0.3	0.3
■ Env - Eco	-0.	-0.	0.2	0.1	0.0	-0.	-0.	-0.	0.1	0.0	0.3	0.0	0.0	0.2	0.5	0.2	0.2	0.0	0.0	0.4	0.2	0.3	0.3
■ Soc - Eco	0.0	-0.	0.0	-0.	0.3	0.3	-0.	0.3	0.4	0.1	0.3	0.0	0.0	0.2	0.1	-0.	0.5	0.2	0.0	0.0	0.4	0.3	0.3
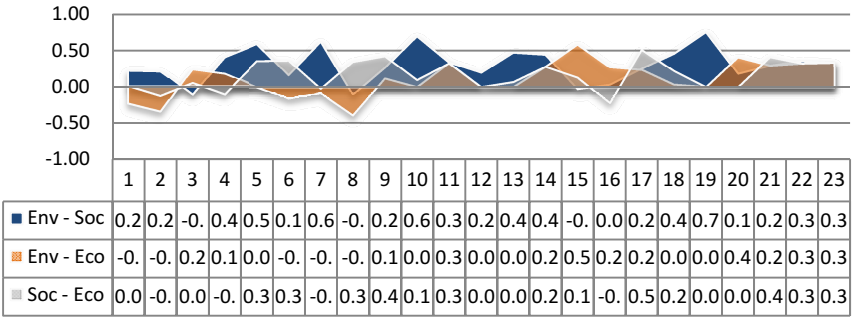
Fig. 4. Interaction indices for each respondent - FEEM SI node

Table 3. Shapley values in each node

Pillar	Node	Criteria	Shapley (%)
	FEEM SI	Environment	35.70
		Society	38.60
		Economy	25.70
Environmental Pillar	Environment	Natural Endowment	35.59
		Energy & Resources	30.70
		Pollution	33.71
	Natural Endowment	Water	48.59
		Biodiversity	51.41
	Biodiversity	Animals	51.07
		Plants	48.93
	Energy & Resources	Material Intensity	32.01
		Energy Intensity	31.93
		Renewables	36.06
	Pollution	GHG p.c.	37.33
		CO2 Intensity	33.65
		Waste	29.02
Social Pillar	Society	Vulnerability	29.47
		Well-Being	41.19
		Transparency	29.34
	Vulnerability	Food Relevance	33.91
		Private Health	32.28
		Energy Security	33.81
	Energy Security	Energy Imported	29.69
		Energy Access	70.31
	Well-Being	Population Density	21.04
		Education	49.09
		Life Expectancy	29.87
	Transparency	Corruption	70.47
		ICT Access	29.53
Economy Pillar	Economy	Growth Drivers	38.34
		Exposure	31.42
		GDP p.c.	30.24
	Growth Drivers	R&D	56.92
		Investment	43.08
	Exposure	Relative Trade Balance	57.69
		Public Debt	42.31

Table 4. Marginal criteria importance in the second level

Pillar	Criteria	Marginal Importance (%)
Environment	Natural Endowment	12.71
	Energy & Resources	10.96
	Pollution	12.03
Society	Vulnerability	11.38
	Well-Being	15.90
	Transparency	11.32
Economy	Growth Drivers	9.85
	Exposure	8.07
	GDP p.c.	7.77

Algorithms Based on Computational Intelligence for Autonomous Physical Rehabilitation at Home

Nunzio Alberto Borghese[1], Pier Luca Lanzi[2],
Renato Mainetti[1], Michele Pirovano[1], and Elif Surer[1]

[1] Department of Computer Science, University of Milan, Milan, Italy
{alberto.borghese,renato.mainetti,michele.pirovano,
elif.surer}@unimi.it
[2] Department of Electronics, Information Science and Bioengineering,
Polytechnic University of Milan, Milan, Italy
pierluca.lanzi@polimi.it

Abstract. Exergames provide efficient and motivating training mechanisms to support physical rehabilitation at home. Nonetheless, current exergame examples lack some important aspects which cannot be disregarded in rehabilitation. Exergames should: (i) modify the game difficulty adapting to patient's gameplay performance, (ii) monitor if the exercise is correctly executed, and (iii) provide continuous motivation. In this study, we present a game engine which implements computer intelligence-based solutions to provide real-time adaptation, on-line monitoring and an engaging gameplay experience. The game engine applies real-time adaptation using the Quest Bayesian approach to modify the game difficulty according to the patient's performance. Besides, it employs a fuzzy system to monitor the execution of the exercises according to the parameters set by the therapists and provides on-line feedback to guide the patient during the execution of the exercise. Finally, a motivating game experience is provided using rewards and adding random enrichments during the game.

Keywords: Rehabilitation, Bayesian optimization, Gamification, Exergames, Fuzzy systems.

1 Introduction

Stroke is the leading cause of death in adults [1] and physical rehabilitation is needed to recover from its consequences [2]. With stroke figures increasing every year, rehabilitation is bound to have an even bigger impact on the expected costs of healthcare providers. On the other hand, healthcare providers strive to reduce those same costs, reducing personnel, discharging patients as soon as possible and reducing support. Rehabilitation is presently based on intensive exercising with an almost daily schedule that is carried out with a therapist; new solutions that can extend traditional rehabilitation, or even replace it, are needed.

Exergaming represents one promising solution to this problem. Exergames merge the therapeutic effects of exercises to the engaging factors of games. They guide

patients to autonomously perform the exercises required by the rehabilitation therapists, thus requiring less supervision and opening to the possibility of self-rehabilitation at home, further reducing costs both for the healthcare provider and for the patient on one side, and supporting intensive rehabilitation on the other side. In addition, exergames add a layer of engagement to the traditionally boring repetitive exercising required by rehabilitation session, allowing the patient to exercise longer and have even fun while doing it [3]. Several instances of successful rehabilitation exergames have been reported [4], but there are still many open problems concerning how to properly structure exergames so that the therapeutic validity of the exercises is not sacrificed and no hazard occurs to the patient. For this reason, exergames are presently used mainly inside the rehabilitation centers or with remote supervision by the therapist (tele-rehabilitation).

Recent commercial video games require the player to move in order to play correctly. For instance, WiiFit is a collection of games produced for the Nintendo Wii, a gaming console that requires the player to use her body movements to play, has been successfully used for in-hospital rehabilitation [5]. However, video games such as Wii Fit have been designed with a different scope in mind than professional rehabilitation and they lack many of the needed features that would make such exergames usable without the presence of a therapist. For instance, commercial games usually impose fixed difficulty settings, making them impractical for impaired people; they do not track nor log the player's movements and do not provide any supervision. These practical considerations are of utmost importance if we want exergames to be useful in an at-home setting, and thus further reduce the costs of healthcare.

We propose here, as a solution to these issues, the application of intelligent algorithms inside a game engine to support the features needed for a correct therapy. Such features can be summarized as follows [6]: a) **Scheduling**, as the configuration of an exercises session must be carefully tailored to the patient's current state; b) **Adaptation**, as the same exergames should be played by patients with varying degrees of impairment and still provide a reasonable challenge. c) **On-line monitoring**, as the movements of the patient must be evaluated in real-time to avoid dangerous movements and maladaptation. d) **Assessment**, as the results of the exercises have to be reviewed by the therapist for tuning and evaluation. e) **Motivation**, as the gaming elements of the exergames must be carefully designed to maximize compliance both in the short and in the long term. Specific intelligence-based algorithms can be applied to all these different features, allowing exergames to be safely used at home. We delineate here how computational intelligence can be used to empower exergames, and we focus especially on the solutions we have introduced for adaptation and monitoring. We incorporated the design considerations we have presented into a game engine that we have fully realized and that is currently being tested by patients.

2   Methodology

In this section, our game engine and its computational intelligence based solutions that mainly address adaptation and monitoring, are presented in detail.

2.1   Intelligent Game Engine for Rehabilitation

The Intelligent Game Engine (IGER) [7] has been developed as part of Rewire platform1. The main objective of Rewire is to develop a low-cost game-based platform that assists patients, discharged from the hospital, to pursue their rehabilitation autonomously at home under remote supervision of the clinicians. The Rewire platform is constituted of three main components: a patient station (PS), a hospital station (HS), and a networking station (NS). The HS allows therapists to monitor remotely the ongoing rehabilitation at patient's home. It also provides features for the therapists to schedule rehabilitation exergames and assess the rehabilitation outcomes. The NS provides advanced data mining tools that analyze common features of rehabilitation treatment among hospitals and regions. The PS is installed at the patient's home and it is the core of the Rewire platform. It guides the patient through exercises prescribed by the HS by games.

The PS has four main modules. The hospital module is used by patients to interact with the therapists at the hospital through audio/video communication and to download their scheduled exergames. The lifestyle module collects data on the patient's daily activity through a body sensor network. The community module allows patients to interact with other patients online. Finally, the IGER module guides the actual rehabilitation using 3D exergames. IGER's structure (Figure 1) and its computational intelligence based features are the focus of the present work.

IGER has at its core a typical game engine, i.e. a software architecture designed to support video games. We have used Panda3D here, but any other game engine (e.g. Unity 3D) can be used. The Game engine has been expanded through scripting language to support a set of features tailored to rehabilitation, and especially post-stroke physical rehabilitation for balance and posture.
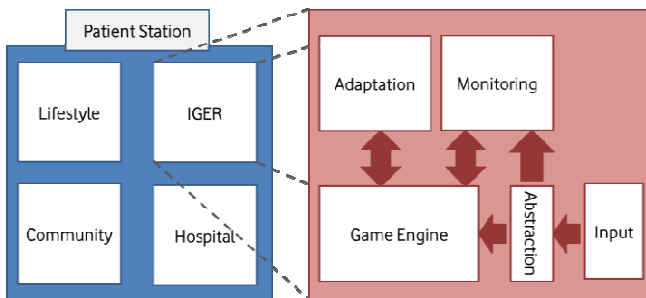
Fig. 1. The components of the PS and the IGER module

The IGER module embodies two components: the game engine and the game control unit. The former provides all the gaming functionalities: input data, animation, collision detection, rendering, scoring and game logic. The latter contains the game control unit that provides the features needed for rehabilitation-oriented exergames. IGER presents modules for therapist-driven scheduling and assessment.

---

1   REWIRE project: http://www.rewire-project.eu

The exercises associated to each rehabilitation session are configured off-line at the hospital by the therapists and their parameters are configured for each patient. The therapists can select which exercises and consequently which games to include in each training session, the duration of each exercise, the number of repetitions, and the difficulty level. In addition, the therapists can also select the most suitable input device for each rehabilitation task. Finally, the therapist can access the detailed outcome of the rehabilitation session and assess the patient's performance and status.

2.2   Adaptation

In video-games field, the widely accepted flow theory [8] defines flow as a state of heightened focus during which he/she is completely immersed in the activity, time flies by, and external inputs are ignored; a person can enter the state of flow while performing an activity, on the condition that the difficulty of the activity matches the person's skills. This state is desirable because it is engaging, and many games strive to provide it by changing their difficulty in real-time, either through their intrinsic design or through specific Dynamic Difficulty Adaptation (DDA) algorithms.

To design a DDA algorithm, one or more parameters can be selected for adaptation, making sure to choose parameters that directly affect the difficulty of the task. A simple method uses heuristics to change dynamically the parameters according to a metric derived from the patient's input, such as the number of correct trials performed during a session of play.

An alternative is to resort to statistical methods to support adaptation. In our system, we perform adaptation through a Bayesian framework, based on the QUEST psychometric method [9], to update one parameter and adapt it to the patient's behavior. The parameter is changes on a trial or epoch basis, analyzing the success rate in the game tasks such that the overall rate approaches a desired score. The optimal parameter value according to the therapist plays the role of a-priori value. This method is based on three assumptions that are all satisfied in our case. First, the function that relates the parameter to the performance in the game must have the same shape under all conditions. This can be obtained by considering an adequate function (in our case, Weibull distribution). Second, the subject's threshold should not vary from trial to trial. Third, individual trials must be statistically independent. The last two assumptions are satisfied, since the adaptation scope is a single session of a game and thus patient's skill improvement can be safely disregarded.

2.3   On-line Monitoring

One of the most important roles of the physical rehabilitation therapist, if not the one of utmost importance, is real-time monitoring of the user's performance and movements. Video-games, in fact, try to challenge the user to move fast and this can become dangerous for patients with limited motion control and capabilities [10]. Monitoring is aimed to avoid mal-adaptation, i.e. it avoids making the patient perform movements that would be detrimental to the therapy while attempting to perform the exercises correctly.

In classical rehabilitation the exercises are performed in presence of the therapist who can correct the posture of the patient when she does not move correctly through verbal feedback or direct intervention. In an autonomous setting, such as at-home rehabilitation, the watching eye of the therapist must be replaced by a suitable software system. In fact, the focus of the exergames alone may be on the games instead of on the exercise (and this is all the more likely if a state of flow is reached), we see that correct and responsive monitoring becomes even more important. In addition, consider that the autonomous and nonintrusive system cannot directly intervene on the patient's movement as a therapist would do, hence why clear visual and audial feedback is needed. Note that correct monitoring cannot be achieved with commercial active games, as they provide feedback based on the game's results but fail to give useful feedback to the patient about her actual movements. This might be one of the causes the leads to a relatively high drop out in the first tests on using exergames for rehabilitation at home. Monitoring can greatly benefit from Computational Intelligence that can provide algorithms suitable to be used to analyze in real-time the movements of the patient and select what feedback to give to the patient as well as its frequency in time.

From a high-level perspective, we can see monitoring as being composed of three parts: the inputs given by the patient (i.e. her movements), the rules that dictate when the movements are correct and what is more important, as chosen by the therapist, and the generated feedback. A fuzzy system [11] has been developed for this purpose (Figure 2).
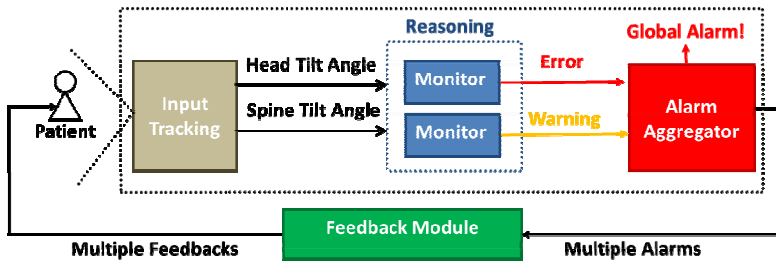
Fig. 2. The scheme of the monitoring mechanism in IGER

The rules on the movements are created by the therapist inside the hospital and incorporate the therapist's knowledge and requirements on the exercises. These rules map the input variables to the alarm level output, according to the monitor configurations; they are translated into fuzzy clauses that are associated with an alarm level. For instance: "If bend laterally_MEDIUM spine, raise alarm_MEDIUM". At run-time the monitoring system receives as input all the motion data and feeds them to a fuzzy inference engine that outputs an alarm level for each chosen monitor. Such alarm level is used by the feedback system to notify the patient of wrong movements. From the different alarm level a single global alarm level is raised through defuzzification. The alarm level is transformed into a color code that is applied to the avatar:

Fig. 3. Color-coded monitoring examples from games (a)Fruit Catcher (b)Animal Hurdler

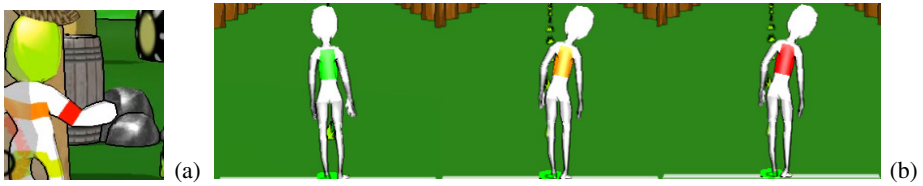white means no error and a color progressively going towards red indicates dangerous postures. The patient can therefore see immediately how she is performing (Figure 3). If the patient movement is beyond the maximum allowed range defined by the therapist, the game pauses and a virtual therapist appears and advice the patient on how to perform the exercise correctly. Afterwards, the game resumes.

3   Results and Discussion

IGER offers efficient exergames for rehabilitation and provides computer-intelligence solutions to adaptation, monitoring and motivation mechanisms as described in the previous sections. These solutions allow us to tailor rehabilitation to the patient's performance, monitor the correctness of the exercise and motivate the patient by combining good game design principles with an engaging gaming environment. We have designed and implemented a total of 18 exergames that incorporate all these features and that are aimed to posture and balance and neglect rehabilitation. These games have been designed starting from the inputs provided by the therapists.

For instance, the Animal Hurdler exergame (Figure 4a) aims at training patients on balance, while providing an additional cognitive load (dual task). In the gameplay, the patient is guided to raise her leg such that the avatar can step over worms that approach it. To step over the creatures, the patient, alternatively, can move laterally (lateral step). Thus, the game supports several exercises: lateral stepping and leg-rising exercises. Leg rising exercise can be tracked through Kinect device and a balance board can also be used to have a direct and reliable estimate of the center of mass. When lateral stepping exercise is performed, motion is tracked only with the Kinect device.

An additional cognitive load can be added to the game, by asking the patient to keep his arm at ninety degrees, asking him to let her avatar to carry logs with its arms (Figure 4b).

The performance of the patient in each exergame is computed by considering how many worms have been caught in respect to how many have been missed on a set of trials. The adaptable parameter of Animal Hurdler is set as the number of approaching worms per minute. Therefore, when the patient performs well during the exergame, the number of animals is incremented so that the user has to perform the same exercise more times in the same time period. The animal size and the speed of the animal are also adaptable parameters. However, it was considered unsafe to ask the patient to perform higher or faster steps and therefore the height of the foot and the speed of the motion are not controlled or adapted.
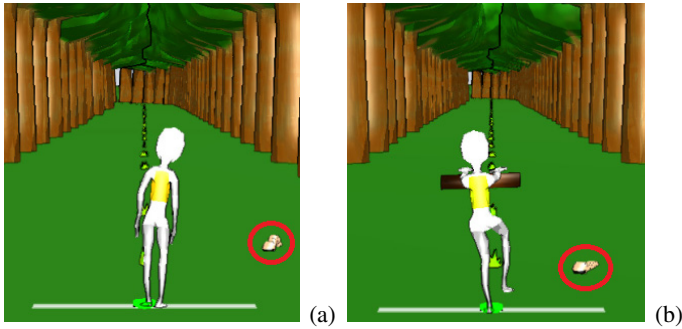
Fig. 4. (a) Animal Hurdler exergame (b) Animal Hurdler exergame with dual task

Monitoring is applied on the avatar's spine and neck, to make sure the patient does not bend them, and on the COP of the patient to make sure she does not move around the playing area. As described in Section 2.3, the therapist sets the optimal range for the monitoring exercises and visual and audial feedback is given to the patient depending on the alarm severity.

Another game example is Fruit Catcher (Figure 5), which is designed for weight-shift and lateral steps exercises. In Fruit Catcher, the player must catch fruits falling from the top of a tree. The player, represented in third view by an avatar, stands below the tree with a basket on her head and can either shift her body to the left and to the right, while keeping the feet still on the ground, or move the body laterally to catch the fruits in the basket; depending on the requisite of the exercise.

Fig. 5. Fruit Catcher exergame

The game can be played with different devices for different exercise goals, with a balance board (weight shift) or a Kinect (lateral steps). The performance of the patient is measured by the number of fruits caught by the patient with respect to the number of fruits missed. The adaptable parameter of Fruit Catcher is set as the frequency of the falling fruits. Monitoring is applied on the avatar's spine and neck, to make sure the patient does not bend them, and on the COP of the patient to make sure she does not move around the playing area or she does not exceed the defined limits. The bending of the knee can be also monitored, as well as the movement of the arms.

Motivation is yet another factor that must be taken into account for physical rehabilitation, as exercising is hard and can be painful. This has even more importance for autonomous rehabilitation, as in absence of a therapist that orders the execution of the exercises, intrinsic motivation of the patient (who is willing to improve) has to be supported with external motivation provided to the patient with different means.

Exergames possess an intrinsic motivational factor in their gaming nature which strives to make the exercises fun to execute, provided the game that is linked to the exercise is well designed. The game alone, however, comes short to provide motivation to the patient, especially due to the limits on the gameplay imposed by the exercise and due to prolonged exercising periods. As such, additional motivational factors can be added to the experience, such as enrichments and reward systems.

Using IGER, we have developed a set of eleven exergames for balance and posture rehabilitation, six of which have been thoroughly tested for usability with post-stroke patients [7]. IGER has also been used to create nine exergames for neglect rehabilitation, highlighting how the engine is flexible enough to be used for the rehabilitation of different pathologies. Two of the games can be used for both neglect and posture rehabilitation. A pilot test with patients in their own homes has started at the end of April 2014. Preliminary results are quite promising. Full results will be available at the end of 2014.

4   Conclusion

Different types of computer intelligence based solutions are implemented in IGER that offer real-time performance-based adaptation, on-line monitoring and a dynamic motivation mechanism. These features provide an ever-changing gameplay experience tailored to the patient's performance. Besides, during the gameplay, correct execution of the exercises is targeted by monitoring and guiding the patient based on the parameters defined by the therapists. These features can make rehabilitation at home an efficient, personalized, safe and engaging option.

References

1. Warlow, C., Sandercock, P., Hankey, G., et al.: Stroke: Practical Management. Blackwell Publishing (2008)
2. Langhorne, P., Coupar, F., Pollock, A.: Motor recovery after stroke: a systematic review. The Lancet Neurology 8(8), 741–754 (2009)
3. Rizzo, A., Kim, G.J.: A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. Presence 14(2), 119–146 (2005)
4. Ruiva, J.A.: Exergames and cardiac rehabilitation: a review. Journal of Cardiopulmonary Rehabilitation and Prevention 34(1), 2–20 (2014)

5. Sugarman, H., Weisel-Eichler, A., Burstin, A., Brown, R.: Use of the Wii Fit system for the treatment of balance problems in the elderly: A feasibility study. In: Virtual Rehabilitation International Conference, pp. 111–116 (2009)

6. Borghese, N.A., Pirovano, M., Lanzi, P.L., Wüest, S., de Bruin, E.D.: Computational Intelligence and Game Design for Effective At-Home Stroke Rehabilitation. Games for Health Journal 2(2), 81–88 (2013)

7. Wüest, S., Borghese, N.A., Pirovano, M., Mainetti, R., van de Langenberg, R., de Bruin, E.D.: Usability and Effects of an Exergame-Based Balance Training Program. Games for Health Journal 3(2), 106–114 (2014)

8. Csikszentmihalyi, M.: The flow experience and its significance for human psychology (1988)

9. Pirovano, M., Mainetti, R., Baud-Bovy, G., Lanzi, P.L., Borghese, N.A.: Self-adaptive games for rehabilitation at home. In: Computational Intelligence and Games, pp. 179–186 (2012)

10. Prosperini, L., Fortuna, D., Giannì, C., Leonardi, L., Marchetti, M., Pozzilli, C.: Home-based Training Using the Wii Balance board: a Randomized, Corssover Pilot Study in Multiple Sclerosis. Neurorehabilitation and Neural Repair 27(6), 516–525 (2013)

11. Mamdani, E.H.: Application of fuzzy algorithms for the control of a simple dynamic plant. Proceedings of IEEE 121(12), 121–159 (1974)

A Predictive Approach Based on Neural Network Models for Building Automation Systems

Davide De March[1,2], Matteo Borrotti[1,2], Luca Sartore[2], Debora Slanz[1,2], Lorenzo Podestà[3], and Irene Poli[1,2]

[1] Department of Environmental Sciences, Informatics and Statistics
University Ca' Foscari of Venice, Dorsoduro 2137, 30123 Venice, Italy
[2] European Centre for Living Technology, San Marco 2940, 30124, Venice, Italy
[3] R & D SYSTEMS S.r.l., via Fornaci 35, 38068, Rovereto (Trento), Italy
{davidedemarch,matteo.borrotti,debora.slanzi,
luca.sartore,irenpoli}@unive.it, lorenzo.podesta@rdsystems.it

Abstract. In this paper we address the problem of developing a control strategy to reduce the building energy consumption and reach indoor comfort levels. For this multiple and conflicting objectives optimisation we develop an approach based on stochastic feed-forward neural network models with ARIMA model predictions considered as input variables for networks. Studying real data from a sensorised office located in Rovereto (Italy) we develop the approach and achieve results exhibiting the very good performance of this predictive procedure.

Keywords: Feed-forward neural networks, Prediction, Time Series models, Building Automation System (BAS), Energy Efficiency.

1 Introduction

Energy consumption and carbon footprint are most fundamental issues that our economies and societies are currently addressing to assure a sustainable development. Buildings in particular, both residential and commercial, are responsible for more than 40% of the energy consumption, and this level will rapidly increase if drastic strategies will not be adopted. A critical target of recent EU policies is to transform existing buildings into nearly zero-energy consumption by 2020. In the renovation of existing buildings, an important role is played by building automation systems (BAS) that require new and efficient strategies to work with low energy levels. Control strategies of these systems have to optimise multiple and conflicting objectives, such as low energy consumption and indoor comfort levels. To develop efficient control strategies with these objectives, we have to address the problem of high dimensionality of the system: a very large set of parameters (dimensions) and a complex interacting network are known to affect the dynamical behaviour of the system and have to be involved in any strategy formulation. In current literature, two main approaches are proposed for modelling BAS energy consumption [1]. The former is based on the study of

the physical system and it is formulated with differential equation sets which describe thermodynamic aspects and features of the environment [2,3]. The latter consists of models (linear and non-linear) which represent the stochastic processes underlying by observed time series with the aim of predicting accurate future dynamics to incorporate in the control strategy [4,5].

Energy observed time series are usually collected by building automation systems and modelled considering both auto-regression (AR) and moving average (MA) components [6,7]. For non-stationary time series, ARIMA models are frequently considered [6], and when exogenous variables are introduced also the AR-MAX class of models is derived [8]. Non-linearity in observed dynamics is commonly modelled with artificial neural networks (ANN) as in [9,10]. Interesting are the studies on comfort management with feed-forward neural networks [11] and on predicting lighting and heating systems with radial basis function neural networks [11] or recurrent networks [12]. Other valuable modelling strategies are based on evolutionary neural networks [13] and fuzzy networks, which combine the advantages of both neural networks and fuzzy logic mostly for blur data [14].

In this paper we develop an approach based on stochastic feed-forward neural networks to provide accurate predictions for three building automation system response variables, studying sets of real data recorded in a sensorised office. The system response variables are: energy consumption, indoor thermal comfort and indoor lighting comfort. To develop an optimal control strategy we first derive a selected set of explanatory variables using different statistical variable selection procedures. We then construct ARIMA predictions on these variables to achieve a data set which includes both observed data and univariate predictions. On this composite data set we further construct neural network models and derive accurate predictions for the three system response variables. The research is developed on data collected in a sensorised office located in Rovereto (Italy). The paper is organized as follows: in Section 2 we present the predictive approach by describing the variable selection procedure adopted to reduce the dimensionality of the problem and the construction of the neural network models to derive accurate predictions. In Section 3 we present the case study and the particular modelling strategy that we adopt. In Section 4 we present the results and we evaluate the accuracy of the achieved predictions and the global performance of the approach.

2   The Predictive Approach

Given a set of time series with N observations and denoting by $x_{1,t}, \ldots, x_{p,t}$ the informative variables at time t, $t = 1, \ldots, N$, and by $y_{j,t}$ the j-th system response variable at time t, we develop a general approach to predict system responses for energy efficient buildings by means of stochastic neural networks based on ARIMA model predictions. This approach involves several statistical procedures that are merged together to enhance the predictive performance of neural network models. In particular, we derive the predicted values $\hat{y}_{j,t+\tau}$, $\tau = 1, \ldots, T$, for each response variable according to the procedure represented in Fig. 1.
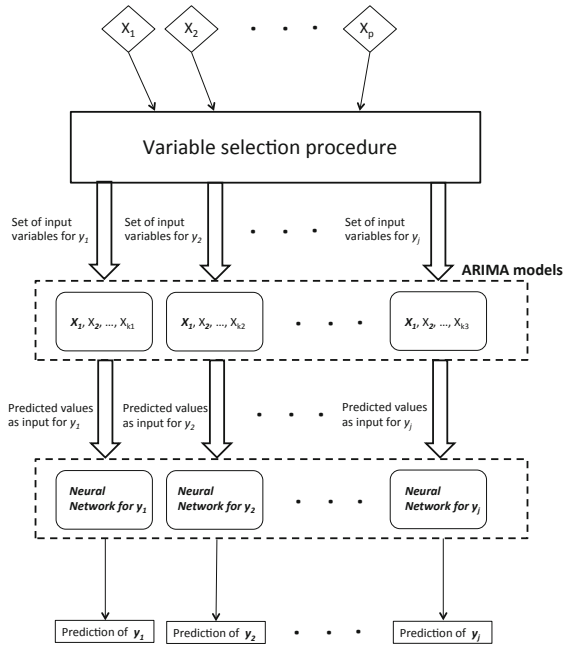
Fig. 1. The general structure of the predictive approach

2.1   Variable Selection Procedure

Addressing the problem of deriving a design for energy efficient buildings, a large set of variables has to be considered and modelled. Some variables represent the endogenous characteristics of a particular state of the system and they are recorded at fixed time intervals. In estimating statistical models to predict the dynamical behaviour of BAS responses, this large set of state variables can be in a *sparse* space where some variables affect more than others the system response. To select just the most informative variables from the very large and frequently noisy set initially collected, we adopt variable selection procedures, combining three different approaches:

- the physical mathematical formulation of each response variable;
- the Spearman correlation index;
- the non-linear variable selection approach based on Random Forests.

Each system response variable is generally described by a physical mathematical formulation based on a set of endogenous variables representing the state of the system [2,3]. Developing our predictive procedure, we decide to select *a priori* all the state variables which are involved in the physical formulation of the system response. Then we identify all the variables that are linearly related with the response by computing the Spearman correlation coefficient ρ [15]. We

select those variables that show a significant linear correlation with the response variable achieving $|\rho| > 0.5$.

To identify the informative variables with a non-linear relationship with the system response, we develop a selection procedure based on Random Forest approach. Random Forests have been introduced by [16] and they can be used to rank the importance of a set of variables in a non linear regression analysis. The principle of Random Forests is to combine many regression trees [17] built using bootstrap training samples and randomly choosing at each node of the tree a subset of state variables; successively a prediction performance is computed. Variables that largely influence the prediction error of the regression trees are considered the most influential for the system response: we select the state variables whose normalised prediction error is larger than 0.5 [18].

2.2   The Predictive Neural Network Based on ARIMA

The second phase involves the construction of ARIMA models for those variables selected by the procedure described in Sect. 2.1 and the prediction of each univariate time series $\hat{x}_{i,t+\tau}$, with $i = 1, \ldots, p$ and $\tau = 1, \ldots, T$.

For each of these variables, an ARIMA model is estimated according to the procedure described in [19], that returns the best model with the lowest AIC value. Each ARIMA model is then used to predict the univariate time series for the following τ observations with $\tau = 1, \ldots, T$. In the third phase of the procedure we build a class of sigmoidal feed-forward neural network models, one for each system response variable. The sigmoidal neural networks use $x_{i,t}$ and the ARIMA time series predictions, $\hat{x}_{i,t+\tau}$, as input variables to predict the system responses $\hat{y}_{j,t+\tau}$, $\tau = 1, \ldots, T$. The network topology involves one hidden layer with a number of nodes changing in a specific finite interval, a sigmoidal activation function between the input and the hidden layer and a linear activation function between the hidden and the output layer [20]. All the networks are trained by means of back-error propagation algorithm [21].

To identify which neural network topology can be used to predict each system response, we adopt a bootstrap procedure with B resamples [22]. At each bth run, with $b = 1, \ldots, B$, we select a set of n_b time series observations of the input variables (with $n_b < N$) and we predict $\hat{x}_{i,n_b+\tau}$ observations with the ARIMA models estimated on the n_b observations. We then estimate all the neural networks, whose number of nodes in the hidden layer changes in the defined interval, and we evaluate their prediction error for the future unknown T values of the response. Therefore, at each bth run of the bootstrap procedure we compute a predictive error for each of the chosen topology. We iterate the procedure for B resamples and we identify the topology which minimises simultaneously the Bootstrap Mean Absolute Error (BMAE), as presented in Eq. 1, and its standard deviation.

Once the topology of the network has been identified with the bootstrap procedure, we then proceed to estimate its parameters (weights). We generate a set of different random weights to initialise the network and we train all the networks on the $N - \mathrm{T}$ observed data by means of back-error propagation. Networks are tested for their predictive performance on the remaining T values. The final network is the one which minimises the Mean Absolute Percentage Error (MAPE) as in Eq. 2.

$$BMAE = \frac{1}{\mathrm{T}B} \sum_{b=1}^{B} \sum_{\tau=1}^{\mathrm{T}} |\hat{y}_{n_b+\tau,b} - y_{n_b+\tau,b}| \tag{1}$$

$$MAPE = \frac{1}{\mathrm{T}} \sum_{\tau=1}^{\mathrm{T}} \left| \frac{\hat{y}_{N-\mathrm{T}+\tau} - y_{N-\mathrm{T}+\tau}}{y_{N-\mathrm{T}+\tau}} \right| \tag{2}$$

3 Predicting Building Automation Systems

We construct and test our approach in a real case study addressing a sensorised office located in Rovereto (Italy). In this office, a set of installed sensors are used to record the most relevant **state variables** that affect the energy consumption and the levels of comfort for the office users. In particular, we consider:

- *indoor state variables*: internal temperature (x_1), humidity (x_2), air velocity (x_3), central mean radiant temperature (x_4), west luminosity (x_5), east luminosity (x_6), CO2 concentration (x_7), occupancy (x_8), window sensor (x_9), door sensor (x_{10}), corridor temperature (x_{11}) and fan coil thermal power (x_{12});
- *outside state variables*: outside temperature (x_{13}), outside illuminance (x_{14}), outside radiation (x_{15}) and outside humidity (x_{16}).

A set of **controllable variables** are identified and codified: the power of the fan coil (f_c), the position of the blinder (b) and two dimmable lights $(d_1$ and $d_2)$.

As **building automation system responses** we measure the total electric power (y_1) - which includes thermal and electric consumptions - and two comfort indices for inhabitants: the Predictive Mean Vote (PMV, y_2) and the Daylight Glare Index (DGI, y_3)[23][24]. PMV measures the level of satisfaction of office users with respect to the thermal environment and it is mostly influenced by temperature, humidity, air velocity and central mean radiant temperature observed in the room. DGI expresses discomfort glare due to the lighting system and depends on luminosity inside the room and electromagnetic radiation given off by the sun.

4   System Response Predictions

We develop the procedure presented in Sec. 2 to achieve the best sigmoidal feed-forward neural network models for predicting the three response variables of the building automation system. Each response variable - Total electric Power (y_1), PMV (y_2) and DGI (y_3) - is predicted by a different neural network topology.

According to the structure presented in Sec. 2, we identify the set of relevant variables for the prediction of each response variable. For Total Electric Power (y_1) we select the following relevant variables: Internal Temperature, x_1; Humidity, x_2; Central mean radiant Temp, x_4; CO_2, x_7; Corridor temperature, x_{11}; Fan Coil thermal power, x_{12}. For PMV (y_2) we select the variables: Internal Temperature, x_1; Humidity, x_2; Central mean radiant Temp, x_4; Corridor temperature, x_{11}. At last, for the third response variable DGI (y_3) we identify the following relevant variables: Humidity, x_2; West luminosity, x_5; East luminosity, x_6; Fan Coil thermal power, x_{12} and Outside Radiation, x_{15}. For each variable, the specific ARIMA model is estimated and used to achieve predicted values $\hat{x}_{i,t+\tau}$, $\tau = 1, \ldots, T$.

We then proceed in the construction of the general approach by selecting a sigmoidal feed-forward neural network with one hidden layer, whose number of nodes ranges from 2 to 20, as described in Sec. 2.2. We select the topology by means of a bootstrap procedure, that has been run for $B = 30$, where each resample uses $n_B = 5000$ time series observations to train the neural network model and the successive T = 72 (6 hours) time series observations to test and validate the results of the predictions. After having identified the topology, we estimate the weights associated to each neural network model, adopting the complete dataset of $N = 29362$ except for the last T = 72 observations which are used as test set. With this general approach, the best model for estimating the Total Electric Power (y_1) involves 12 neurons in the hidden layer and the predicted values $\hat{y}_1(t)$ can be described as a function of the following variables:

$$\hat{y}_1(t) = f(d_1(t), d_2(t), b(t), fc(t), \hat{x_1}(t), \hat{x_2}(t), \hat{x_4}(t), \hat{x_7}(t), \hat{x_{11}}(t), \hat{x_{12}}(t),$$
$$\hat{y_2}(t), \hat{y_3}(t), y_1(t-1)).$$

In this expression, d_1 and d_2 are dimmable lights levels, b is the position of the blinder, fc is the fan coil level and the others represents the selected input variables, predicted by ARIMA models. The response variable itself has been used in the model with a temporal one-lag delay to provide auro-regressive information (as suggested in [9]) and the predicted values of $y_2(t)$ and $y_3(t)$ are also introduced in the model.

With the same procedure, we identify the best topology for PMV, y_2, which is characterised by 11 neurons in the hidden layer and takes the following form:

$$\hat{y}_2(t) = f(d_1(t), d_2(t), b(t), fc(t), \hat{x_1}(t), \hat{x_2}(t), \hat{x_4}(t), \hat{x_{11}}(t), y_2(t-1)),$$

Similarly we identify the best topology for DGI, y_3. The final model has 13 neurons in the hidden layer and takes the following form:
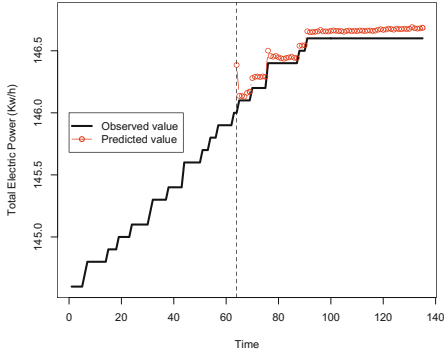
$$\hat{y_3}(t) = f(d_1(t), d_2(t), b(t), fc(t), \hat{x_2}(t), \hat{x_5}(t), \hat{x_6}(t), \hat{x_{12}}(t), \hat{x_{15}}(t), y_3(t-1))$$

To be confident about the behaviour of the estimated neural network models, we evaluate a class of radial basis neural networks and a class of recurrent Elman networks to develop a comparison among different approaches. In the radial basis neural network models we adopt the same strategy to evaluate the best predictive models: a three layer topology is created and a number of neurons from 2 to 20 in the hidden layer is tested. The bootstrap procedure is then run to obtain the BMAE value for each of the topology on the remaining T = 72 observations. In addition, we build a recurrent neural network topology and derive the Elman network architecture by adding a context layer (with the same number of neurons identified for the sigmoidal feed-forward neural networks) to a standard three layered feed-forward network and train them by means of back propagation. We compare these network models in their prediction accuracy on a test set composed of the remaining T = 72 observations (6 hours). The results, as described in Tab. 1, show that sigmoidal feed-forward neural networks are the models that better predict all the responses [10,25].
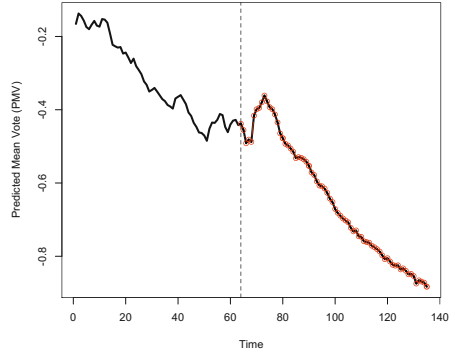
Table 1. Predictive performance metrics of sigmoidal and radial basis neural networks and Elman networks on a 6 hours prediction. In bold the best results obtained for each response. Standard deviations are presented in brackets.

Method	y_1		y_2 (PMV)		y_3 (DGI)	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
Sigmoidal FFN	**0.07** *(0.04)*	$\simeq$ **0** *($\simeq$ 0)*	$\simeq$ **0** *($\simeq$ 0)*	$\simeq$ **0** *($\simeq$ 0)*	$\simeq$ **0** *($\simeq$ 0)*	$\simeq$ **0** *($\simeq$ 0)*
Radial BFN	11.58 *(0.17)*	0.08 *($\simeq$ 0)*	0.51 *(0.20)*	0.79 *(0.20)*	3.35 *(0.23)*	5.10 *(1.28)*
Elman N	0.40 *(0.17)*	$\simeq$ **0** *($\simeq$ 0)*	0.08 *(0.05)*	0.14 *(0.13)*	0.21 *(0.09)*	0.32 *(0.18)*

Our approach based on sigmoidal feed-forward neural network gives very good predictions since all the criteria are equal or very close to 0. The sigmoidal feed-forward results exhibit also better performances in comparison with the results obtained using radial basis function network and Elman Network. We present in Fig. 2 the predicted responses in comparison with the actual values recorded by the sensors (Figure 2 presents only the subsample of the last 12 hours in order to make the plots easier to read). We notice that the predicted values for the three responses are very close to the actual real values. In particular we can notice that the prediction of y_2 and y_3 (Figs. *2b* and *2c*) perfectly overlaps the observed values, while only a very small difference can be noticed in Fig. *2a* for the prediction of y_1.

(a) Total eletric power consumption

(b) Predicted Mean Vote
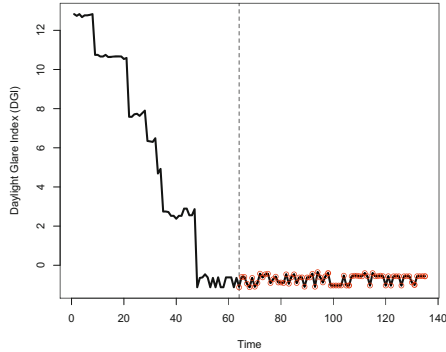
(c) Daylight Glare Index

Fig. 2. Comparison between observed and predicted values for the last 12 recorded hours of the three responses: (a) The Total Electric Power consumption, (b)The Predicted Mean Vote, and (c) The Daylight Glare Index.
Black lines describe the observed values recorded by sensors and red lines describe the predicted values of the sigmoidal feed-forward neural network.

References

1. Afram, A., Janabi-Sharifi, F.: Review of modeling methods for HVAC systems. Applied Thermal Engineering 67(1), 507–519 (2014)
2. Chen, Y., Treado, S.: Development of a simulation platform based on dynamic models for HVAC control analysis. Energy and Buildings 68, 376–386 (2014)
3. Avci, M., Erkoc, M., Rahmani, A., Asfour, S.: Model predictive HVAC load control in buildings using real-time electricity pricing. Energy and Buildings 60, 199–209 (2013)
4. Soyguder, S., Alli, H.: Predicting of fan speed for energy saving in HVAC system based on adaptive network based fuzzy inference system. Expert Systems with Applications 36(4), 8631–8638 (2009)
5. Oldewurtel, F., Parisio, A., Jones, C.N., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., Morari, M.: Use of model predictive control and weather forecasts for energy efficient building climate control. Energy and Buildings 45, 15–27 (2012)
6. Box, G.E.P., Jenkins, G.M.: Time series analysis, forecasting and control [by] George EP Box and Gwilym M. Jenkins. Holden-Day, San Francisco (1970)
7. Mustafaraj, G., Chen, J., Lowry, G.: Development of room temperature and relative humidity linear parametric models for an open office using BMS data. Energy and Buildings 42(3), 348–356 (2010)
8. Yiu, J.C.M., Wang, S.: Multiple ARMAX modeling scheme for forecasting air conditioning system performance. Energy Conversion and Management 48(8), 2276–2285 (2007)
9. Kusiak, A., Xu, G.: Modeling and optimization of HVAC systems using a dynamic neural network. Energy 42(1), 241–250 (2012)
10. Kusiak, A., Zeng, Y., Xu, G.: Minimizing energy consumption of an air handling unit with a computational intelligence approach. Energy and Buildings 60, 355–363 (2013)
11. Ferreira, P.M., Ruano, A.E., Silva, S., Conceição, E.Z.E.: Neural networks based predictive control for thermal comfort and energy savings in public buildings. Energy and Buildings 55, 238–251 (2012)
12. Dounis, A.I., Caraiscos, C.: Advanced control systems engineering for energy and comfort management in a building environment—a review. Renewable and Sustainable Energy Reviews 13(6), 1246–1261 (2009)
13. Zemella, G., De March, D., Borrotti, M., Poli, I.: Optimised design of energy efficient building façades via evolutionary neural networks. Energy and Buildings 43(12), 3297–3302 (2011)
14. Orosa, J.A.: A new modelling methodology to control HVAC systems. Expert Systems with Applications 38(4), 4505–4513 (2011)
15. Spearman, C.: The proof and measurement of association between two things. The American Journal of Psychology 15(1), 72–101 (1904)
16. Breiman, L.: Random forests. Machine Learning 45(1), 5–32 (2001)

17. Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A.: Classification and regression trees. CRC Press (1984)
18. Genuer, R., Poggi, J.M., Tuleau-Malot, C.: Variable selection using random forests. Pattern Recognition Letters 31(14), 2225–2236 (2010)
19. Hyndman, R.J., Khandakar, Y.: Automatic time series forecasting: The forecast package for r. Journal of Statistical Software 27(3), 1–22 (2008)
20. Haykin, S.S.: Neural networks and learning machines, vol. 3. Pearson Education, Upper Saddle River (2009)
21. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. (eds.) Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1, pp. 318–362. MIT Press, Cambridge (1986)
22. Efron, B., Tibshirani, R.J.: An introduction to the bootstrap, vol. 57. CRC Press (1994)
23. Bellia, L., Cesarano, A., Iuliano, G.F., Spada, G.: Daylight glare: a review of discomfort indexes. In: Proceedings of the International Workshop and 7th IEA Annex 45 Expert Meeting: Visual Quality and Energy Efficiency in Indoor Lighting: Today for Tomorrow (2008)
24. Fanger, P.O.: Thermal comfort: analysis and applications in environmental engineering. R.E. Krieger Pub. Co., Malabar (1982)
25. Sfetsos, A.: A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renewable Energy 21(1), 23–35 (2000)

Part VI

Emotional Expressions and Daily Cognitive Functions

Effects of Narrative Identities and Attachment Style on the Individual's Ability to Categorize Emotional Voices[*]

Anna Esposito[1,2], Davide Palumbo[1], and Alda Troncone[1]

[1] Department of Psychology, Second University of Naples, Caserta, Italy
[2] International Institute for Advanced Scientific Studies (IIASS), Vietri sul Mare, Salerno, Italy
iiass.annaesp@tin.it, da.palumbo@outlook.it,
alda.troncone@unina2.it

Abstract. This research aimed to assess individual's abilities in decoding emotional vocal expressions according to attachment styles and Narrative Identities. To this aims 30 students (15 females, 15 males; mean age = 21.4 ± 2.47) were recruited at the Second University of Naples (Italy) and underwent an emotional-voice-decoding task after being tested through the "Experience in Close Relationships" (ECR) and Personality Meaning (PMQ) Questionnaire to assess their attachment styles and Narrative Identities. The results showed that Outward subjects were more accurate in decoding joy and surprise especially in the group of individuals with an Insecure attachment style, suggesting that emotional regulation dynamics and attachment parameters shape the ways individuals develop their ability to decode other emotional feelings.

Keywords: Emotional vocal expressions, emotional voice decoding, attachment style, Narrative Identities.

1 Introduction

The human capacity to recognize emotions is considered a fundamental innate activity for social communication and survival [1]. In the abundant literature that covers the factors at the basis of this ability, the role of attachment style is of significant interest. The Attachment theory hypothesizes that Secure individuals have a greater capacity to recognize emotions when compared to those individuals that are Insecure [2]. Insecure individual demonstrate particular difficulty especially in identifying negative emotions such as sadness [3,4]. However, to date, the research offers contrasting results that do not allow for an unequivocal explanation. Secure individuals are not always the most able in recognizing emotions, and there is not always an agreement on which attachment style is the less accurate in this task [5]. For example, the Fearful attachment style, in some studies [6] is associated with a poor ability to recognize emotions. However, according to other authors it is the second most accurate closely behind the Secure attachment style [7].

---

[*] The names of the authors are in alphabetic order since each made a significant contribution to the research reported.

In this context, what has been neglected by current research investigations is the influence the Narrative Identity theory may play on the individual's attachment styles and her/his ability to recognize emotions [8,9,10,11].

The Narrative Identity theory was proposed in the context of the post-rationalist cognitive theory in order to "*mediate between the continuous aspect of identity and the variable….of personal experience*" (from [12], page 261).

According to this theory, the process of creating a Narrative Identity depends on both emotional regulation dynamics [13] and attachment parameters, in terms of the ways individuals develop their own emotional experiences shaped in turn by their attachment styles [14]. On the basis of this framework, individuals can be grouped into two basic categories of constructing identity: "Outward" or "Inward" (Narrative) Identity according to their ability to anticipate parental responses to their affective requests during childhood. For Inward individuals the ability to foresee the nature of the affective exchanges is quite apart from the quality and efficacy of the relationship with the caregiver. Outward individuals are capable of understanding their emotions by exploiting the same cognitive processes they exploit to understand situations and emotional responses of the caregiver. These basic identities differ in the regulation of emotional and cognitive processes: the Inward individuals are more focused on their inner experience whereas the Outward ones are more focused on external referential contexts.

Emotional voices are considered relevant signals/signs to understand other people's emotional states playing an important role in social interactions [15].

In the light of the above considerations, the present work hypothesizes that an individual with different attachment styles and Narrative Identities will show significant differences in her/his ability to decode emotional voices.

The hypothesis to be tested are:

• Secure subjects are more accurate than Insecure ones in decoding vocal emotional expressions;
• Inward/Outward Identities will play a role on the individuals' ability to recognize vocal emotional expressions. Outward individuals will be more accurate than Inward ones;
• There can be possible interactions between attachment styles and Narrative Identities.

2   Method

2.1   Participants

A sample of 30 subjects, equally balanced by gender and aged between 18 and 29 years was recruited at the Second University of Naples (Italy). The subjects (15 males, 15 females) with a mean age of 21.4 years (SD= 2.47) are mainly students at the Department of Psychology (76.7 %), and others departments (23.3%).

2.2   **Experimental Tools**

Attachment. The participants' attachment style was assessed by using the "Experience in Close Relationships" (ECR) questionnaire proposed by [16] and standardized for the Italian population by [17,18].

Narrative Identities. The individual's Narrative Identities were assessed through the Personality Meaning Questionnaire (PMQ) [19] identifying the key cognitive themes characterizing Inward or Outward personalities.

Vocal Emotional Voice Recognition Task. In order to evaluate the ability to recognize emotional feelings from voices, the subjects were asked to listen to a set of 20 emotional vocal stimuli associated with five out of the six basic emotions defined by [20]. The "emotional voices" were selected by one of the authors from a database of emotional voices already assessed and published in literature and details are therein [21,22].

2.3   **Procedure**

Each participant first filled in and signed a consent form providing his/her general information and completed the ECR and PMQ questionnaires for assessing the attachment style and the Narrative Identity. After filling out the questionnaires, they underwent the emotional-voice-decoding task. A suitable neutral setting was created in the laboratory, free of distractions and disturbing events. Each participant, after being informed of the ongoing experiment, was asked to listen to the emotional stimuli, that were randomly presented through headphones and asked to attribute one and only one of the following emotional labels: fear, sad, happy, anger, surprise, to each of the vocal stimuli listened by crossing the corresponding box on an answer grid reporting the five emotional labels. The stimuli were equally balanced among the 4 different emotion categories with 4 samples (two produced by an actor, two produced by an actress) for each emotion. Participants were allowed to listen to the stimuli no more than 3 times before selecting their answers.

2.4   **Data Analysis**

To assess the significance of the attachment style on the emotional decoding accuracy, an ANOVA analysis was performed with the attachment style as a between subject variable (2 levels of attachment Secure/Insecure) and the emotion categories as a within subject variable (5 levels for the 5 emotions considered). To assess the effects of the Inward/Outward Identity on the emotional decoding accuracy an ANOVA analysis was performed with the Narrative Identity as a between subject variable (2 levels, Inward/Outward) and emotions as a within subject variable (5 levels for the 5 emotions considered). Finally, to assess effects due to interactions between attachment styles and Narrative Identities, an ANOVA was performed on the percentage of correct emotional labels attributed to the each vocal stimulus by the Inward/Outward individuals with Secure and Insecure attachment style.

3   Results

Table 1 illustrates the distribution of attachment style and Narrative Identities among the participants. From such data it is possible to observe that 56.7% (n= 17) of the participants were Securely attached. Alternatively, 13.3 % (n=4) of the participants in the Insecure group were assessed as Avoidant, 26.7% (n=8) as Preoccupied and 3.3 % as Fearfully attached. With respect to Narrative Identities, 33.3% (n=10) were classified as Inward and 66.7% (n=20) as Outward individuals.

Table 1. Distribution of the attachment style and Narrative Identities among the participants

	Attachment style	
	Secure (n=17)	**Insecure** (n=13)
Gender	8 m, 9 f	7 m, 6 f
Inward (n=10)	6	4
Outward (n=20)	11	9

The overall percentages of correct identification obtained by all participants for each emotion category are reported as confusion matrices in Tables 2.

Table 2. The confusion matrix reporting the percentages of accuracy (on the diagonal) in the emotional voice decoding task obtained for the entire sample

Emotion to identify			% Answers		
	Joy	**Fear**	**Anger**	**Surprise**	**Sadness**
Joy	**80**	1.6	0.8	11.7	5.9
Fear	0.9	**90.8**	2.5	3.3	2.5
Anger	0.8	4.2	**86.7**	0	8.3
Surprise	4.2	1.6	1.7	**90.8**	1.7
Sadness	0	14.2	0.8	4.2	**80.8**

3.1   Attachment Style Effects

A one-way ANOVA was performed to test differences between Secure and Insecure subjects in the decoding accuracy of vocal emotional expressions. Results showed no significant differences between Secure and Insecure groups in each emotional category (joy $F_{(1,30)}$ =3.275, p=n.s.; fear $F_{(1,30)}$ =1.543, p=n.s.; anger $F_{(1,30)}$ =.191, p=n.s.; surprise $F_{(1,30)}$ =2.033, p=n.s.; sadness $F_{(1,30)}$ =.755, p=n.s.).

3.2   Narrative Identities Effects

Figure 1 displays the means of correct answers (for each emotional category under examination) obtained by grouping participants as Inward (white bars) and Outward (red bars) individuals. The data highlights a tendency of Outward individuals to perform better that Inward ones.
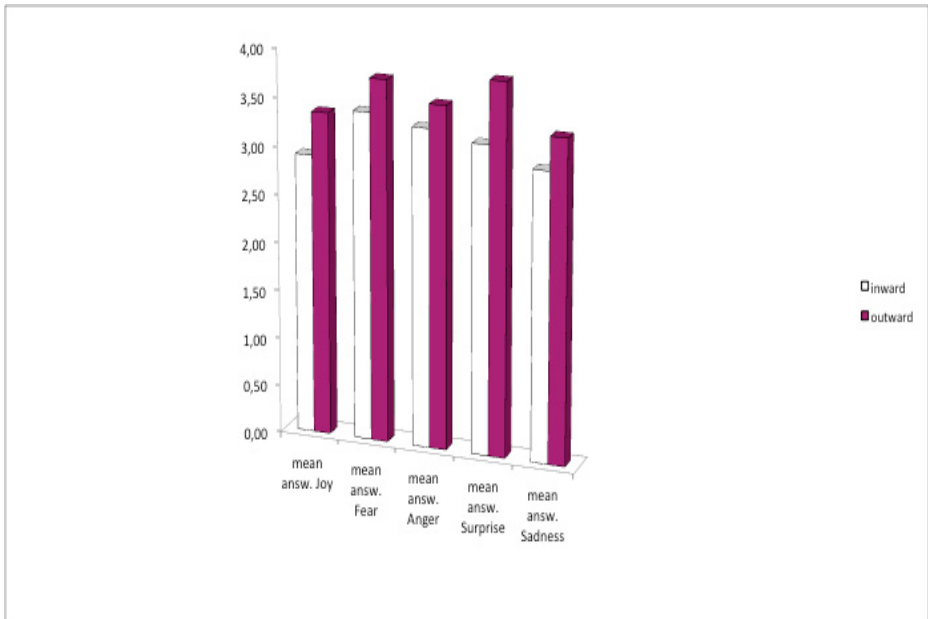
Fig. 1. Means of emotion's correct labeling by Inward and Outward individuals

A deeper analysis performed through a one-way ANOVA showed that Outward significantly outperformed Inward individuals only in the recognition of surprise (F(1,30) =4.345, p=.046, η^2=.1). No significant differences were found in the decoding of the remaining emotional categories (joy F(1,30) =2.447, p=n.s.; fear F(1,30) =1.882, p=n.s.; anger F(1,30) =.684, p=n.s.; p=n.s.; sadness F(1,30) =1.113, p=n.s.).

3.3   Attachment Style and Narrative Identities Interaction

In order to check if the attachment style and Narrative Identities affected the emotion-decoding task, a 2 x 2 (attachment style [Secure, Insecure] x Narrative Identities [Outward, Inward]) mixed analysis of variance (ANOVA) was conducted on each of the five emotional categories under examination, using the number of correct answers-to-emotional-voices as the dependent variable. The ANOVA revealed that Secure individuals outperformed Insecure ones in the decoding of joy (F= (3,30) =3.107, p=.023, η^2=.185) and surprise (F= (3,30) =3.613, p=.038, η^2=.155), whereas no effects were found for the remaining emotional categories (fear F(3,30) = 1.578, p=n.s.; anger F(3,30) =1.036, p=n.s.; sadness F(3,30) =1.142, p=n.s.). In addition, the ANOVA revealed that these effects were due to the fact that Insecure-Outward individuals significantly outperformed Insecure-Inward ones in the recognition of joy F(3,30) =4.063, p=.054, η^2=.135), and surprise (F(3,30) =6.563, p=.017, η^2=.202). No main effects of Narrative Identities were found for the remaining emotional categories (fear F(3,30) = 2.565, p=n.s.; anger F(3,30) = 1.108, p=n.s.; sadness F(3,30) = 1.660, p=n.s.). This data suggested that Outward subjects were more accurate in decoding joy and surprise especially in the group of individuals with an Insecure attachment style.

No interaction effects were found between attachment style and Narrative Identities for each emotion under examination (joy ($F_{(3,30)}= 2.461$, p=n.s.; fear $F_{(3,30)} = 1.017$, p=n.s.; anger $F_{(3,30)} = 2.251$, p=n.s.; surprise $F_{(3,30)}= 3.044$, p=n.s.; sadness $F_{(3,30)} = 1.435$, p=n.s.). Table 2 illustrates the details of this analysis.

Table 3. Averaged correct responses and standard deviations (SD) to the emotional vocal stimuli obtained by participants with Secure and Insecure attachment styles and Inward/Outward Narrative Identities.

| Emotional Voices | Narrative Identities | Attachment style | | |
| | | Secure | Insecure | |
		Mean (SD)	Mean (SD)	total
Joy	Inward	3.33 (.82)	2.25 (.96)	2.90 (.94)
	Outward	3.45 (.52)	3.22 (.67)	3.35 (.59)
	total	**3.41 (.62)$_a$**	**2.92 (.86)$_b$**	
Fear	Inward	3.67 (.52)	3.00 (1.41)	3.40 (.97)
	Outward	3.82 (.40)	3.67 (.50)	3.75 (.44)
	total	3.76 (.44)	3.46 (.88)	
Anger	Inward	3.50 (.84)	3.00 (1.41)	3.30 (1.06)
	Outward	3.36 (.67)	3.78 (.44)	3.55 (.61)
	total	3.41 (.71)	3.54 (.88)	
Surprise	Inward	3.67 (5.16)	2.50 (1.91)	**3.20 (1.32)$_a$**
	Outward	3.91 (.30)	3.78 (.44)	**3.85(.37)$_b$**
	total	**3.82 (.39)$_a$**	**3.38 (1.19)$_b$**	
Sadness	Inward	3.33 (.82)	2.50 (1.73)	3.00 (1.25)
	Outward	3.36 (.67)	3.33 (.50)	3.35 (.59)
	total	3.35 (.70)	3.08 (1.04)	

Means with different subscripts within a row or column are significantly different at p<.05.

4   Conclusions

This research aimed to assess individual's abilities in decoding emotional vocal expressions according to her/his attachment styles and Narrative Identities. It was discovered that Narrative Identities play a significant role, in particular for individuals with an Insecure attachment style. This legitimate the theoretical constructs and suggests that both emotional regulation dynamics and attachment parameters shape the ways individuals develop their own emotional experiences and their ability to decode other emotional feelings. These are however, the results of a pilot study. More data are needed to increase our understanding on how emotions are decoded and relate to the individual's personality style and experience.

References

1. Izard, C.E.: Innate and Universal Facial Expressions: Evidence from Developmental and Cross-Cultural Research. Psychol. Bull. 115, 288–299 (1994)
2. Harris, P.L.: Individual Differences in Understanding Emotion: The Role of Attachment Status and Psychological Discourse. Attach. Hum. Dev. 1, 307–324 (1999)
3. Laible, J.D., Thompson, R.A.: Attachment and Emotional Understanding in Preschool Children. Dev. Psychol. 34, 1038–1045 (1998)
4. Suslow, T., Dannlowski, U., Arolt, V., Ohrmann, P.: Adult Attachment Avoidance and Automatic Affective Response to Sad Facial Expressions. Aust. J. Psychol. 62, 181–187 (2010)
5. De Rosnay, M., Harris, P.L.: Individual Differences in Children's Understanding of Emotion: The Roles of Attachment and Language. Attach. Hum. Dev. 4, 39–54 (2002)
6. Colle, L., Del Giudice, M.: Patterns of Attachment and Emotional Competence in Middle Childhood. Soc. Dev. 20, 51–69 (2011)
7. Steele, H., Steele, M., Croft, C.: Early Attachment Predicts Emotion Recognition at 6 and 11 Years Old. Attach. Hum. Dev. 10, 379–393 (2008)
8. Arciero, G.: Sulle Tracce di Sé. Bollati Boringhieri, Torino (2006)
9. Arciero, G., Bondolfi, G.: Selfhood, Identity and Personality Styles. Wiley BlackWell, Hoboken (2011)
10. Guidano, V.F.: Le Dimensioni del Sé. Una lezione sugli sviluppi del modello post-razionalista. Alpes Italia, Roma (2010)
11. Nardi, B.: CostruirSi. Sviluppo e Adattamento del Sé nella Normalità e nella Patologia. Franco Angeli, Milano (2007)
12. Arciero, G., Gaetano, P., Maselli, P., Gentili, N.: Identity, Personality and Emotional Regulations. In: Freeman, A., Mahoney, M.J., Devito, P. (eds.) Cognition and Psychotherapy, 2nd edn., ch. 12, pp. 261–269. Springer, Heidelberg (2004)
13. Gross, J.J.: Emotion Regulation in Adulthood: Timing is Everything. Curr. Dir. Psychol. Sci. 10, 214–219 (2001)
14. Arciero, G., Gaetano, P., Maselli, P., Mazzola, V.: Le Organizzazioni di Significato Personale. In: Bara, B. (ed.) Nuovo Manuale di Psicoterapia Cognitiva, vol. 1, pp. 17–38. Bollati Bolinghieri, Torino (2005)
15. Russell, J., Bachorowski, J.A., Fernandez-Dols, J.M.: Facial and Vocal Expressions of Emotions. Annu. Rev. Psychol. 54, 329–349 (2002)
16. Brennan, K.A., Clark, C.L., Shaver, P.R.: Self Report Measurement of Adult Attachment: An Integrative Overview. In: Simpson, J.A., Rholes, W.S. (eds.) Attachment Theory and Close Relationships, pp. 46–76. Guilford Press, New York (1998)
17. Picardi, A., Vermigli, P., Toni, A., D'Amico, R., Bitetti, D., Pasquini, P.: Evidence of the Validity of the Italian Version of the Questionnaire "Experiences in Close Relationships" (ECR), a Self-Report Instrument to Assess Adult Attachment. Ital. J. Psychopathol. 8, 282–294 (2002)
18. Picardi, A., Bitetti, D., Puddu, P., Pasquini, P.: La scala Experiences in Close Relationships (ECL), un Nuovo Strumento per la Valutazione dell'Attaccamento negli Adulti: Traduzione, Adattamento, e Validazione della Versione Italiana. Riv. Psichiatr. 3, 114–120 (2000)

19. Picardi, A.: First Steps in the Assessment of Cognitive-Emotional Organization within the Framework of Guidano's Model of the Self. Psychother. Psychosom. 72, 363–365 (2003)
20. Ekman, P., Friesen, W.V., Hager, J.C.: The Facial Action Coding System, 2nd edn. Weidenfeld & Nicolson, Salt Lake City (2002)
21. Esposito, A.: The Perceptual and Cognitive Role of Visual and Auditory Channels in Conveying Emotional Information. Cogn. Comp. 1, 268–278 (2009)
22. Esposito, A., Riviello, M.T.: The New Italian Audio and Video Emotional Database. In: Esposito, A., Campbell, N., Vogel, C., Hussain, A., Nijholt, A. (eds.) COST 2102. LNCS, vol. 5967, pp. 406–422. Springer, Heidelberg (2010)

Cogito Ergo Gusto: Explicit and Implicit Determinants of the First Tasting Behaviour

Vincenzo Paolo Senese, Augusto Gnisci, and Antonio Pace

Second University of Naples, Department of Psychology, Caserta, Italy
{vincenzopaolo.senese,augusto.gnisci,antonio.pace}@unina2.it

Abstract. In this study, we investigated how a new food label forms explicit and implicit attitudes toward a product, and through which processes, these attitudes influence consumer behaviours. To this aim, 215 adults (85% female) implicitly and explicitly evaluated labels representing two products: water and chocolate. The labels were presented either in basic form or as having one of four additional symbols representing, respectively, the origin of the product, the respect of the environment, the wellness information, and the shelf life. Results showed that the additional symbolic information creates more of a negative implicit impression but more of a positive explicit attitude toward the products than the basic label does. Moreover, the analysis showed that for the chocolate only were both implicit and explicit reactions critical in driving the approach behaviour toward that food. The theoretical implications of these results are discussed.

Keywords: Consumer psychology, Implicit evaluations, First impression, SC-IAT, Approach behaviours.

1   Introduction

One of the most relevant problems for food makers is to create a label that immediately and efficiently communicates their products' specific features, driving users to positive approach behaviours and ultimately to purchasing decisions. The choice of the label is even more crucial if the product and brand are new for consumers. Indeed, in this case, the consumer behaviour cannot be driven by previous knowledge or experiences of the product or the brand [1], and so it is mainly determined by the individuals' attitudes toward information presented on the label [2].

The relationship between attitudes and behaviours has been widely investigated. According to the Theory of Planned Behaviour [3], attitudes, subjective norms, and perceived behavioural control influence individuals' intentions and behaviours. A more recent dual process model, the Motivation and Opportunity as DEterminants (MODE) [4], specifies that attitudes can also guide behaviours in a spontaneous and automatic manner, out of individual awareness. Indeed, the MODE model [4] assumes that the brain processes information through two operating systems: the spontaneous and the deliberative. The former is a top-down cognitive process, automatically activated by the memory upon the individual's encountering the attitude

object; the latter is a bottom-up process, based on a cognitive effort to evaluate the target object in order to compare and to adopt different behavioural alternatives. A better understanding of how people implicitly and explicitly evaluate a target would explicate the subsequent decision-making process and help in predicting individuals' behaviours toward an evaluated target [3,4]; see also [5,6,7,8,9]. Several studies have been conducted in different research fields, such as political psychology [10,11], economy, and consumer psychology [1,2],[12,13,14,15]. These studies all show that implicit reactions influence behaviours over and above the explicit attitude.

For example, a study on drinking behaviours showed that implicit attitudes toward some traditional brands (such as Coca Cola or Pepsi) can be used to predict consumers' choices and consequent uses [1]. The brands in the study were well-known and thus the automatic evaluations were probably shaped by the consumers' past experiences. In addition, an investigation into the role of both impulsive and reflective evaluations toward a new clothing brand [2] showed that emotional, implicit, and explicit evaluations (whose effects are mediated by intentions) affect approach behaviours. In the latter case, the automatic evaluation was not a consequence of the experience with the product or brand, but it was tied to graphical and perceptual cues owned by the label. The automatic evaluation was, in fact, a first impression [2],[12].

Given the scarce research on the influence of implicit reactions toward a new label on individual approach behaviours, we conducted an experimental study in which we presented to participants some never-before-seen labels of foods (in our case water and chocolate) and then observed their reactions in terms of implicit and explicit attitudes, intentions to buy the product, interest in it, and tasting behaviours. We chose water and chocolate since they represent two different categories of food: the first is regarded as essential and vital (a primary need), the second as a desire or pleasure (almost secondary). In fact, being thirsty means essentially wanting to satisfy the need of drinking water, while being hungry does not have to do with the will to eat chocolate. The food labels were presented to participants either basic or paired with one of four different symbols representing the geographical origins of the product, the respect for the environment and workers in manufacturing it, wellness, and shelf life.

The first scope of the study was to compare the participants' implicit and explicit reactions to the basic labels with their reactions to the labels presented with one of the four abovementioned additional symbols. We expected differences between the implicit and explicit reactions given that automatic mechanisms work better with few visual cues and less semantic information [1,2,3,4,5,6,7]. The second scope of the present study was to evaluate the role of explicit and implicit attitudes, the intention to buy the food, and the interest in the products in predicting the participants' tasting behaviours. We introduced the interest in the product, since we hypothesized that never-before-seen brands' labels might raise curiosity and interest rather than a firm intention to buy the food. In line with other studies [1,2],[12,13,14,15], we expected a strong and positive effect of the implicit reaction on the tasting behaviour, over and above the explicit reactions. That is, we expected that irrespective of the explicit evaluations, the immediate implicit reaction toward the label would work in a strong and independent way in predicting the tasting behaviour.

2   Method

2.1   Sample

A total of 215 first year undergraduate university students (183 females, 32 males) participated in a within-subject design experiment. Their ages ranged from 18 to 43 years (*M*=19.87, *SD*=3.05). All the participants were tested individually in sessions lasting about 45 minutes each.

2.2   Procedure

At their arrival, each participant was welcomed and given instructions about the experiment, and then we trained them on the meanings of the new labels and the symbolic information. The real experiment consisted of two phases. In the first phase, participants were submitted to a computerized task, implemented with the Inquisit 3.0 software [16].

The computerized task was divided into three trials. The implicit reaction to the label, the explicit attitude toward the label, and the intention to buy the products were collected in the first, second, and third trials, respectively. For each trial, participants were presented six targets: two labels presented alone, one relative to water and one to chocolate, and four labels paired with one symbol each. Both the basic labels and the labels with symbols were new and unknown to each participant. Symbols represented a feature of the manufacturing of the food: the geographical origins (O), the respect of the environment and the workers (R), the wellness information (W), and the shelf life (S) (see Fig. 1). Each participant evaluated two labels with symbolic information for the water and two for the chocolate, for a total of four labels with symbols. Labels and symbols were randomly paired for each participant.

The stimuli (first the basic labels and then the labels with symbols) and the trials (first the implicit measure, then two measures of the explicit attitude, and the intention; see below) were presented in a blocked order to investigate how the basic labels or the labels with symbolic information influenced subjective reactions and/or intentions toward the products, and to avoid any influence of the explicit task on the implicit measures.

In the second phase of the experiment, participants were presented some fliers and samples of the products (water and chocolate) and were invited to pick up fliers and to taste the products. After the second phase, participants were debriefed and thanked.

2.3   Measures

Implicit Reaction toward Labels and Symbols. To evaluate the implicit reactions toward each basic label and each label with a symbol, the Single Category Implicit Association Test (SC-IAT) [2],[17] was administered. Participants completed six different SC-IATs, each one with a different target: the new labels of water and chocolate, two labels with symbols for the water, and two labels with symbols for the chocolate. To administer the SC-IATs, we defined ten different attribute categories

each for the positive (positive, joy, beauty, happy, heaven, present, pleasant, friend, laughing, loving) and the negative (negative, pain, ugly, sad, hell, disaster, unpleasant, enemy, crying, hating) dimensions (Cronbach's αs > .71). For each SC-IAT, we computed a single score that expressed the implicit evaluation of the target: negative values indicated a negative implicit attitude, values around 0 indicated a neutral reaction, and positive values indicated a positive implicit attitude.

Fig. 1. Labels according to "basic" and "with symbolic information" as a function of the target

Explicit Attitude toward Basic Labels and Labels with Symbols. To evaluate the explicit attitudes toward the basic labels and the labels with symbols, a Semantic Differential scale [18] and the Warmth and Competence Rating Scale (WCS; from the Stereotype Content Model [19]) were administered.

For the semantic differential, respondents evaluated each stimulus using fifteen bipolar couples of adjectives on a seven-point scale, from 0 = completely negative to 6 = completely positive (Cronbach's αs > .85). As regards the WCS, participants evaluated the targets using fourteen adjectives on an eleven-point scale, from 0 = "not at all" to 10 = "completely" (Cronbach's αs > .90).

Each participant completed six semantic differentials and six WCSs, with the same stimuli used for the implicit measures (see above for the procedures).

Intention to Buy the Product. At the end of the computerized task, we evaluated the participants' intention to buy the product by an ad hoc five-item scale. The items asked participants if they were interested in knowing which shops sell the product, if they had the intention to buy the product, etc., on an eleven-point rating scale, from 0 = "not at all" to 10 = "completely" (Cronbach's αs > .90). Participants completed this scale six times, each time evaluating a different stimulus.

Interest in the Product and Tasting Behaviour. In the second phase of the experimental procedure, two different participant behaviours were observed: interest and taste.

Given that participants were presented with fliers about the product and invited to pick them up, the interest in the product was measured by their behaviour of taking or not taking the fliers. For tasting behaviour, participants were invited to taste some samples of the products, and their behaviour (tasting or not) was coded by two independent observers (Cohen's ks > .90).

2.4   Data Analysis

Principal-component factor analyses were preliminarily executed on the two explicit attitude scales (Semantic Differential and WCS). In both cases the single factor solution was preferred because it explained more than 50% of the variance; therefore, a single factor score was computed for each dimension and used in the analysis.

To compare reactions toward to the basic labels with reactions to the labels with one of the four symbols, paired sample t-tests (α=.05) were conducted on implicit and explicit attitudes, and the intention to buy the product.

To investigate if the water- or chocolate-tasting behaviours (a dichotomous dependent variable) were predicted by the implicit or explicit reactions toward new food labels, and/or by intention and interest, we executed two four-step hierarchical logistic regressions, one for the water and one for the chocolate. In the first step, the implicit reaction was examined, and then the two explicit evaluations (step 2), the intention (step 3), and the interest (step 4) were added to the model separately. Finally, blocks with two-way, three-way, four-way, and five-way effects were used in testing interactions.

3   Results

3.1   Comparison between Labels According to "Basic Labels" and "Labels with Symbols"

The water t-tests showed that the labels with the symbols elicited different evaluations from the basic labels: a) for the implicit measure, participants evaluated more negatively the labels with symbols "O", "R", and "W" than they did the basic labels; b) for the semantic differential scale, participants evaluated more positively the labels with symbols "O", "R", and "W" than they did the basic labels; c) for the WCS, participants evaluated more positively all the labels with symbols than they did the basic labels; d) for the intention, participants showed a stronger intention to buy the product when the labels were presented with symbols than when the labels were basic (see Table 1).

Table 1. Mean values (SD) and paired t-tests (α = .05) for the water labels, as a function of labels and measures. °"Origin" [O]; "Respect" [R]; "Wellness" [W]; "Shelf life" [S]; [*]$p < .05$; [**]$p < .01$; [***]$p < .001$.

| | Stimulus | | | |
| | Basic Label | Label with [symbol]° | | |
Measures	M (SD)	M (SD)	df	t
Implicit	0.25 (0.40)	[O] 0.15 (0.35)	108	2.21[*]
	0.22 (0.40)	[R] 0.10 (0.35)	104	2.65[**]
	0.27 (0.37)	[W] 0.17 (0.33)	94	2.01[*]
	0.21 (0.36)	[S] 0.15 (0.31)	98	1.22
Explicit (Differential)	46.71 (13.47)	[O] 49.04 (13.10)	113	-3.62[***]
	46.56 (14.14)	[R] 48.80 (13.14)	106	-2.26[*]
	47.59 (13.25)	[W] 51.63 (11.51)	96	-4.95[***]
	45.01 (13.42)	[S] 42.54 (14.01)	103	1.97
Explicit (WCS)	60.28 (35.00)	[O] 73.44 (33.41)	113	-6.23[***]
	62.05 (40.31)	[R] 80.62 (38.56)	106	-6.91[***]
	64.43 (36.39)	[W] 76.96 (35.54)	100	-5.25[***]
	60.62 (38.10)	[S] 68.85 (35.68)	103	-3.13[**]
Intention	21.51 (12.34)	[O] 26.63 (12.30)	111	-6.44[***]
	22.10 (13.74)	[R] 32.63 (12.91)	108	-9.12[***]
	21.33 (12.92)	[W] 30.72 (13.14)	100	-7.61[***]
	21.27 (12.83)	[S] 23.80 (13.31)	103	-2.59[*]

The chocolate t-tests also showed that the labels with symbols elicited different evaluations than did the basic labels: a) for the implicit measure, participants evaluated more negatively the label with the symbol "W" than they did the basic label; b) for the semantic differential scale, participants evaluated more positively the labels with symbols "R" and "W", but more negatively the label with the symbol "S", than they did the basic labels; c) for the WCS, participants evaluated more positively all the labels with symbols than they did the basic labels; d) for the intention, participants showed a stronger intention to buy the product when the symbols "O", "R", and "W" were present on the label than they did for the basic label (see Table 2).

3.2   Relation between Implicit and Explicit Attitudes, Purchasing Intention/Interest, and Taste

Regarding the water, no model showed significant effects or interactions. However, the logistic regression on the chocolate showed significant effects among the predictors, with a final Nagelkerke index of R^2=.16 (see Table 3). Results showed that implicit reaction, semantic differential scale, and interest predicted, in an additive and independent way, the chocolate-tasting behaviour. The final model predicted correctly 84.9% of the non-tasting behaviour and 33.3% of the tasting behaviour. Data showed that positive explicit or implicit attitudes were associated with a higher percentage of tasting behaviours, and showed that the higher the interest, the higher the percentage in the tasting behaviour.

Table 2. Mean values (*SD*) and paired *t-tests* (α = .05) for the chocolate labels, as a function of labels and measures. °"Origin" [O]; "Respect" [R]; "Wellness" [W]; "Shelf life" [S]; $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$.

| Measures | Stimulus | | | |
	Label by basic M(SD)	Label with [symbol]° M(SD)	df	t
Implicit	0.19 (0.37)	[O] 0.11 (0.35)	97	1.41
	0.10 (0.38)	[R] 0.07 (0.37)	104	0.80
	0.16 (0.39)	[W] 0.07 (0.34)	111	2.02*
	0.11 (0.39)	[S] 0.13 (0.37)	108	-0.29
Explicit (Differential)	46.60 (11.71)	[O] 47.65 (11.37)	98	-0.94
	47.66 (10.96)	[R] 50.32 (10.04)	105	-3.22**
	47.03 (11.27)	[W] 49.04 (11.26)	111	-2.49*
	47.93 (10.81)	[S] 45.89 (12.00)	106	3.41**
Explicit (WCS)	61.33 (38.48)	[O] 71.81 (39.12)	98	-4.88***
	64.50 (36.24)	[R] 80.65 (33.26)	104	-6.31***
	59.07 (35.06)	[W] 72.70 (34.35)	111	-4.84***
	65.14 (35.09)	[S] 70.28 (34.74)	108	-2.59*
Intention	27.61 (13.92)	[O] 34.84 (11.09)	98	-6.49***
	26.22 (12.65)	[R] 33.74 (11.00)	105	-7.26***
	26.78 (13.72)	[W] 35.88 (12.15)	111	-8.21***
	27.68 (13.41)	[S] 28.37 (13.17)	108	-0.75

Table 3. Hierarchical logistic regression for the chocolate, with tasting behaviour as dependent variable, and implicit evaluation (step 1), explicit evaluations (step 2), intention (step 3), and interest (step 4) as predictor variables. $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$

Steps	Measures	B	SE	Exp(B)	X^2_{Block}	df	X^2_{Model}	df	R^2
1	Implicit	0.32	0.15	1.38*	4.52*	1	4.52*	1	.03
2	Implicit	0.34	0.16	1.41*	9.96**	2	14.48**	3	.10
	Explicit (Differential)	0.42	0.17	1.52*					
	Explicit (WCS)	0.19	0.16	1.21					
3	Implicit	0.34	0.16	1.40*	0.53	1	15.01**	4	.10
	Explicit (Differential)	0.39	0.18	1.48*					
	Explicit (WCS)	0.17	0.17	1.18					
	Intention	0.12	0.17	1.13					
4	Implicit	0.32	0.16	1.38*	9.39**	1	24.40***	5	.16
	Explicit (Differential)	0.43	0.18	1.54*					
	Explicit (WCS)	0.13	0.17	1.14					
	Intention	0.11	0.17	1.12					
	Interest	0.48	0.16	1.61**					

Regarding the interaction effects, only the block of four-way effects was significant, $X^2(5) = 11.75$, $p < .05$. In particular, the "Implicit × Differential × WCS × Interest" interaction effect was significant, Exp(B) = 2.63, $p < .05$. The four-way effect confirmed that predictors interact in driving tasting decisions, and it showed that the positive effect of the implicit attitude is stronger when there is a negative explicit attitude but a positive interest in the product.

4   Discussion and Conclusions

The aims of the present paper were twofold: to investigate the effect of symbolic information on new food labels on explicit and implicit reactions toward the product, and to examine through which cognitive processes these attitudes influence consumer behaviours.

Results showed that additional symbolic information on labels related to the manufacturing of the product influences both emotional and cognitive reactions toward water and toward chocolate but in opposite ways. The symbols "origin" (only for the water), "respect" (only for the water) and "wellness" (for both the products) were implicitly evaluated in a negative way. However, all of the labels with the considered symbols were more positively evaluated than were the basic labels in an explicit way, and the symbolic information strengthened the consumers' intention to buy the product. A possible interpretation of the contrasting effect of the informative labels on the implicit and explicit reactions is that symbolic information needs a deep semantic processing behind the one devoted to processing the basic label; therefore, they have a positive effect when there is enough time to elaborate on their meaning (i.e., in the explicit evaluations), while they have a negative effect when a rapid evaluation is requested. In sum, informative symbols can be an obstacle when a first, immediate impression of a food is at stake, but they can contribute to a positive evaluation of a food when a more reflexive and prolonged judgment is requested.

Regarding the second aim, the hierarchical logistic regressions confirmed that both first, emotion-based reactions to the labels and the cognition-based attitudes toward these labels are critical in driving the immediate desire to taste [3,4,5,6,7,8,9]; however, this is true for the chocolate and not for the water. That is, the type of product moderates this effect.

In this study we tested the first consuming behaviour toward new water and toward chocolate products. Consuming water is a vital need, so people might not be influenced by label information to decide whether to taste it. People simply taste water when they are thirsty. Indeed, in our study only a small percentage of participants tasted the water (19%). This small variance can also explain why the model was not verified for this product. On the other hand, chocolate is not a vital product. Its use more probably reflects a desire or a pleasure, so the label and its effects on participants (first impression, explicit attitude, and interest) become a critical element in positively or negatively orienting the consumers' approaching behaviour of tasting. Indeed, in our data, a larger portion of participants (36%) tasted the product.

Regarding the processes that drive the first tasting behaviour, the results on the chocolate tests showed that when people see a new product label, the likelihood that they will taste the product is a function of both conscious and unconscious processes. Indeed, if the label is associated with a positive implicit or explicit reaction or with a higher interest in the product, then the percentage of tasting behaviours increases. Moreover, if there is a negative explicit attitude but a positive interest in the product, the implicit attitude becomes even more critical in driving the tasting behaviour.

From a theoretical perspective, our results are in line with the MODE model [4] and confirmed that attitudes also guide behaviours in a spontaneous and automatic manner. Moreover, the results of this study showed that when consumer behaviour is not based on previous knowledge or experience with the product or the brand, the influence of the immediate impulsive or reflective reactions depends on the characteristics of the food, or rather its being a primary or a secondary food. When consuming food corresponds with fulfilling a pleasure, the simple exposure to the product label can influence the implicit and explicit evaluations and the interest in the product. Then, in turn, the subjective reactions can directly influence the tasting behaviour. Interestingly, different from what expected according to the Theory of Planned Behaviour [3,4], the interest in the food, not the intention, was the proximal predictor of the tasting behaviour. A possible explanation of this effect could be that the intention has to be based on some experience with the brand or the product to be a proximal predictor of the tasting behaviour [1]. In our study, participants observed new labels, so they did not have enough time or experience with the brand or the product to consolidate a clear intention toward it. Therefore, we suggest that a new label of an unknown brand might act on the curiosity and the interest more than the intention.

In sum, this experimental study confirms the role of explicit and implicit reactions in orienting short-term tasting behaviours (i.e., when the consumer choice immediately follows the presentation of targets). Further studies with more participants and balanced by sex should replicate our results and investigate explicit and implicit immediate reactions in predicting long-term consumer behaviours.

References

1. Maison, D., Greenwald, A.G., Bruin, R.H.: Predictive Validity of the Implicit Association Test in Studies of Brands, Consumer Attitudes, and Behavior. Journal of Consumer Psychology 14, 405–415 (2004)
2. Di Conza, A., Gnisci, A.: First Impression in Mark Evaluation: Predictive Ability of the SC-IAT. In: Esposito, A., Esposito, A.M., Vinciarelli, A., Hoffmann, R., Müller, V.C. (eds.) COST 2102. LNCS, vol. 7403, pp. 353–364. Springer, Heidelberg (2012)
3. Ajzen, I.: The Theory of Planned Behavior. Organizational Behavioral and Human Decision Processes 50, 179–211 (1991)
4. Fazio, R.H.: Attitudes as Object-Evaluation Associations of Varying Strength. Social Cognition 25, 664–703 (2007)
5. Smith, E.R., DeCoster, J.: Dual Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems. Personality and Social Psychology Review 4, 108–131 (2000)
6. Strack, F., Deutsch, R.: Reflective and Impulsive Determinants of Social Behavior. Personality and Social Psychology Review 8, 220–247 (2004)

7. Gawronski, B., Bodenhausen, G.V.: Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change. Psychological Bulletin 132, 692–731 (2006)
8. Lodge, M., Taber, C.S.: The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis. Political Psychology 26, 455–482 (2005)
9. Greenwald, A.G., Poehlman, T.A., Uhlmann, E., Banaji, M.R.: Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive validity. Journal of Personality and Social Psychology 97, 17–41 (2009)
10. Di Conza, A., Gnisci, A., Perugini, M., Senese, V.P.: Atteggiamento Implicito ed Esplicito e Comportamenti di Voto. Le Europee del 2004 in Italia e le Politiche del 2005 in Inghilterra. Psicologia Sociale 2, 301–329 (2010)
11. Di Conza, A., Gnisci, A., Senese, V.P., Pagano, P., Schiavone, S.: La Partecipazione Politica Modera l'Effetto degli Atteggiamenti Impliciti sul Voto? Uno Studio sulle Elezioni Politiche del 2006 e 2008. Giornale Italiano di Psicologia 3, 627–648 (2011)
12. Shapiro, S.: When an Ad's Influence is Beyond our Conscious Control: Perceptual and Conceptual Fluency Effects Caused by Incidental Ad Exposure. Journal of Consumer Research 26, 16–36 (1999)
13. Maison, D., Greenwald, A.G., Bruin, R.H.: The Implicit Association Test as a Measure of Implicit Consumer Attitudes. Polish Psychological Bulletin 32, 1–9 (2001)
14. Brunel, F.F., Tietje, B.C., Greenwald, A.G.: Is the Implicit Association Test a Valid and Valuable Measure of Implicit Consumer Social Cognition. Journal of Consumer Psychology 14, 385–404 (2004)
15. Isen, A.M., Labroo, A.A., Durlach, P.: An Influence of Product and Brand Name on Positive Affect: Implicit and Explicit Measures. Motivation and Emotion 28, 43–63 (2004)
16. Inquisit 3.0 [Computer software]. Millisecond Software, Seattle (2007)
17. Karpinski, A., Steinman, R.B.: The Single Category Implicit Association Test (SC-IAT) as a Measure of Implicit Consumer Attitudes. European Journal of Social Science 7, 32–42 (2008)
18. Osgood, C.E., Suci, G., Tannenbaum, P.: The Measurement of Meaning. University of Illinois Press, Urbana (1957)
19. Fiske, S.T., Cuddy, A.J.C., Glick, P., Xu, J.: A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow from Perceived Status and Competition. Journal of Personality and Social Psychology 82, 878–902 (2002)

Coordination between Markers, Repairs and Hand Gestures in Political Interviews

Augusto Gnisci, Antonio Pace, and Anastasia Palomba

Second University of Naples, Department of Psychology, Caserta, Italy
{augusto.gnisci,antonio.pace}@unina2.it,
anastasiapalomba@libero.it

Abstract. This study aims at exploring the coordination of linguistic features (markers and repairs) with hand gestures, assuming that this concerns rhythm and cohesion of the discourse, rather than the semantic content. Two independent observers coded a sample of 29 broadcasts "Tribuna Politica" (10h 28m) during the campaign of 2004 European elections in Italy. Descriptive data showed that: meta-textual markers were used more than interactive ones; repairs were equally subdivided into modifications and repetitions; conversational gestures (i.e., rhythmic and cohesive) and adaptors were used more than ideational gestures. Correlations between markers/repairs and gestures revealed that these linguistic features are prevalently associated to conversational gestures and adaptors. Probably, when the conversation is fluid, markers and repairs are coordinated with conversational gestures in order to accompany the flow of speech; when the conversation is blocked, they are coordinated with adaptors in order to re-establish the previous flow.

Keywords: Political speech, Nonverbal behavior, Markers, Repairs, Hand gestures, Television interviews.

1   Introduction

The persuasiveness of a political discourse is paramount in the modern democracy, where the use of an effective speech helps politicians to gain the approval of electorate. The importance of studying the characteristics that make the political speech persuasive is linked to the large diffusion of media. In particular, the television is the tool through which politicians present themselves to their potential electors [1-5].

Several verbal and non-verbal features (for example, rhetoric, intonation, body movements, face expressions, etc.) contribute to make the communication effective during the television interviews. Thus, studying how verbal and non-verbal elements are associated and which functions they accomplish, favors the comprehension of the dynamics that make the political speech persuasive. Above all, the combination of aspects with similar functions improves the power of communication, increasing its impact on audience. For example, hand gestures are closely related to intonation and rhetorical strategies in order to give rise to the audience applause [6].

The present contribution aims at exploring the coordination between verbal and non-verbal elements into the politicians' television performances during the campaign of 2004 European elections in Italy. Specifically, we investigated the association of markers and repairs with hand gestures, in order to verify whether this coordination concerns rhythm and cohesion of the speech, rather than the semantic content.

1.1   Markers, Repairs and Hand Gestures

Verbal and non-verbal elements are often coordinated in speech. Here, we consider three elements with similar functions: markers, repairs and hand gestures.

As regards markers, they are words partly deprived of their original meaning. They favor the correct interpretation of a sentence, without contributing to the propositional meaning of it. Bazzanella [7-8] provided a useful linguistic framework where markers have two general functions, interactive and meta-textual. Interactive markers underline a shared background with the interlocutor, strengthening perception of social cohesion. Meta-textual markers concern the articulation of speech contents. Redeker [9] proposed a similar categorization, distinguishing pragmatic (playing an interpersonal function) and ideational markers (linked to the discourse elaboration).

Interactive markers [7-8] are distinguished into: a) turn management (turn taking, turn maintenance and handover); b) attention requests; c) phatisms and modulation mechanisms. Turn taking signals (e.g., *so, then, therefore, well*) are used to obtain the turn; turn maintenance cues to keep the turn; handover signals to indicate that the turn is finished and a new interlocutor is going to be selected. Attention requests (e.g., *hey, listen, look*) are used to capture attention of the interlocutor. Phatisms (e.g., *friend, my dear, as you know*) have a cohesive function; modulation mechanisms modify the impact of semantic content, for example expressing uncertainty or mitigating disagreement.

Meta-textual markers [7-8] include: a) demarkers; b) focalisers; c) reformulations (paraphrases, corrections and exemplifications). Demarkers (e.g., *in short, in sum, however, nonetheless*) give cohesion to the different parts of speech, indicating a change of topic. Focalisers (e.g., *exactly, so there*) highlight the main elements of a discourse, focusing the interlocutor's attention on them. Paraphrases (e.g., *that is, I mean*) are used to make the correspondence between a first formulation and its reformulation explicit; corrections (e.g., *no, on the contrary*) to substitute an incorrect term; exemplifications (e.g., *let's say, for example*) to make a point of view more explicit.

As regards repairs, they are corrections acted by the speaker to rephrase a problematic piece of the discourse. They can be interpreted as indicators of uncertainty or, on the contrary, used to emphasize a part of the speech. Moreover, they can be started by the speaker (self-inducted) or solicited by the interlocutor (hetero-inducted).

The repairs are distinguished into: 1) modifications, used when the speakers start their turn, but they stop before the conclusion, in order to modify a part of it; 2) repetitions, when some words or a whole sentence are formally repeated.

As regards the non-verbal feature considered here – hand gestures – they are unanimously considered a relevant aspect of the communicative process, in coordination with the verbal elements. Several functions have been traditionally assigned to them, such as giving rhythm and cohesion to the speech, illustrating into the space concepts verbally expressed or satisfying psychological needs [10-13]. Unifying a wide, heterogeneous literature, Maricchiolo and collaborators [13] proposed a classification system based on a single criterion: the presence (or not) of a link with the speech. This taxonomy includes: 1) rhythmic; 2) cohesive; 3) ideational; 4) adaptors.

Rhythmic and cohesive (hank, weaving, star, whirlpool, brush, pincers and pliers) are conversational gestures, referring to the structure of speech and contributing to give coherence and continuity to it [12]. Ideational gestures (emblems and illustrators, the latter subdivided into iconics, metaphorics and deictics) refer to the semantic content of speech. Adaptors are not linked to the speech, neither its structure nor its content, and include: a) hetero-addressed (object, person); b) self-addressed.

1.2   Examples of Coordination between Verbal and Non-verbal Features

Verbal elements often overlap to non-verbal cues with a similar function. In the following excerpt from our corpus data, a "modification" (subtype of the repairs) is coordinated with a "whirlpool" (subtype of the cohesive gestures). This linguistic feature signals an obstacle in the communicative performance and the whirlpool helps to recover continuity of the discourse.

Excerpt from Tribuna Politica (31st May, 2004), Annamaria Baccarelli interviews Giuseppe Pizza (DC-Paese Nuovo).

Pizza: We have declared and still declare our friendship with the U.S., but we cannot agree with this <u>unilatism–unilateralism</u>, mortifying the natural political role of Europe.
           Abbiamo professato e continuiamo a professare l'amicizia con gli Stati Uniti, ma non possiamo essere d'accordo con quest'<u>unilaterismo-unilateralismo</u>, che mortifica il ruolo politico naturale dell'Europa.

In the next excerpt, a "repetition" (the other subtype of repairs) is coordinated with a "metaphoric" gesture (subtype of the ideational gestures). This linguistic feature is often associated with this category of gestures, probably because both aim at underlining specific elements of the speech.

Excerpt from Tribuna Politica (3rd June, 2004), Simona Sala interviews Carlo Fatuzzo (Partito Pensionati).

Fatuzzo: This Maroni reform unfortunately is also a killing-pension reform, meaning that everyone has to work three years longer to have the pension and this is unfair, because who worked and paid have to get back their own money whenever they want <u>and don't-and don't pay</u> when they have to pay and don't have their own money back when they get to the retirement age.
           Questa riforma Maroni è purtroppo anche questa una riforma ammazza pensioni, significa che si deve lavorare tutti tre anni di più per avere la pensione e questo non è giusto perché chi ha lavorato e pagato deve riavere i propri soldi quando vuole <u>e non-e non pagare</u> quando si deve pagare e non prendere i soldi che sono suoi quando è arrivata l'età della pensione.

Verbal and non-verbal elements contribute to give coherence, making the communicative exchange understandable. When the message is meant to be persuasive, they can be coordinated in order to strengthen their peculiar functions, as showed in the examples above.

This study aims at exploring the most frequent patterns of coordination between two linguistic features (markers and repairs) and hand gestures employed by politicians during the interviews preceding 2004 European Elections, in order to verify whether this coordination concerns more with rhythm and cohesion of the discourse, rather than with the semantic content.

2   Method

2.1   Sample

Twenty-nine political interviews broadcasted by "Tribuna Politica" during the Italian campaign for 2004 European elections, from May 21st and June 10th, were video-recorded. The whole duration is 10h 28m, for a total of 358 markers and/or repairs.

2.2   Observation Procedure and Category Systems

Two independent observers first transcribed and then coded the interviews following three category systems (below described): markers, repairs and hand gestures.

Occurrence of markers and repairs was registered, together with the concurrent hand gestures. Therefore, the observer coded the whole markers and repairs in the interviews, but only the gestures executed (or not executed) in co-occurrence with these linguistic features. The aim of the present study, in fact, was to explore the coordination between verbal and non-verbal elements, and not their simple distribution.

Markers (Cohen's k=.90) are: 1) interactive; 2) meta-textual. Interactive markers include: turn management (turn taking, turn maintenance and handover); attention requests; phatisms and modulation mechanisms. Meta-textual markers are: demarkers; focalisers; reformulations (paraphrases, corrections and exemplifications).

Repairs (k=.98) include: 1) modifications; 2) repetitions.

Hand gestures (k=.88) are: 1) rhythmic; 2) cohesive; 3) ideational; 4) adaptors. Cohesive gestures include: hank; weaving; star; whirlpool; brush; pincers; pliers. Ideational gestures include: emblems; illustrators (iconics, metaphorics and deictics). Adaptors are: hetero-addressed (object, person); self-addressed.

2.3 Coding Procedures and Data Analyses

The video-recorded material was observed, transcribed and codified following these three category systems.

Occurrences of markers and repairs, as well as co-occurrences of hand gestures, were reported in a file .doc and then saved in a SDIS format (Sequential Data Interchange Standard; [14]), in order to obtain descriptive data.

Afterwards, this file was imported in SPSS to compute correlations between markers/repairs and concurrent gestures by Pearson's coefficient.

3 Results

3.1 Descriptive Data

We identified 358 linguistic elements: 248 markers (among them, 24% interactive, 76% meta-textual) and 110 repairs, almost equally subdivided into modifications and repetitions (48% and 52%). In Table 1, more details about the distribution of markers subtypes are presented.

Table 1. Distribution of markers

Markers	Frequency	Percentage
Turn Taking	17	6.85%
Turn Maintenance	13	5.24%
Handover	5	2.02%
Attention requests	0	.00%
Phatisms	1	.40%
Modulation mechanisms	24	9.68%
Interactive	**60**	**24.19%**
Demarkers	77	31.05%
Focalisers	46	18.55%
Paraphrases	26	10.48%
Corrections	8	3.23%
Exemplifications	31	12.50%
Meta-textual	**188**	**75.81%**
Total	**248**	**100.00%**

Among the linguistic elements coded, only 10% (36 out of 358) were not accompanied by gestures. Among these 322 concurrent gestures, we found: 36% rhythmic, 31% cohesive, 23% adaptors, and only 6% ideational. In Table 2, more details about the distribution of gestures subtypes are presented.

Table 2. Distribution of concurrent gestures

Hand gestures	Frequency	Percentage
Rhythmic	**117**	**36.34%**
Hank	3	.93%
Weaving	2	.62%
Star	11	3.42%
Whirlpool	13	4.04%
Brush	9	2.80%
Pincers	55	17.08%
Pliers	7	2.17%
Cohesive	**100**	**31.06%**
Emblems	8	2.48%
Iconics	2	.62%
Metaphorics	8	2.48%
Deictics	5	1.55%
Ideational	**23**	**7.14%**
Object-adaptors	58	18.01%
Person-adaptors	1	.31%
Self-adaptors	23	7.14%
Adaptors	**82**	**25.47%**
Total	**322**	**100.00%**

3.2   Coordination between Markers/Repairs and Hand Gestures

Interactive markers correlated significantly with: cohesive gestures ($r=.40$, $p<.05$); adaptors ($r=.39$, $p<.05$).

As regards the subtypes of interactive markers, turn management signals and modulation mechanisms showed some significant correlations. Turn management cues correlated with cohesive gestures ($r=.44$, $p<.05$). Modulation mechanisms correlated with the absence of gestures ($r=.36$, $p=.053$).

Meta-textual markers correlated significantly with: rhythmic gestures ($r=.57$, $p<.01$); cohesive gestures ($r=.73$, $p<.001$), in particular with pliers ($r=.69$, $p<.001$); adaptors ($r=.51$, $p<.01$), especially with object-addressed ($r=.50$, $p<.01$).

As regards the subtypes of meta-textual markers, demarkers, focalisers and reformulations showed some significant correlations. Demarkers correlated with: rhythmic gestures ($r=.48$, $p<.01$); adaptors ($r=.48$, $p<.01$), in particular with object-addressed ($r=.58$, $p<.01$). Focalisers correlated with: rhythmic gestures ($r=.41$, $p<.05$); cohesive gestures ($r=.60$, $p<.001$), especially with pliers ($r=.69$, $p<.001$). Reformulations correlated with: rhythmic gestures ($r=.67$, $p<.001$); cohesive gestures ($r=.67$, $p<.001$), in particular with pliers ($r=.56$, $p<.01$); adaptors ($r=.43$, $p<.05$).

Modifications correlated significantly with: rhythmic gestures ($r=.60$, $p<.01$); cohesive gestures ($r=.48$, $p<.01$), in particular with pliers ($r=.38$, $p<.05$).

Repetitions correlated significantly with: rhythmic gestures ($r=.48$, $p<.01$); ideational gestures ($r=.40$, $p<.05$), in particular with metaphorics; adaptors ($r=.67$, $p<.001$), especially with object- ($r=.58$, $p<.01$) and self-addressed ($r=.59$, $p<.01$).

4   Discussion and Conclusions

The present contribution aimed at exploring the coordination between two verbal features (markers, repairs) and one non-verbal element (hand gestures) into the politicians' television performances during the campaign of 2004 European elections.

As regards the descriptive data about linguistic features, repairs were almost equally subdivided into modifications and repetitions, whereas meta-textual markers were used much more than interactive ones.

Therefore, in our sample the majority of markers did not concern relational aspects, but the structuring of speech (i.e., the way through which the contents are expressed). It should not surprise that relational markers are less used in the context of a broadcast like "Tribuna Politica", where several rules regulate the interactive exchange: for example, times are fixed and equally distributed, and alternation of turns almost prefixed.

As regards the descriptive data about hand gestures concurrent to these linguistic features, conversational gestures (i.e., rhythmic and cohesive) were used more than adaptors, and even more than ideational gestures (only 6%). The percentage of ideational gestures is very low if compared with that of studies in which gestures were coded with no regard to coordination with linguistic features. In these studies, in fact, the distribution shows a frequency of ideational gestures above 50% [15] or at least 23% [16].

Considering that rhythmic and cohesive gestures are the most frequent among the gestures concurrent to markers and repairs, probably these linguistic features tend to be used more with a conversational aim, which regard essentially the development of a well-structured speech (like rhythmic and cohesive gestures), and less with the aim to clarify the semantic content (like ideational gestures). In addition, the most frequent concurrent gestures, after the conversational ones, were the adaptors. If rhythmic and cohesive gestures give rhythm and cohesion to the discourse [12], adaptors manifest presumably uncertainty, obstacles and problems [17,18].

As regards the correlations between markers/repairs and hand gestures, the specific coordination patterns provide further insights regarding the use of these verbal and non-verbal elements into the political discourse.

For interactive markers, the modulation mechanisms were the unique signals associated with the absence of gestures. Their modal use has probably enough strength also without the reinforcement of a gesture.

For meta-textual markers, both demarkers and focalisers were coordinated with the rhythmic gestures (which give emphasis on a specific point of the speech), whereas only the focalisers were associated with the cohesive gestures (which highlight an important point of the discourse). Presumably, the tendency of verbal (demarkers, focalisers) and non-verbal elements (rhythmic and cohesive gestures) to be used

contemporarily, confirm that the combined use of aspects with similar functions might improve the power of communication [6]. The association between reformulations and adaptors (which express uncertainty), instead, shows that markers might be used also when the conversation is temporarily blocked.

For repairs, both the subtypes were coordinated with rhythmic gestures. However, modifications were also correlated with cohesive gestures; repetitions with adaptors and ideational gestures (one of the few correlations found for this category).

In sum, the values of correlations indicate that markers and repairs seem to be coordinated more with conversational gestures (rhythmic, cohesive) and adaptors, and much less with ideational gestures. In particular, when the conversation is fluid, markers and repairs might be coordinated with rhythmic and cohesive gestures in order to accompany the flow of speech; when the conversation is blocked, markers and repairs might be coordinated with adaptors to restart the previous flow.

In conclusion, the patterns arisen from the coordination between linguistic features of the political discourse and non-verbal behaviors show that the politicians pay great attention to the way through which the content is structured. Their aim is to build a fluid speech, with basic points to be marked and focused and other moments in which re-establishing the previous flow, overcoming difficulties typical of the public improvised spoken speech. Although we regard these conclusions as coherent and fruitful, we should note as a final comment that they are based on a descriptive and correlational study. Future studies should therefore deepen the matter of the coordination between verbal and nonverbal features.

References

1. Gnisci, A., Bonaiuto, M.: Grilling Politicians. A Study on Politicians' Answers to Questions Comparing Televised Political Interviews and Legal Examinations. Journal of Language and Social Psychology 29, 384–413 (2003)
2. Gnisci, A., Di Conza, A., Zollo, P.: Political Journalism as a Democracy Watchman. In: Herrmann, P. (ed.) Democracy in Theory and Action, pp. 205–230. NOVA Publishers, New York (2011)
3. Gnisci, A., Zollo, P., Perugini, M., Di Conza, A.: A Comparative Study of Toughness and Neutrality in Italian and English Political Interviews. Journal of Pragmatics 50, 152–167 (2013)
4. Gnisci, A., Van Dalen, A., Di Conza, A.: Interviews in a Polarized Television Market: The Anglo-American Watchdog Model Put to the Test. Political Communication 31, 112–130 (2014)
5. Gnisci, A., Bull, P., Graziano, E., Ciancia, M.R., Errico, D.: Un Sistema di Codifica delle Interruzioni per l'Intervista Politica Italiana 1, 107–128 (2011)
6. Bull, P.: The Use of Hand Gesture in Political Speeches: Some Cases Studies. Journal of Language and Social Psychology 5, 103–118 (1986)
7. Bazzanella, C.: Le Facce del Parlare. Un Approccio Pragmatico all'Italiano Parlato. La Nuova Italia, Firenze (1994)
8. Bazzanella, C.: Linguistica e Pragmatica del Linguaggio. Un'introduzione. Laterza, Roma-Bari (2008)
9. Redeker, G.: Linguistic Markers of Discourse Structure. Linguistics 29, 1139–1172 (1991)

10. Ekman, P., Friesen, W.V.: The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. Semiotica 1, 49–98 (1969)
11. Kendon, A.: Some Uses of Gestures. In: Tannen, D., Saville-Troike, M. (eds.) Perspective on Silence, pp. 215–234. Ablex, Norwood (1985)
12. Contento, S.: Sincronismo Verbale-Gestuale in Sequenze Conversazionali. QVR 8, 89–98 (1996)
13. Maricchiolo, F., Gnisci, A., Bonaiuto, M.: Coding Hand Gestures: A Reliable Taxonomy and a Multi-media Support. In: Esposito, A., Esposito, A.M., Vinciarelli, A., Hoffmann, R., Müller, V.C. (eds.) COST 2102. LNCS, vol. 7403, pp. 405–416. Springer, Heidelberg (2012)
14. Bakeman, R., Quera, V.: Analyzing Interaction. Sequential Analysis with SDIS and GSEQ. Cambridge University Press, New York (1995)
15. Maricchiolo, F., Bonaiuto, M.: Lo Stile Comunicativo dei Leader Politici: Analisi di Alcuni Parametri Verbali e Non Verbali nelle Interviste Televisive Durante la Campagna Elettorale delle Elezioni Politiche 2001. In: Sensales, G. (ed.) Rappresentazioni della "Politica". Ricerche in Psicologia Sociale della Politica, pp. 189–208. Franco Angeli, Milano (2005)
16. Maricchiolo, F., Gnisci, A., Bonaiuto, M.: Political Leaders' Communicative Style and Audience Evaluation in an Italian General Election Debate. In: Poggi, I., D'Errico, F., Vincze, L., Vinciarelli, A. (eds.) PoIitical Speech 2010. LNCS, vol. 7688, pp. 114–132. Springer, Heidelberg (2013)
17. Maricchiolo, F., Gnisci, A., Bonaiuto, M., Ficca, G.: Effects of Different Types of Hand Gestures in Persuasive Speech on Receivers' Evaluations. Language of Cognitive Processes 24, 239–266 (2009)
18. Gnisci, A., Pace, A.: The Effects of Hand Gestures on Psychosocial Perception: A Preliminary Study. In: Bassis, S., Esposito, A., Morabito, F.C. (eds.) Recent Advances of Neural Networks Models and Applications. SIST, vol. 26, pp. 305–314. Springer, Heidelberg (2014)

Making Decisions under Uncertainty
Emotions, *Risk* and *Biases*

Mauro Maldonato and Silvia Dell'Orco

University of Basilicata, Potenza, Italy
`{m.maldonato,silviadellorco}@gmail.com`

Abstract. The difficulty in deciding and facing up to uncertainty is not only linked to the inadequacy of the architecture of our minds but also to an 'external' model of uncertainty which does not correspond to the way in which our mind naturally functions. New conceptual paradigms and new programmes for experimental research are called for in order to redefine the role of internal and external restrictions on human action (resources and available information, limitations on calculation ability, on the capacity of memory, cognitive styles, gender differences and so on). All this should be contemplated in a more general theoretical framework – natural logic – based not on metaphysical assumptions but on the concrete evidence provided by cognitive neurosciences.

Keywords: Decision-making, risk, gender differences, biases, uncertainty.

1 Introduction

During the 20th century economists and mathematicians went to great lengths to neutralise risk and associated concepts such as uncertainty and unforeseeability. The demonstration of the limits of the neoclassical paradigm based on the simple calculation of costs and benefits made it more difficult to arrive at a scientific evaluation of risk and uncertainty. From the seventies onwards a large quantity of theoretical and empirical studies have investigated the heuristic principles and cognitive strategies which individuals use to deal with risky and uncertain situations. This research has shown how the explicative and predictive shortcomings of normative risk analysis depend in many respects on undervaluing the continuous interaction between the individual and the environment. These are factors which day by day represent significant obstacles in decision making (1). Conventionally, when one speaks of uncertainty one refers to situations in which the individual knows the outcomes of the choice but not the probabilities involved. The problem of uncertainty is central to the study of decision-making processes because the consequences of the actions an individual undertakes are often prolonged into the future, and one can never be completely sure that the hoped-for outcome will in fact be achieved. Although uncertainty is a key concept in discussions of decision making, there is no real consensus of opinion as to its meaning. One can find as many definitions of it as there are ways of

approaching it (2). In order to clarify the nature of the uncertainty, (3) identified three basic situations:

1. uncertainty is the sense of doubt that blocks or delays action. We can identify three essential features in this definition: 1) it is subjective (different people can be subject to different doubts in identical situations); 2) it is inclusive (no particular form of doubt, such as ignorance of future results, is specified); 3) it conceptualises uncertainty in terms of its effect on action (hesitation, indecision, procrastination).
2. the uncertainty with which decision makers must cope depends on the model of decision making adopted. In other words, models implemented which have different informational requisites will be blocked or delayed by different doubts;
3. different types of uncertainty can be classified according to their issue (what the decision maker is unsure about) and source (what determines the uncertainty). The fundamental problems include results, situations and alternatives. As for the causes, incomplete information is the most commonly cited cause of uncertainty. On occasions, however, decision makers are incapable of acting not so much out of lack of information but because they are disoriented by conflicts generated by the surfeit of meanings the information gives rise to. Moreover the causes of uncertainty are not limited to incomplete information and inadequate comprehension. Decision makers may be prevented from acting even if they have understood the alternatives perfectly but are unable to differentiate between them.

2   Risk and Biases

The concept of uncertainty goes hand in hand with that of risk: a risky situation is always determined by a certain degree of uncertainty concerning the results of future actions.The first scientific study on the perception of risk was carried out by the director of the Atomic International Division, Starr. Published in "Science" in 1969 (4), it looked at safety in nuclear power stations and proposed a procedure for calculating the level of technological risk acceptable to society in view of the attendant social benefits. Even though it relied on a mathematical/probabilistic evaluation of risk, the results revealed an enormous discrepancy between the objective risk and the perception on the part of the population. The variant of "social acceptance" soon proved to be complex, eluding concrete estimates and classifications, leading researchers to talk about different levels of risk. In particular, it was shown that the risks perceived as voluntary (such as risks associated with smoking or the failure to prevent certain illnesses) were considered more acceptable and less probable than the risks perceived as involuntary or imposed (as for example those of nuclear power stations). Moreover, as is shown by Starr's correlation function (1969), when the events are very familiar, objective and perceived risk coincide; as they become less frequent, the perceived risk increases unduly; and finally, in cases of extreme rarity, it diminishes unduly.

Subsequently Starr's research discussion of risk, which had been restricted to the sphere of technological safety, spread to such sectors as psychology and sociology.

Psychology has contributed considerably to risk analysis, progressing from the classic concept of the calculation of probabilities of an undesirable event to the concept of subjective risk based on perception and individual evaluation. In this line of research the most commonly used methodology is known as the psychometric paradigm, proposed by Slovic and his group. The main aim of this research is to identify the mental strategies people use in formulating risk assessments. According to Kahneman, Slovic and Tversky (5), heuristic judgement often constitutes the only practical way to evaluate uncertain elements. In fact, unlike what happens in formal calculus, heuristic evaluation of probability is generally based on immediate solutions which do not consider all the factors at stake, but only the peculiar features of the object being evaluated, the way in which the problem has been formulated, the clarity with which the situation has been described, the degree of control, the seriousness of the consequences, previous knowledge and experiences and so on. These factors, whether separately or in conjunction, influence decision-making behaviour and can easily lead to distortions of judgement or biases (6). Among others, particular attention has been paid to the following phenomena:

Confirmation bias
In interpreting events there is a general tendency to attribute little importance to the contradictory information or else to only contemplate events which are coherent with one's expectations. We often appear to base our judgement on information that confirms our hypotheses rather than the contrary. These affirmations are borne out by the well known experiment devised by Wason (7) known as the four-card selection task. Participants were shown the following four cards. Each card has a letter on one side and a number on the other. Participants were informed of the rule that if a card has a vowel on one side, it must have an even number on the other. The task is to say how many cards need to be turned over in order to verify whether this rule holds good or not. The correct answer, rarely given, is to turn over only two cards: card E and card 7. In fact, if on the back of E there is an odd number, the rule is false; if on the back of 7 there is a vowel, again the rule is false. In other words, any card with a vowel on one side and an odd number on the other breaks the rule. On the contrary, opting for card 4 and card B, the ones most often chosen by participants, is pointless because the rule states "if there is a vowel then there is an even number" and not "only if" there is a vowel, so that on the other side of 4 there could be either a vowel or a consonant, just as on the other side of B there could be either an odd or an even number. This experiment demonstrates the tendency, very common in inferential tasks, to acquire information which only goes to confirm a hypothesis, without checking the falsifying cases.

Consent heuristic
The consent heuristic is a cognitive strategy based on the observation that when a reasonably large number of people reach a consensus on the assessment of an event, individuals taken one by one undergo a sort of psychological pressure and tend to adopt the common point of view in a reaction which is gregarious and conformist but rationally inexplicable. This cognitive strategy is adopted more commonly if the subject

is unfamiliar or there is low motivation or limited possibility of processing the information. In the specific context of risk behaviour it has been observed that, if we are given information on the preventive behaviour of others, this modifies our intentions concerning the use of safety measures in a directly proportional ratio. In a classic experiment, the psychologist Asch (8) asked a group of participants to state whether two lines were the same length. All the participants were his accomplices apart from one. Asch discovered that it was enough for the other participants to answer in a certain way, even if it was patently wrong, for the judgement of the individual whose behaviour he was studying to be influenced.

Illusion of control

The illusion of control is defined «as an expectancy of a personal success probability inappropriately higher than the objective probability would warrant» (9). In other words, people tend to believe that the risks inherent in such behaviour, such as driving at high speed, can be controlled by their own ability. This betrays an excessive and unjustified belief in oneself (overconfidence), since even an expert driver cannot control all the factors which contribute to causing a road accident. A series of studies was conducted to elucidate this phenomenon. A common example is smoking. Those who smoke, in fact, believe they can control their behaviour more than is actually the case in real life. Several studies have shown that among occasional smokers only 15% believed that over the next 5 years they would become heavy smokers. In reality, 5 years on about 43% of them had done so, showing a significant over-estimation of their ability to control events. Among heavy smokers, on the other hand, about 32% believed that over the next 5 years they would still be smoking, and 68% thought they would have given up. In reality 5 years on 70% continued to smoke.

Unrealistic optimism

Unrealistic optimism is closely connected to the illusion of control. It represents the difference between what we consider risky for ourselves and what we consider risky for others (10). Numerous experiments have shown that this bias derives from two dynamics. The first (cognitive) consists in overestimating the number and efficacy of the precautions you yourself take with compared to those taken by others. The reason is that one's own behaviour is more readily accessible in one's memory than that of others with the consequence that the evaluation is distorted by a recollection that favours oneself. The second dynamic (motivational) shows how the individual also uses optimistic distortions to safeguard self esteem. If there were no such distortions, in fact, we would perceive the risks inherent in consciously dangerous activities – such as smoking or driving without a seatbelt – and this would reflect badly on our self image. Interestingly, in some conditions not only does the optimistic bias disappear but it is replaced by pessimistic bias (11), a tendency which is apparently correlated to the nature of the risk. If, in fact, the optimistic bias characterises risks which are incidental, potential and familiar, pessimistic bias corresponds to risks perceived as common, real and unfamiliar (for example the health effects linked to radiation following a nuclear accident). In terms of adaptation, in fact, in the first case an optimistic attitude can free us from anxiety and help us to cope more serenely with everyday activities; while in the second case one is induced to pay more attention to the risks.

Expertise

Numerous experiments show that the level of expertise - where the term refers not to actual experiences of dangerous situations, but to the competence of individuals acquired during their professional activity - generally influences risk evaluation. For example, the studies carried out by Klein and colleagues (12) on decision making by experts (doctors, fire fighters, pilots and others) have shown how in critical situations they tend not to follow normative models, but they "photograph" the current situation and act on the grounds of intuition deriving from past experience. A significant example is provided by the so-called circumstantial paradigm typical of medical semiotics. It is based not on analytical reasoning but on an intuitive activity enabling the medical expert to diagnose pathologies which are inaccessible to direct observation on the basis of superficial symptoms that are insignificant to the untrained eye. It is one of the gifts of the expert: being able to make a correct diagnosis at a glance, in next to no time and with very few elements to go on. In this type of knowledge there are imponderable elements which come into play: flair, instinct, intuition. Some elements only reveal themselves to a scrupulous, practised observer, endowed with that "third eye" which is sometimes called a "clinical gaze" and which is developed in the course of time, through experience. Moreover numerous experiments have shown that – even in the presence of high quality scientific information and data like those provided by the EBM – in many routine clinical decisions (for example the interpretation of a diagnostic test, the choice between different therapeutic options, the identification of a patient's preference, and so on) cognitive errors are commonly made. Furthermore, even when both the exact percentage of error for a certain diagnostic test and the general frequency of an illness are known, doctors are often unable to infer the probability that a patient showing a positive outcome from a certain test actually has that illness. Gigerenzer (13) has labelled "statistical illiteracy" the inability to interpret probabilistic problems and draw inferences based on Bayesian calculus.

3   Risk Assessment and Emotions

The emotions represent an important system of monitoring for relations between the individual and the environment because they pinpoint situations that regard us directly, highlighting what is at stake and which resources we can call on in order to modify these situations. The emotions fulfill both a communicative and a motivational function. In the former the subject is rapidly alerted to the situation with respect to his/her needs and goals, showing third parties, by means of non verbal language, the affective reaction in progress; the latter consists in preparing the organism to react to the emotive situation, adopting appropriate modes of behaviour, which may involve inaction or the rejection of inter-relations, as when somebody is feeling demoralised. Of relevance, in this sense, is the notion of 'risk as feelings' (14), which refers to our fast, instinctive, and intuitive reactions to danger. In other words, the choices made in situations of risk are in part the result of the direct influence of the emotive reactions on the cognitive process. The studies carried out by Loewenstein suggest that in conditions of risk, emotive and rational reactions can diverge on account of

risk assessment. Nonetheless judgement is often determined by the former rather than the latter. In a state of anger, for example – as demonstrated by Lerner and Keltner (15) – angry people express more optimistic risk assessments and manifest risk-seeking behaviour. This conclusion is coherent with Lerner and Keltner's theory of assessment, whereby anger is associated with the perception of greater certainty and control over the outcome of one's behaviour and decisions. On the contrary, sadness seems to be characterised by a lack of physiological excitation and thus scarce propensity to action, associated with a sense of resignation and impotence. This sensation reduces risk aversion, and the consequences of one's decisions are often attributed to the situation rather than to personal factors. Fear and anxiety, while not being synonyms – fear in fact refers to knowable causes, while in anxiety the threat is represented by uncertainty regarding future states or situations concerning individual well-being – produce a common impulse concerning action: evasion or flight. In a state of anxiety there is no concrete threat prompting evasion or flight. The behavioural correlates of anxiety are more common, and the effects on behaviour more pervasive and long-lasting (16). Fear and anxiety derive from assessments of uncertainty and lack of control over the situation. Unlike anger, they are associated with pessimistic evaluations of the environmental conditions. Thus people manifest a contrary impulse to action: instead of being optimistic with respect to risk, they display risk aversion and a pessimistic assessment of the situation. A manager prone to fear or anxiety, for example, is likely to pay more attention to their own behaviour and arrive at a negative risk assessment (17).

There is ample evidence to indicate that joy and happiness favour a sociable, cooperative attitude towards others, reducing interpersonal conflicts. Happiness induces a sense of security and control in people's perception of the environment, making them more ready to adopt risky decisions.

4   Gender Differences and Brain

Right from infancy, hormones such as estrogens and testosterone play a part in the development of the brain, highlighting the differences between the genders. From the outset, the study of faces and the immediate environment models and moulds the cerebral development of males and females alike. Numerous studies show that through eye contact and the observation of faces, the skills of female infants within the first three months are much more developed than those of males. Furthermore, in subsequent phases of development, females tend, for example, to look at their mothers' faces, seeking signs of approval or disapproval, 10 to 20 times more than males in the control group (18). Similarly, during puberty, estrogens and testosterone continue to influence development. If estrogens drive male adolescents to exert more energy in building relationships and in competing for sex, testosterone in boys gives rise to a tendency towards solitude. In fact, testosterone reduces their desire to socialize, except when searching for sex, sport, independence-related challenges and authoritativeness through competitive behaviour. These behaviour models will influence men and women throughout adulthood.

On the evolutionary level, consequently, the female brain is supposedly "programmed" in order to maintain social harmony, while for males it serves to compete, reproduce and transmit its own genes. This dynamic is the basis of our social system: males engaged in competition to become fathers, females in tasks related to child rearing.

The current problem is that while it may be true that our society has radically changed, our brain is still controlled by the same hormonal mechanisms. Deborah Tannen (19) has shown how businesswomen in western cultures still search for eye contact and look into people's faces for approval or disapproval. Men often interpret this behaviour as a sign of insecurity more than as the ability to observe and assess. In the workplace men, due to their tendency to compete, show little inclination to regard a woman as a leader, especially if she is not fiercely competitive. However, for a woman the psychological stress originating from a situation of conflict is very deep-seated, and therefore, it is not surprising that even the most competitive business-woman tends not to attack others or to engage in shouting matches. One of the main reasons why women leave organizations is that they do not wish to involve themselves in political power struggles, because they experience these as a sheer waste of energy.

One of the structural differences between the male and female brain is the size of the amygdale: in the adult, in fact, that of the male is bigger than that of female (18). In women, the amygdale, which is smaller, but better connected to the cortex, is a key factor in the decision-making process, giving greater emphasis to the emotional co potent than is the case with men. (20). The latter, as a matter of fact, tend to process surrounding reality by making particular use of the rational, logical and linear left hemisphere. Women, on the other hand, use mainly the right hemisphere for multi-tasking operations, underlining their strong intuitive faculties. Interestingly, the regions of the brain which differ in size between men and women are the same as those which contain high concentrations of sexual hormone receptors, all of which goes to prove the importance of identifying the sizes of specific regions of the brain during development.

The corpus callosum (a thick structure made up of myelinated fibres joining both cerebral hemispheres) is denser in females, permitting them to use both hemispheres in a more integrated fashion than men (21). In the case of women, in particular, these junctures allow them to express their emotions more effectively, remember details of emotive events and to communicate them: here the hippocampus is decisive – a structure strongly involved in learning, memory and the emotions – larger and more active in the female brain (particularly sensitive to estrogens) (18).

In the sphere of the decision-making process it is interesting to note how the pre-frontal cortex, the brain's 'work space' devoted to decision-making, is larger in women and matures more rapidly than in men. This difference, combined with the fact that women have lower levels of testosterone and higher levels of estrogens, allows the latter to seek solutions to conflict, often causing them to stand back in order that the situation may be resolved. Men, on the other hand, tend to emerge as winners.

The anterior cingulate cortex, another important part of the brain's decision-making workspace which weighs up options, is also larger in women, and has been

defined as the "apprehension centre" of the feminine brain. Numerous studies, in fact, show that anxiety is four times more common in women than it is in men and this leads them to be extremely cautious and collaborative, especially with regard to defending their young ones. This caution today, especially in the workplace and the business world, may be interpreted by men as an indicator of insecurity when it comes to taking on and assessing uncertain and risky situations.

5   Men vs Women: Cognitive and Decision-Making Styles

Individuals do not possess uniform and stable cognitive faculties within time. The current scientific debate tends to consider individual differences in terms of cognitive styles, than cognitive ability. In the realm of studies on decision-making it has become apparent that in most situations and problems related to our day to day life, at any level of complexity, there is no single solution. In fact, given the same situation, different people will act in different ways and adopt different cognitive strategies. But how is one to decide? What are the neural correlates pertaining to differences in cognitive and decision-making styles? In a study conducted by Podell and coll. (22) such issues were addressed by designing the so-called Cognitive Bias Task (CBT). One group of healthy subjects were shown stimuli depicting various simple geometric designs differentiated by colour (red/blue), outline (indicated and coloured in), number (one/two), shape (circular/square) and size (big/small), for a total of 32 possible combinations. Each test consisted of three stimuli: one target stimulus and two possible choices aligned vertically beneath it (Fig. 1). The target figure was presented for two seconds, followed by the simultaneous presentation of two possible choices. The subject, seated in front of a computer was asked to look at a target card and then to select one of two alternatives. The subjects were exposed to different models of response which centred on distinct strategies.
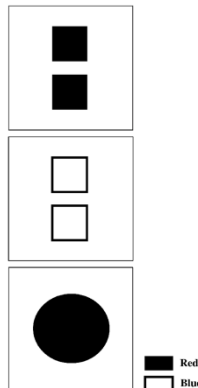
Fig. 1. In the Cognitive Bias Task, a subject is instructed to look at the target form at the top (in this case, a filled-in red circle) and is then asked either to choose the bottom form that is most similar to it (or most different from it) or to choose the bottom form he or she likes best. (From: Goldberg, 2001)

Certain subjects tend to link their choice to the target and when this changed, so did their preference. Such a decision-making strategy is called contest-dependent. Other subjects, on the other hand, tended to make a decision based on stable preferences, irrespective of the target: that is, they always chose blue, red, a circle or a square. This last decision-making strategy is instead known as contest-dependent. An interesting aspect is that males and females presented their choices in surprisingly different ways: males were more dependent on contest than were females (Fig. 2).
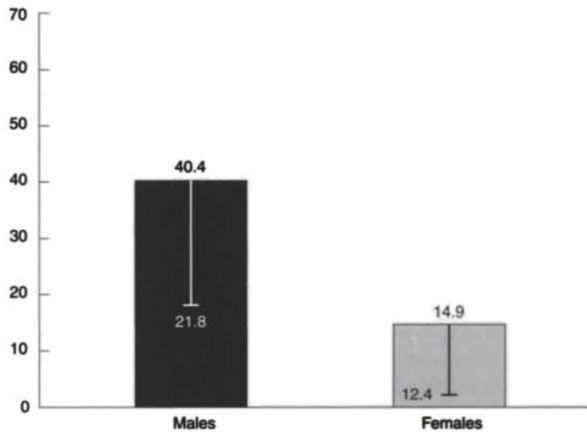
Fig. 2. Sex differences in actor-centered decision-making. Males exhibit a more context-dependent response selection pattern on the Cognitive Bias Task. Females exhibit a more context independent response selection pattern (23).

In the CBT experiment Goldberg (23) examined whether the observed differences between the genders might correspond to real life situations. The contest dependent strategy may be considered a universal default strategy, an attempt on the part of the individual to formulate the best answers in all possible real life situations. The organism accumulates a repertoire of responses corresponding to the sum total of its own life experiences that are slowly but surely updated with new experiences. The problem with such a strategy is that often real-world situations are so different from one another that any attempt to adapt old strategies to new problems becomes meaningless. Nevertheless, such a default strategy may represent the best solution when one is confronted with a totally new situation, for which there is no specific experience or knowledge with which to deal with it. On the other hand, a contest-dependent strategy reflects the propensity to capture the specific properties of the situation and personalise the individual response. Confronted by another situation the organism attempts to recognize it as a familiar model, with known features. However, confronted by a situation which is completely new, such an attempt can only have a negative outcome. In this case, an organism guided by a contest related strategy will seek to capture the properties unique to the situation, even though the available information may be insufficient. The optimum decision-making strategy is probably reached by means of a dynamic balance between the two approaches. Few people, in fact, adhere to either strategy in its pure form, but rather adopt a mixture of strategies, depending on the situation.

References

1. Thompson, J.: Organizations in Action. McGraw Hill, New York (1967)
2. Argote, L.: Input uncertainty and organizational coordination in hospital emergency units. Administrative Science Quarterly 27, 420–434 (1982)
3. Lipshitz, R., Strauss, O.: Coping with uncertainty: A naturalistic decision making analysis. Organizational Behavior and Human Decision Processes 69(2), 149–163 (1997)
4. Starr, C.: Social benefit versus technological risk. What is our society willing to pay for safety? Science 165(19), 1232–1238 (1969)
5. Kahneman, D., Slovic, P., Tversky, A.: Judgment under Uncertainty: Heuristics and biases. Cambridge University Press, New York (1982)
6. Maldonato, M., Dell'Orco, S.: Toward an Evolutionary Theory of Rationality. World Futures 66(2), 103–123 (2010)
7. Wason, P.C.: On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology 12(3), 129–140 (1960)
8. Asch, S.: Social Psychology. Prentice-Hall, New York (1952)
9. Langer, E.J.: The illusion of control. Journal of Personality and Social Psychology 32(2), 311–328 (1975)
10. Slovic, P., Fischhoff, B., Lichtenstein, S.: Facts and fears: Understanding perceived risk. In: Schwing, R., Albers Jr., W.A. (eds.) Societal Risk Assessment: How Safe is Safe Enough?. Plenum Press, New York (1980)
11. Dolinsky, D., Gromsky, W., Zawinsza, E.: Unrealistic pessimism. Journal of Social Psychology 127(5), 511–516 (1986)
12. Klein, G.A.: Recognition-primed decisions. In: Rouse, W.B. (ed.) Advances in Man-Machine Systems Research. JAI Press, Greenwich (1993)
13. Gigerenzer, G.: Quando i numeri ingannano. Cortina, Milano (2003)
14. Loewenstein, G., Weber, E., Hsee, C., Welch, N.: Risk as Feelings. Psychological Bulletin 127(2), 267–286 (2001)
15. Lerner, J.S., Keltner, D.: Beyond valence: Toward a model of emotion-specific influences on judgment and choice. Cognition and Emotion 4(4), 473–493 (2000)
16. Lazarus, R.S.: Emotion and Adaptation. Oxford University Press, New York (1991)
17. Gigerenzer, G.: Reckoning with Risk: Learning to live with uncertainty. Allen Lane/Penguin, London (2002)
18. Goldstein, J.M., Seidman, J.L., Horton, N.J., Makris, N., Kennedy, D.N., Caviness, V.S., et al.: Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. Cer. Ctx. 11, 490–497 (2001)
19. Tannen, D.: Gender and Family Interaction. In: Holmes, J., Meyerhoff, M. (eds.) The Handbook on Language and Gender, pp. 179–201. Basil Blackwell, Oxford (2003)
20. Hamann, S.: Sex differences in the responses of the human amygdala. Neuroscientist 11, 288–293 (2005)
21. Leonard, C.M., Towler, S., Welcome, et al.: Size matters cerebral volume influences sex differences in neuroanatomy. Cereb Cortex 18(12), 2920–2931 (2008)
22. Podell, K., Lovell, M., Zimmerman, M., Goldberg, E.: The Cognitive Bias Task and lateralized frontal lobe functions in males. Journal of Neuropsychiatry and Clinical Neuroscience 7, 491–501 (1995)
23. Goldberg, E.: The Executive Brain: Frontal Lobes and the Civilized Mind. Oxford University Press (2001)

Influence of Induced Mood on the Rating of Emotional Valence and Intensity of Facial Expressions

Evgeniya Hristova and Maurice Grinberg

Department of Cognitive Science and Psychology, Research Center for Cognitive Science
New Bulgarian University, Sofia, Bulgaria
ehristova@cogs.nbu.bg, mgrinberg@nbu.bg

Abstract. The current study investigates the influence of mood (sad, happy, or neutral) on the valence/intensity ratings of facial expressions of sad, happy, and neutral emotions. The study uses video clips for mood induction and color photographs of emotional facial expressions. Under these conditions, the results show that participants give more extreme ratings to the emotion displayed (happy or sad) when in happy or sad mood without mood-congruence effects. This effect supports the conclusion that arousal alone may play a role in emotion valence/intensity rating (in contrast to results showing mood congruence in other tasks like emotion recognition and detection of emotional expression change). The explanation proposed in the paper is that experienced arousal might guide judgments about the intensity of emotions expressed by other people – when in a more aroused state, a person tends judge that other people also experience more intense emotions.

Keywords: emotions, mood induction, emotional facial expressions, evaluation of emotions.

1 Introduction

In everyday life, we constantly judge the emotions experienced by other people. One of the most informative cues in such judgments is the emotional facial expression. There is a great interest in studying how the information about the emotions in facial expressions is acquired – which facial features are most informative, what is the pattern of looking at these features, etc.

There is a lot of research demonstrating that mood influences human cognitive processing. Affective states have an effect on judgments, decision, perception, and thinking [1-4]. This research has shown that affective states influence judgments by shifting them towards the experienced emotion (e.g. when participants feel happy, they give more favorable judgments). Positive mood also influences memory, reasoning, and visual perception: people tend to use more global processing when in happy mood (global processing is demonstrated by greater reliance on heuristics, scripts, and holistic strategies). On the other hand, people use more local features when they are sad. Schmid et al. [5] investigated the influence of the mood in emotion recognition tasks using eye-movement recordings. For the purpose, they induced either happy or

sad mood and after that presented photos for emotion recognition task (happiness, sadness, anger, and fear). Results showed that participants in happy mood used more global processing styles than participants in negative mood.

There are also studies with a special focus on the influence of mood on emotion recognition. Some of them [6, 7] have found that depressed patients demonstrate impaired emotion recognition and also tend to provide more negative ratings. Schmid & Mast [8] manipulated the mood of healthy participants and studied how that manipulation changes emotion recognition. They found that the induced mood impaired the recognition of the mood-incongruent facial expressions. Other research [9] has demonstrated that participants with induced sad mood tend to perceive more sadness and less happiness in sad and happy faces, respectively. Also, sad and happy participants were more sensitive to and detected earlier changes in sad and happy facial expressions, respectively, when they were gradually changed from sad and happy to neutral or from sad to happy and happy to sad [10, 11]. In the present study, we wanted to explore the influence of induced mood on emotion recognition in a new task – emotion rating – in experimental settings similar to the previously used in this type of research. We used mood induction using short movie episodes following successful and well established procedures [12, 13]. Next, the task of the participants was to rate the valence/intensity of expressed emotion of human faces taken from the FACES database [14]. The induced mood and the selected facial expressions were happy, neutral, and sad. The human faces were presented in relatively natural settings (in color, with hair and background) to make the task as close as possible to real situations. In our opinion, this is a promising way of addressing directly some issues of interest like mood-congruity in emotion rating and like stronger ecological validity related to the presence of non-relevant features like hair and background. As will be discussed below, our results give evidence that the influence of mood on emotion rating may be due to arousal and not to mood-congruence as research using other tasks (e.g. emotional expression change) has previously suggested.

2   Goals and Hypothesis

In the present study, the influence of the mood experienced by participants on their valence/intensity ratings of facial emotional expressions is explored.

In this field of research, as the discussion in the previous section has shown, emotion recognition is naturally a central task of interest due to its importance in social interactions and its evolutionary value. On the other hand, in our opinion, the evaluation of the expressed emotion valence/intensity cannot be underestimated as not only the perceived emotion but also its intensity is an important factor for action and decision making in a larger social context.

While [8] and [5] used black and white photos and only the faces were presented (the stimuli were oval-shaped with hair and background removed), in the current study, we used color photographs without removal of any information (hair or background). This is done in order to explore the emotion recognition in more natural

settings. As emotion recognition is a process that takes place in everyday life and its study in natural settings seems important to us.

Therefore, our main goal is to look for possible congruency effects of the induced mood and valence/intensity rating of facial emotional expressions in relatively natural settings. More precisely, the study aims at exploring how participants' ratings of emotional valence/intensity are affected by their mood. The hypothesis, based on research on emotion recognition discussed in the preceding section, is that valence/intensity ratings will be shifted towards the experienced mood, e.g. neutral emotional expressions will be rated as more negative in negative mood compared to positive mood.

The second goal is to study how the mood influences visual information acquisition. We hypothesize that negative mood will provoke more elaborate processing than positive mood which will lead to longer observation times in sad induced mood compared to happy induced mood, based on the differences in processing information discussed previously.

3   Method

3.1   Stimuli

In the current study, the mood of the participants was manipulated between-subjects on 3 levels – happy, neutral, and sad.

Short video-clips were used for mood induction. Meta-reviews demonstrate that this is one of the most effective techniques for inducing emotional states [15, 16]. Clips, shown to elicit the target moods, are used based on [12, 13]. Video-clips from the following movies were selected: from 'When Harry met Sally' (duration 2' 35'') for happy mood induction; from 'Return to me' (3'36'') for sad mood induction; and from 'Hannah and her sisters' (1'30'') for neutral mood induction. All video-clips are taken from movies with actors in English with Bulgarian subtitles.

Three types of emotional facial expressions are used – happy, neutral, and sad – taken from the FACES database [14]. Photographs from 9 female (IDs 20, 48, 54, 63, 71, 90, 115, 152, and 182) and 9 male (IDs 8, 13, 37, 57, 109, 114, 127, 147, and 153) actors are used, each actor presenting each of the emotional expressions. The stimuli were presented in 3 pseudo-randomized sequences, each consisting of 18 photographs, with an equal number of faces with happy, neutral, and sad emotional expressions. A photograph of a given actor was used only once in a given sequence. In order to get the participants accustomed to the experimental procedure, in the beginning of each list, 2 additional photographs (1 with a male and 1 with a female actor) with neutral expressions are included (IDs 10 and 89). Each of these 3 presentation sequences is preceded by a happy, neutral, or sad mood induction resulting in a total of 9 presentation conditions.

3.2   Design and Procedure

The study employs a 3×3 factorial design with induced mood ('happy' vs 'neutral' vs 'sad') as a between-subjects factor and facial emotion expression ('happy' vs 'neutral' vs 'sad') as a within-subjects factor.

The video-clips were presented with the instruction to watch them for subsequent rating of liking. After the end of the video-clip, the participants were asked to rate their mood on a 9-point Likert scale ranging from '−4' = 'extremely sad' to '+4' = 'extremely happy'. In order to avoid reactivity in the subsequent mood induction ratings, no specific instruction with regard to the mood induction was given before watching the video-clips.

After the mood induction phase, each participant was presented with color photographs of human faces with sad, neutral, or happy expressions in a self-paced presentation. For each stimulus the participant had to rate the emotional expression of the face in the photo on a 7-point Likert scale ranging from '−3' = 'very sad' to '+3' = 'very happy'. The duration of the presentation of the emotional expressions was self-paced. The observation time was taken to be the time during which participants looked at each photograph. It was measured from the beginning of the stimulus presentation to the mouse click after which a screen with the rating scale appeared.

Thus, in order to answer the main questions addressed in the paper, namely whether there are differences in valence/intensity ratings depending on the induced mood and how the induced mood influences the observation time needed for making the rating, the following metrics are used:

- mean emotional valence/intensity rating for the facial expression;
- mean observation time.

3.3   Participants

93 participants (34 male, 59 female) took part in the experiment: 32 in the happy mood condition; 30 in the neutral mood condition; and 31 in the sad mood condition. Participants' age ranged from 18 to 44 years (average 24 years). The participants were university students taking part in the study for partial fulfillment of course requirements or voluntarily.

4   Results

4.1   Mood Induction

First, a manipulation check was performed. The results demonstrate that the mood manipulation was successful (see Table 1) the average ratings given by the participants about their mood for happy, neutral, and sad moods differ significantly ($F(2, 90)$ = 87.99, $p < .001$). Bonferoni post-hoc test shows that all differences are statistically significant at .001 level.

Table 1. Mood rating after the manipulation on a 9-point Likert scale (from '−4' = 'very sad' to '+4' = 'very happy')

Target Mood	Mean	SD
Happy	2.8	1.2
Neutral	0.5	1.0
Sad	−1.1	1.3

4.2 Face Emotion Ratings

In Fig. 1, the ratings of facial emotion expression intensity depending on the induced mood are shown. Mean valence/intensity ratings are analyzed in a repeated-measures ANOVA with facial emotional expression ('happy' vs 'neutral' vs 'sad') as a within-subjects factor and 'induced mood' ('happy' vs 'neutral' vs 'sad') as a between-subjects factors.
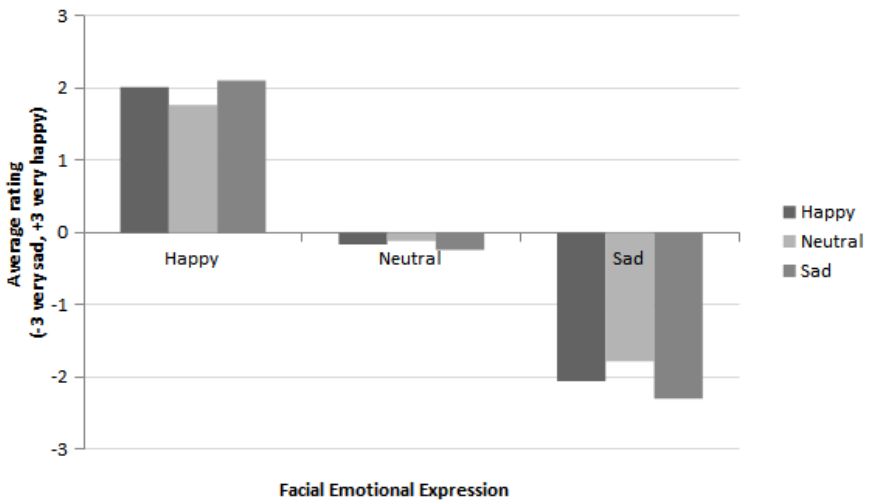
Fig. 1. Face emotion ratings with respect to induced participants' mood.

There is a main effect of the facial emotional expression ($F(2, 180) = 699.94$, $p < .001$): the mean ratings are 1.96 for happy emotional expressions, −0.18 for neutral emotional expressions, and −2.05 for sad emotional expressions, respectively (post =-hoc tests show that all differences are significant at .001 level, Bonferoni correction applied). There is no main effect of induced mood on valence/intensity ratings ($p = .98$). There is an interaction between induced mood and facial emotional expression ($F(4, 180) = 13.3$, $p < .001$). To explore the interaction found, additional analyses were performed separately for happy and for sad faces.

For the ratings of happy faces, there is a main effect of induced mood ($F(2, 90) = 19.19$, $p < .001$). Bonferoni post-hoc test shows that there are significant differences between happy and neutral ($p < .001$) and neutral and sad ($p < .001$) moods,

respectively. Happy faces are rated as more happy by the participants in happy or sad mood compared to the participants in neutral mood (see Fig. 1).

Similarly, for the ratings of sad faces, there is a main effect of induced mood ($F(2, 90) = 7.95$, $p < .001$) and the ratings in neutral mood are significantly higher (less negative) than the ratings in happy ($p < .001$) or sad mood ($p < .004$). Sad faces are rated as more sad by the participants in happy or sad mood compared to the participants in neutral mood.

4.3   Observation Times

The stimuli observation time is measured from the time of appearance of the picture until its disappearance after participants had pressed the mouse left button. The data was not normally distributed and a logarithmic transformation was applied.

Mean observation times were analyzed with repeated-measures ANOVA with facial emotional expression (happy vs. neutral vs. sad) as a within-subjects factor and 'induced mood' (happy vs. neutral vs. sad) as a between-subjects factors. When needed, Bonferoni correction for multiple comparisons was applied.

There is a main effect of the facial emotional expression ($F(2, 180) = 10.92$, $p < .001$). The observation times for the happy faces are shorter than the ones for neutral and sad faces ($p < .001$ and $p = .002$, respectively). It takes less time for the participants to rate a happy facial expression. The means of the observation times for each facial expression are shown in Fig. 2.

There are no significant differences between the observation times with respect to induced mood, nor any interaction between induced mood and the emotional facial expressions of the photographs. Thus, no conclusion about differences in processing depending on induced mood can be drawn for valence/intensity rating.
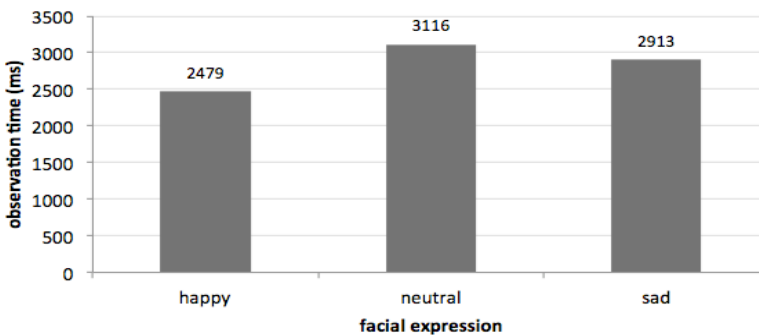
Fig. 2. Comparison of observation times for happy, neutral, and sad faces.

5   Discussion and Conclusion

In the present study, the influence of induced mood on valence/intensity ratings of facial emotional expressions is studied. Based on the previous studies on emotion

recognition, the expectation was that mood-congruency effects will be observed. However, such effects were not found. Instead, higher valence/intensity ratings are obtained for both sad and happy emotional expressions for participants in both happy and sad induced mood.

This led us to conclude that when rating emotional valence/intensity, participants are influenced by the arousal associated to the induced mood without depending on its specific type – happy or sad. As this is a novel effect, not reported in the literature to our knowledge, it requires a deeper and more systematic analysis. It might be the case that the specific settings of our experiment have led to lower sensitivity to the specific mood or emotion expression. One possibility could be that the emotions expressed by the actors in the photographs are too extreme and do not allow for noticeable influences from the mood of the perceivers. But, this cannot explain the difference between the ratings in neutral vs non-neutral mood. In order to understand better the observed effects, stimuli with more graded emotional expression should be used. Moreover, the valence/intensity rating scale used should be split in the usual valence, arousal, and dominance scales in order to be able to identify and measure finer effects and their origin.

However, there are also good reasons to consider that the effect found is due to the influence of the arousal experienced by the participants. As the now classical study of Schachter & Singer [17] showed, physiological arousal is interpreted in the light of the environmental context in order determine the experienced emotion by the person herself. In our view, experienced arousal could influence the judgments not only about the emotions experienced by the subject, but also the judgments of emotions experienced by others. For instance, when in a more aroused state, a person tends to perceive and judge that other people also experience more intense emotions. This is a novel explanation, which has not been proposed previously and deserves further experimental efforts.

Our expectations about observation times, related to previous research, were not met either. They were based on the hypothesis of selective influence of the induced mood on information processing relating information acquisition in a happy mood to more global (fast) processing and in a sad mood to more local feature based (slow) one, respectively. In our experiment, the observation times in all moods were not significantly different with respect to the mood induced. The only significant result was that happy face expressions require shorter observation time. To draw more general conclusions about the dependence of processing on mood, more data is needed.

In summary, the current study contributes to the field of perception of facial expressions of emotions by using a new task – valence/intensity rating. An effect of the induced mood is found, namely, an increase in valence/intensity ratings when the induced mood is different from neutral independent of the non-neutral mood type (sad or happy). This new experimental finding and its theoretical implications need to be explored in experiments manipulating the level of arousal and its dependence on a variety of tasks related to emotion recognition.

References

1. Gasper, K., Clore, G.L.: Attending to the Big Picture: Mood and global verasus local processing of visual information. Psychological Science 13, 34–40 (2002)
2. Gasper, K.: Do you see what I see? Affect and visual information processing. Cognition and Emotion 18, 405–421 (2004)
3. Schwarz, N., Bless, H.: Happy and Mindless, But Sad and Smart? The Impact of Affect States on analytic reasoning. In: Forgas, J. (ed.) Emotion and Social Judgement, pp. 55–71. Pergamon, Oxford (1991)
4. Schwarz, N., Clore, G.L.: Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States. Journal of Personality and Social Psychology 45, 513–523 (1983)
5. Schmid, P.C., Schmid Mast, M., Bombari, D., Mast, F.W., Lobmaier, J.S.: How Mood States Affect Information Processing During Facial Emotion Recognition: An Eye Tracking Study. Swiss Journal of Psychology 70, 223–231 (2011)
6. Hale III, W.W.: Judgment of Facial Expressions and Depression Persistence. Psychiatry Research 80, 265–274 (1998)
7. Persad, S.M., Polivy, J.: Differences Between Depressed and Nondepressed Individuals in the Recognition of and Response to Facial Emotional Cues. Journal of Abnormal Psychology 102, 358–368 (1993)
8. Schmid, P.C., Mast, M.S.: Mood Effects on Emotion Recognition. Motivation and Emotion 34, 288–292 (2010)
9. Bouhuys, A.L., Bloem, G.M., Groothuis, T.G.: Induction of Depressed and Elated Mood by Music Influences the Perception of Facial Emotional Expressions in Healthy Subjects. Journal of Affective Disorders 33, 215–226 (1995)
10. Niedenthal, P.M., Halberstadt, J.B., Margolin, J., Innes-Ker, Å.H.: Emotional State and the Detection of Change in Facial Expression of Emotion. European Journal of Social Psychology 30, 211–222 (2000)
11. Niedenthal, P.M., Brauer, M., Halberstadt, J.B., Innes-Ker, Å.H.: When Did Her Smile Drop? Facial Mimicry and the Influences of Emotional State on the Detection of Change in Emotional Expression. Cognition & Emotion 15, 853–864 (2001)
12. Rottenberg, J., Ray, R.R., Gross, J.J.: Emotion Elicitation Using Films. In: Coan, J.A., Allen, J.B. (eds.) Handbook of Emotion Elicitation and Assessment. Oxford Univ. Press, New York (2007)
13. Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., Bartussek, D.: A Revised Film Set for the Induction of Basic Emotions. Cognition & Emotion 19, 1095–1109 (2005)
14. Ebner, N.C., Riediger, M., Lindenberger, U.: FACES—A Database of Facial Expressions in Young, Middle-Aged, and Older Women and Men: Development and Validation. Behavior Research Methods 42, 351–362 (2010)
15. Westermann, R., Spies, K., Stahl, G., Hesse, F.W.: Relative Effectiveness and Validity of Mood Induction Procedures: A Meta-Analysis. European Journal of Social Psychology 26, 557–580 (1996)
16. Gerrards-Hesse, A., Spies, K., Hesse, F.W.: Experimental Inductions of Emotional states and Their Effectiveness: A Review. British Journal of Psychology 85, 55–78 (1994)
17. Schachter, S., Singer, J.: Cognitive, Social, and Physiological Determinants of Emotional State. Psychological Review 69, 379–399 (1962)

A Multimodal Approach for Parkinson Disease Analysis

Marcos Faundez-Zanuy[1], Antonio Satue-Villar[1], Jiri Mekyska[2],
Viridiana Arreola[3,4], Pilar Sanz[3], Carles Paul[1], Luis Guirao[3], Mateu Serra[3,4],
Laia Rofes[4], Pere Clavé[3,4], Enric Sesa-Nogueras[1], and Josep Roure[1]

[1] Fundació Tecnocampus, Avda. Ernest Lluch 32, 08302 Mataró, Spain
{faundez,satue,paul,sesa,roure}@tecnocampus.cat
[2] Brno University of Technology, Brno, Czech Republic
[3] Hospital de Mataró, Consorci Sanitari del Maresme, Spain
{msanz,mserra}@csdm.cat
[4] Centro de Investigación Biomedica en Red de Enfermedades Hepaticas y Digestivas.
Barcelona, Spain
{laia.rofes,pere.clave}@ciberehd.org

Abstract. Parkinson's disease (PD) is the second most frequent neurodegenerative disease with prevalence among general population reaching 0.1-1 %, and an annual incidence between 1.3-2.0/10000 inhabitants. The mean age at diagnosis of PD is 55 and most patients are between 50 and 80 years old. The most obvious symptoms are movement-related; these include tremor, rigidity, slowness of movement and walking difficulties. Frequently these are the symptoms that lead to the PD diagnoses. Later, thinking and behavioral problems may arise, and other symptoms include cognitive impairment and sensory, sleep and emotional problems. In this paper we will present an ongoing project that will evaluate if voice and handwriting analysis can be reliable predictors/indicators of swallowing and balance impairments in PD. An important advantage of voice and handwritten analysis is its low intrusiveness and easy implementation in clinical practice. Thus, if a significant correlation between these simple analyses and the gold standard video-fluoroscopic analysis will imply simpler and less stressing diagnostic test for the patients as well as the use of cheaper analysis systems.

Keywords: Speech analysis, dysphagia, Parkinson disease, database.

1 Introduction

In this study we will focus on three kinds of signals, but the first step will befocused on speech signals and dysphagia.

It is based on a collaboration be-tween an engineering faculty and a Hospital.

1.1 Voice Analysis

In the PD patient, dysphagia is usually accompanied by other oro-bucal symptoms such as hypokinetic dysarthria. Some studies have reported that the presence of both

311

symptoms usually correlates and that voice disorders could be anticipatory of swallowing impairment [1]. Other studies concluded that a clear post-swallow voice quality provides reasonable evidence that penetration-aspiration and dysphagia are absent [2]. Voice analysis is a safe, non-invasive, and reliable screening procedure for patients with dysphagia which can detect patients at high risk of clinically significant aspiration [6]. The volume-viscosity swallow test (V-VST) was developed at the Hospital de Mataró to identify clinical signs of impaired efficacy (labial seal, oral and pharyngeal residue, and piecemeal deglutition) and impaired safety of swallow (voice changes, cough and decrease in oxygen saturation ≥3 %) [4]. The V-VST allows quick, safe and accurate screening for oropharyngeal dysphagia (OD) in hospitalized and independently living patients with multiple etiologies. The V-VST presents a sensitivity of 88.2 % and a specificity of 64.7 % to detect clinical signs of impaired safety of swallow (aspiration or penetration). The test takes 5-10 min to complete and is an excellent tool to screen patients for OD. It combines good psychometric properties, a detailed and easy protocol designed to protect safety of patients, and valid end points to evaluate safety and efficacy of swallowing and detect silent aspirations [3]. However, nowadays voice assessment is usually done by subjective parameters and a more exhaustive and objective evaluation is needed to understand its relationship with dysphagia and aspiration, as well as the potential relevance of voice disorders as a prognostic factor and disease severity marker. Hypokinetic dysarthria is a speech disorder usually seen in PD which affects mainly respiration, phonation, articulation and prosody. Festination is the tendency to speed up during repetitive movements. It appears with gait in order for sufferers to avoid falling down and also in handwriting and speech. Oral festination shares the same pathophysiology as gait disorders [7]. Voice analysis allows the assessment of all these parameters and has been used to evaluate the improvement of PD after treatment [8-11,17]. Voice impairments appear in early stages of the disease and may be a marker of OD even when swallow disorders are not clinically evident, which would allow to establish early measures to prevent aspiration and respiratory complications. Oropharyngeal dysphagia is a common condition in PD patients. In a recent meta-analysis, the prevalence of PD patients who perceive difficulty in swallowing was estimated at 35% but when an objective swallowing assessment was performed, the estimated prevalence of OD reached 82% [20]. This underreporting calls for a proactive clinical approach to dysphagia, particularly in light of the serious clinical consequences associated to OD in these patients. Dysphagia can produce two types of severe complications; a) alterations in the efficacy of deglutition that may cause malnutrition or dehydration which may occur in up to 24% of PD patients [21], and b) impaired safety of swallow, which may lead to aspiration pneumonia with high mortality rates (up to 50%) [22-23]. Aspiration pneumonia remains the leading cause of death among PD patients.

1.2   Balance and Falls

Postural instability is one of the cardinal signs in PD. It becomes more prevalent and worsens with disease progression and represents one of the most disabling symptoms

in the advanced stages of PD, as it is associated with falls and loss of independence [12]. Balance impairments represent a major burden with high impact on individual's functional capacity, mobility, quality of life and survival. Overall, more than half of patients with PD experience falls. Falls are a major milestone in the evolution of PD because their severe consequences such as bone fractures or head injuries leading to disability, institutionalization and death [13]. Most falls occur during posture changes and are unrelated to extrinsic factors, but are dependent upon intrinsic deficits of balance control. However, pathophysiology of balance disorders and postural instability in PD is not well understood. Posturography allows an objective assessment of balance parameters and posturographic studies have contributed to significant advances in understanding the pathophysiology of postural instability in PD, but it still remains to be fully clarified, partially due to the difficulty to distinguish between the disease process and the compensatory mechanisms and also due to the lack of standardized techniques to measure balance. Dopaminergic treatments can provide improvements in postural instability in early- to mid-stage of PD but the effects tend to decrease with time consistent with spread of the disease process to non-dopaminergic pathways.

1.3   Handwriting

Handwriting skill degradation appears in early stages of PD so handwriting analysis is also of interest in the assessment of the disease progression. Alterations of central dopaminergic neurotransmissions adversely affect movement execution during handwriting and automatic execution of well-learned movements. Drawing exercises in a digitalized tablet allows the accurate evaluation and quantification of size, velocity, acceleration, stroke duration and other parameters of handwriting [18-19]. Although beneficial effects of dopaminergic treatments in kinematics of handwriting movements have been reported, PD patients do not reach an undisturbed level of performance, suggesting that dopamine medication results in partial restoration of automatic movement execution [14-15]. Some authors have shown altered parameters in PD as well as a recovery to the skill of a healthy person after medication with apomorphine [16]. Handwriting tests are useful for assessing the effect of medication and for determining the dosage of drugs for a specific patient.

2   Dysphagia and Speech Analysis

This multimodal analysis has started on speech signals in the context of dysphagia test. A large number of group of people suffer dysphagia, as summarized in table 1, as well as their effect.

The medical term for any difficulty or discomfort when swallowing is dysphagia. A normal swallow takes place in four stages, and involves 25 different muscles and five different nerves. Difficulties at different stages cause different problems and symptoms. The four stages of swallowing are the following ones:

1. The sight, smell, or taste of food and drink triggers the production of saliva, so that when you put food in your mouth (usually voluntarily) there is extra fluid to make the process of chewing easier.

2. When the food is chewed enough to make a soft bolus, your tongue flips it towards the back of the mouth to the top of the tube, which leads down to your stomach. This part of your throat is called the pharynx. This part of swallowing is also voluntary.
3. Once the bolus of food reaches your pharynx, the swallowing process becomes automatic. Your voice box (the larynx) closes to prevent any food or liquid getting into the upper airways and lungs, making the food bolus ready to pass down your throat (known as the oesophagus).
4. The oesophagus, which is a tube with muscular walls that contract automatically, then propels the food down to the stomach.

Table 1. Groups affected by dysphagia and its sympthoms

Group of people	Effects
Elderly people	45 % find some difficulty in swallowing; 65 % of those living in residential or nursing homes: chewing and swallowing muscles are weaker, loss of teeth and saliva production reduced.
Stroke sufferers	40 % nerves, muscles and cognitive/brain function affected
Multiple sclerosis or Parkinson's disease	Nervous system and muscles affected
Alzheimer's disease sufferers and severe depressives	Cognitive/brain function affected
Motor Neuron Disease	Nervous system, nerves and muscles affected
People with cancer of the throat and/or mouth	Nerves and muscles damaged by disease and treatment
People with head and neck injuries	Nerves and muscles damaged

Some signs of dysphagia are:

1   Swallow repeatedly.
2   Cough and splutter frequently.
3   Voice is unusually husky and you often need to clear your throat.
4   When you try to eat you dribble. Food and saliva escape from your mouth or even your nose.
5   Find it
6   eep old food in your mouth, particularly when you have not had a chance easier to eat slowly.
7   Quite often kto get rid of it unseen.
8   Feel tired and lose weight.

2.1   Gold Standard for Dysphagia Analysis

The courrent approach for dysphagia analysis has been developed by some of the medical authors of this paper, and can be summarized in figure 1.
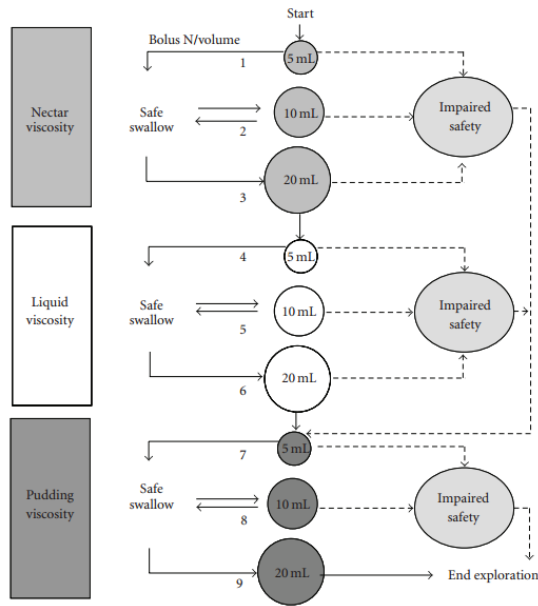
Process for dysphagia analysis based on three liquids of different viscosity and three different volumes per liquid. After swallowing each liquid and volume a word is pronounced by the patient and a speech therapist evaluates the voice quality in a subjective way (just listening to the speech signal).

In those cases where a possible dysphagia problem exists, a videofluoroscopic analysis is perfomed. This diagnose is more invasive as it implies radiation, but it is the procedure to have physical evidence of swal-lowing problems. Figure 2 shows Videofluoroscopic pictures and oropharyngeal swallow response during the ingestion of a 5 mL nectar bolus in: (a) a healthy individual; (b) an older patient with neurogenic dysphagia and aspiration associated with stroke. An increased total duration of the swallow response may be seen, as well as a delayed closure of the laryngeal vestibule and delayed aperture of the upper sphincter. The white dot indicates the time when contrast penetrates into the laryngeal vestibule, and the red dot indicates passage into the tracheobronchial tree (aspiration). GPJ = glossopalatal junction, VPJ = velopalatal junction, LV= laryngeal vestibule, UES = upper esophageal sphincter.

The main goal of the first step of this study is to evaluate if an automatic tool based on speech analysis can be developed to support medical decision during the test depicted in figure 1.

Fig. 1. Process for dysphagia analysis based on three liquids of different viscosity and three different volumes per liquid
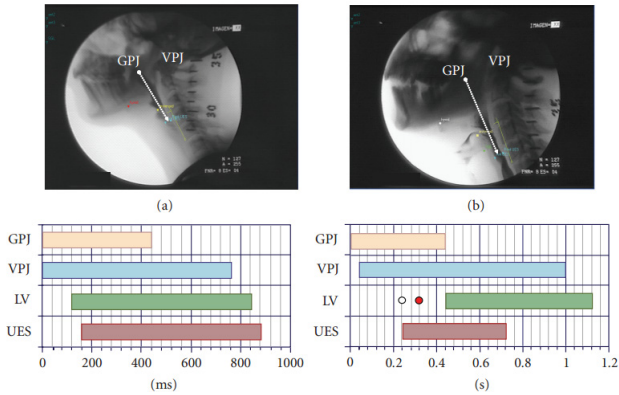
Fig. 2. Videofluorescence images for (a) a healthy individual; (b) an older patient with neurogenic dysphagia and aspiration associated with stroke

3   Database Acquisition and Future Lines

At this moment a speech database is being acquired in the protocol depicted in figure 1, one sample after deglutition of each liquid and volume. Thus, a total of 9 realizations per patient are acquired.

Figure 3 shows the acquisition scenario at Mataro's Hospital. The acquisition setup is based on a capacitor microphone Rode NT2000 (positioned at a distance of approximately 20 cm from the speaker's mouth) and external sampling card (M-AUDIO, FAST TRACK PRO Interface audio 4x4) operating at 48 kHz sampling rate, 16 bit per sample, monophonic recording. Currently we are acquiring 3 patients per week.

The signal processing approach, after database collection will be based on:

(a) Voiced/unvoiced classification and then to check the harmonic to noise ratio (HNR) on the vowels, jitter, shimmer, etc.
(b) To align the sample before and after eating using Dynamic Time Warping. The higher the distance between both realizations, the higher the probability to have deglutition problems.
(c) Some complexity measures

3.1   Gold Standard

Videofluoroscopy (VFS) is the gold standard to study the oral and pharyngeal mechanisms of dysphagia. VFS is a dynamic exploration that evaluates the safety and efficacy of deglutition, characterizes the alterations of deglutition in terms of videofluoroscopic signs, and helps to select and assess specific therapeutic strategies. Since the hypopharynx is full of contrast when the patient inhales after swallowing. Thereafter, VFS can determine whether aspiration is associated with impaired glossopalatal seal (predeglutitive aspiration), a delay in triggering the pharyngeal

swallow or impaired deglutitive airway protection (laryngeal elevation, epiglottic descent, and closure of vocal folds during swallow response), or an ineffective pharyngeal clearance (post swallowing aspiration) [5].

Fig. 3. Acquisition scenario at Mataro's Hospital

References

1. Perez-Lloreta, S., Negre-Pages, L., Ojero-Senarda, A., et al.: Oro-buccal symptoms (dysphagia, dysarthria, and sialorrhea) in patients with Parkinson's disease: preliminary analysis from the French COPARK cohort. European Journal of Neurology 19, 28–37 (2012)
2. Waito, A., Bailey, G.L., Molfenter, S.M., et al.: Voice-quality abnormalities as a sign of dysphagia: validation against acoustic and videofluoroscopic data. Dysphagia 26(2), 125–134 (2011)
3. Rofes, L., Arreola, V., Clavé, P.: The volume-viscosity swallow test for clinicalscreening of dysphagia and aspiration. Nestle Nutr. Inst. Workshop Ser. 72, 33–42 (1998), doi:10.1159/000339979; Epub September 24, 2012, PubMed PMID: 23051998
4. Clavé, P., Arreola, V., Romea, M., Medina, L., Palomera, E., Serra-Prat, M.: Accuracy ofthe volume-viscosity swallow test for clinical screening of oropharyngeal dysphagia and aspiration. Clin. Nutr. 27(6), 806–815 (2008), doi:10.1016/j.clnu.2008.06.011; Epub September 11, 2008, PubMed PMID: 18789561
5. Rofes, L., Arreola, V., Almirall, J., Cabré, M., Campins, L., García-Peris, P., Speyer, R., Clavé, P.: Diagnosis and management of oropharyngeal Dysphagia and its nutritional and respiratory complications in the elderly. Gastroenterol Res. Pract. 2011, 818979 (2011), doi:10.1155/2011/818979; Epub August 3, 2010, PubMed PMID: 20811545; PubMed Central PMCID: PMC2929516

6. Ryu, J.S., Park, S.R., Choi, K.: Prediction of laryngeal aspiration using voice analysis. Am. J. Phys. Med. Rehabil. 83(10), 753–757 (2004)
7. Moreau, C., Ozsancak, C., Blatt, J.-L., Derambure, P., Destee, A., Defebvre, L.: Oral Festination in Parkinson's Disease: Biomechanical Analysis and Correlation with Festination and Freezing of Gait. Movement Disorders 22(10), 1503–1506 (2007)
8. Nagulic, M., Davidovic, J., Nagulic, I.: Parkinsonian voice acoustic analysis in real-time after stereotactic thalamotomy. Stereotact. Funct. Neurosurg. 83(2-3), 115–121 (2005)
9. Gobermana, A.M., Coelho, C.: Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy. NeuroRehabilitation 17, 237–246 (2002)
10. Gobermana, A.M., Coelho, C.: Acoustic analysis of Parkinsonian speech II: L-Dopa related fluctuations and methodological issues. NeuroRehabilitation 17, 247–254 (2002)
11. Stewart, C., Winfield, L., Junt, A., Bressman, S.B., Fahn, S., Blitzer, A., Brin, M.F.: Speech dysfunction in early Parkinson's disease. Movement Disorders 10(5), 562–565 (1995)
12. Kim, S.D., Allen, N.E., Canning, C.G., et al.: Postural instability in patients with Parkinson's disease. Epidemiology, pathophysiology and management. CNS Drugs 27, 97–112 (2013)
13. Grabli, D.: Normal and pathological gait: what we learn from Parkinson's Disease. J. Neurol. Neurosurg. Psychiatry 83 (2012), doi:10.1136/jnmp-2012-302263
14. Tucha, O., Mecklinger, L., Thome, J., Reiter, A., Alders, G.L., Sartor, H., Naumann, M., Lange, K.W.: Kinematic analysis of dopaminergic effects on skilled handwriting movements in Parkinson's disease. J. Neural Transm. 113, 609–623 (2006)
15. Tucha, O., Mecklinger, L., Walitza, S., Lange, K.W.: The effect of caffeine on handwriting movements in skilled writers. Human Movement Science 25(4-5), 523–535 (2006)
16. Eichhorn, T.E., Gasser, T., Mai, N., Marquardt, C., Arnold, G., Schwarz, J., Oertel, W.H.: Computational analysis of open loop handwriting movements in Parkinson's disease: A rapid method to detect dopamimetic effects. Movement Disorders 11(3), 289–297 (1996)
17. Eliasova, I., Mekyska, J., Kostalova, M., Marecek, R., Smekal, Z., Rektorova, I.: Acoustic evaluation of short-term effects of repetitive transcranial magnetic stimulation on motor aspects of speech in Parkinson's disease. J. Neural Transm. 120(4), 597–605 (2013)
18. Drotar, P., Mekyska, J., Smekal, Z., Rektorova, I., Masarova, L., Faundez-Zanuy, M.: Prediction potential of different handwriting tasks for diagnosis of Parkinson's. In: E-Health and Bioengineering Conference (EHB), pp. 1–4 (2013)
19. Drotar, P., Mekyska, J., Rektorova, I., Masarova, L., Smekal, Z., Faundez-Zanuy, M.: A new modality for quantitative evaluation of Parkinson's disease: In-air movement. In: 2013 IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 1–4 (2013)
20. Kalf, J.G., et al.: Prevalence of oropharyngeal dysphagia in Parkinson's disease: a meta-anaysis. Parkinson and Related Disorders 1, 311–315 (2012)
21. Sheard, J.M., et al.: Prevalence of malnutrition in Parkinsom's disease: a systematic review. Nutrition Reviews 69, 520–532 (2011)
22. Fernandez, H.H., Lapane, K.: Predictors of mortality among nursing home residents with a diagnosis of Parkinson's disease. Med. Sci. Monit. 8, CR241–CR246 (2002)
23. Williams-Gray, C.H., Mason, S.L., Evans, J.R., et al.: The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. J. Neurol. Neurosurg. Psychiatry 84, 1258–1264 (2013)

Are Emotions Reliable Predictors of Future Behavior? The Case of Guilt and Other Post-action Emotions

Olimpia Matarazzo[*] and Ivana Baldassarre

Department of Psychology, Second University of Naples, Italy
{olimpia.matarazzo,ivana.baldassarre}@unina2.it

Abstract. This study had two goals: 1) to establish the relative importance of the violation of a moral norm and of the damage done to another person in the genesis of guilt and other post-action emotions; 2) to investigate if the post-action emotions are reliable predictors of future behavior in conditions similar to the ones that elicited them the first time. Through the scenario technique, four typical antecedents of guilt were built, in which the intentionality of the norm violation and of the damage to others were manipulated. In all scenarios the protagonist acted in such a way as to elicit guilt and the other emotions that participants were asked to assess. Thus, we presented a similar situation happening a few months later in which he had to choose whether to behave in the same way as he had behaved previously or in the opposite way. We expected that: (1) moral emotions would stimulate a different behavior from the previous one, whereas selfish emotions should lead to repeat the same behavior; (2) emotions should influence future behavior in an indirect way, through the cognitive mediation of thoughts preceding the decision; (3) the norm violation would have analogous relevance in eliciting guilt to harming another person. On the whole, the results corroborated our predictions, with some exceptions that were discussed.

Keywords: emotion, guilt, post-action emotions, future behavior.

1 Introduction

The idea that emotion plays a central role in decision-making and, more generally, in orienting future behavior is largely accepted in psychological literature [e.g. 1,2,3]. In particular, the "feeling for doing" approach [3] emphasizes the motivational function of emotions and their instrumental value in pursuing goals. According to this approach, the variety of emotions experienced in daily life, each of them with its own peculiar experiential content, has the essential function of guiding behavior, even though the debate about the ways by which emotion influences behavior is open [4]. The most widely accepted theory posits that emotion directly causes behavior and that its function is to lead the organism to behave in such a way as to deal with the emotional event [e.g. 5,6]. The competing theory [4], based on a dual-process model

---

[*] Corresponding author.

distinguishing between "automatic affect" – simple, fast and often not conscious – and "conscious emotion" – a more complex phenomenon entailing the awareness of subjective experience – argues that only the former shapes behavior directly, while emotion affects behavior indirectly, as a feedback system. According to this perspective, conscious emotion influences cognitive processes which in turn affect decision making and behavior regulation [4]. The main function of emotion is to anticipate the emotional outcome of future behavior on the basis of the "affective residue" left by past actions. Thus, emotion is able to enhance behaviors leading to desired outcomes and modify the ones leading to undesired outcomes.

The paradigmatic example of this view are moral emotions (e.g. guilt, shame, pride, embarrassment), i.e. the emotions that are generated by the conformity with, or the transgression of, the norms that regulate behavior and feelings of the members of a society [e.g. 7,8]. Moral emotions are typically post-action emotions, in the sense that their insurgence provides feedback on the adequacy or inadequacy of actual behavior and prompts individuals to reflect on their conduct and eventually envisage counterfactual courses of actions (which constitutes the cognitive experience of such emotions). However, the influence that moral emotions exert on decision making and on future behavior comes from the anticipatory feedback of its emotional consequences [8,9]. More precisely, it is the goal to avoid or repeat previous emotional experiences that motivates the choice between adopting or preempting a course of actions similar to the one that had generated those emotions.

Perhaps the most prototypical moral emotion (at least in Western culture) is guilt, as evidenced by the considerable debate about its nature [see for review 9, 10]. According to the intrapsychic approach [e.g. 11,12], guilt is elicited by the violation of internalized moral norms and motivates self-punishing or reparative behavior. According to the interpersonal approach [e.g. 13-15], guilt is rooted in empathic concern for the well-being of others and is elicited by causing harm to another person, deliberately or unintentionally: thus, guilt-based behavior aims to restore the relationship with the damaged person.

Along the empathic axis lies one of the main lines of demarcation between guilt and another typical moral emotion, shame. While guilt is usually seen as an adaptive emotion, aiming at re-establishing potentially broken relationships, shame is envisaged as a maladaptive emotion tending to reduce rather than restore interpersonal relationships [8]. So, guilt motivates pro-social and reparative behavior [8,16], while shame motivates withdrawal, even if some findings reveal that shame can also induce restorative behavior [17]. It is noteworthy that empirical research has scarcely corroborated the traditional distinctions between guilt and shame based on transgression of moral norms vs. social or communitarian values, and on private vs. public nature of transgression [8,18]. The differences between the two emotions are rather envisaged in their phenomenological experience (guilt would be centered on the action, while shame would be centered on the entire self), in their level of pervasiveness (shame would be more pervasive than guilt) and in the behavior they motivate (other- or relationship-preserving vs. self-preserving) [8,12]. In our view, however, the literature on moral emotions, in particular on guilt, has emphasized the interpersonal nature of this emotion, based on the empathic concern for others, and has neglected the intrapsychic

point of view, according to which guilt arises from the transgression of a moral norm, regardless of the harm caused to others. To investigate the respective importance of interpersonal and intrapsychic component in the genesis of guilt (and other moral emotions) is one of the goals of this study, the other being to determine whether the emotions that arise after an action are good predictors of future behavior.

2   Overview of the Experiment

This study aimed at pursuing two goals: 1) to compare the two main approaches about the intrapsychic vs. interpersonal nature of guilt; 2) to investigate whether and to what extent the emotions felt after an action can be reliable predictors of future behavior, in conformity with the hypothesis of the "feeling for doing" approach. As to the first goal, the way chosen to compare the two approaches was to assess the intensity of guilt elicited by the typical situational antecedents of this emotion. To this end, through the scenario technique, we built four antecedents of guilt by varying its structural determinants: the norm violation and the harm to another. The third determinant we varied was the presence/absence of material advantage the male protagonist could gain from his actions. We created situations similar to those of daily life and capable of eliciting a wide range of emotions, other than guilt (relief, shame, anger toward oneself, satisfaction, sorrow), in order to assess their reciprocal relations. More specifically, since in three of our scenarios the protagonist had to face a dilemma concerning the choice between moral vs. selfish options and preferred the selfish one, we selected relief and satisfaction to assess the amount of positive emotional reaction generated by the decision to pursue his own material concerns even at the expense of moral values. Sorrow was chosen to gauge the broad range of emotional negative states deriving from hurting another person or from an incorrect action. Anger toward oneself was selected to evaluate the amount of aggressive auto-directed reaction associated with the awareness of having committed wrongdoing or unjust actions. Finally, shame was included in this study because it represents one of the two main moral emotions (the other being guilt): in order to contribute to the debate on the distinction between guilt and shame, we examined the extent to which these emotions were elicited by undiscovered breaches of moral norms and/or by damage caused to other people. We expected that the *norm violation* would have a similar relevance in eliciting guilt to *harming another person*, and that both antecedents would be more specific of guilt rather than shame (and other emotions).

In order to pursue the second goal of this study, we subdivided the experiment into two parts: at time 1, we presented a situation in which the protagonist acted in a way capable of eliciting guilt and/or the other envisaged emotions; at time 2, we presented a situation similar to that of time 1 in which the protagonist had to choose whether to behave in the same way as he did the first time or in the opposite way. We expected that negative and moral emotions (anger toward oneself, sorrow, guilt, and shame) would stimulate a different behavior from the previous one, whereas positive selfish emotions (relief and satisfaction) should lead to repeat the same behavior. In conformity with Baumeister *et al.*'s [4] model, we expected that emotions should influence future behavior in an indirect way, through the cognitive mediation of thoughts preceding the decision.

3   Method

Participants, Materials and Procedure. Two hundred undergraduates, aged from 18 to 30 years (Mean = 25.46; SD = 2.58), participated in this study, as unpaid volunteers. No Psychology student was recruited so that prior knowledge about emotion would not influence the results. Each participant was randomly assigned to one of the 4 experimental conditions, except for gender for which they were paired.

By means of the scenario technique, four typical antecedents of guilt were built, in which the intentionality of the norm violation and of the damage to others were manipulated. The third manipulated variable was the presence/absence of the material self-advantage. In the first scenario (henceforth S1) the protagonist deliberately violated a moral norm to gain an advantage for himself but without causing harm to others; in the second scenario (S2) the protagonist deliberately violated a moral norm and caused harm to another person to gain an advantage for himself; in the third scenario (S3) the protagonist deliberately violated a moral norm by causing harm to another person without the aim of gaining any material advantage for himself; in the fourth scenario (S4) the protagonist unintentionally caused harm to a friend: so, no moral norm was violated nor material advantage was pursued. While in the first three scenarios the protagonist had to reach a decision through facing a moral dilemma, in S4 the damage was unintentional: so, no moral dilemma was posed. In all scenarios the nature and seriousness of the norm violation, of the damage to others, and of material self-advantage (if present) were kept constant so as to eliminate or reduce spurious effects due to these contingent variables. After reading the scenario, participants were asked to identify with the protagonist and to specify on a 9-point scale (1= not at all, 9 = extremely) the extent to which he felt the following 6 emotions: satisfaction, anger toward oneself, shame, relief, sorrow, guilt (Time 1). Participants were then asked to read a second scenario where, a few months later (Time 2), a situation similar to the initial one occurred. In the moral dilemmas, the protagonist had to decide whether to behave in a similar way to the previous one or in an opposite way. In the situation of unintentional damage, he has to decide whether to adopt precautionary measures to avoid the possibility of causing harm to another person or not. After the scenario, the protagonist's two possible thoughts (moral/altruistic – henceforth moral - vs. selfish) and two choice options (moral vs. selfish) were presented. Participants were asked to evaluate, on a 9-point scale (1 = strongly disagrees, 9 = extremely agrees) the protagonist's degree of agreement with each of the two thoughts and to predict his choice between two options: moral vs. selfish. The order of thoughts, types of choice, and emotions was randomized.[1]

4   Results

In order to test the effect of the four scenarios on the six emotions, a mixed ANCOVA was performed, with the scenario as between variable and emotion as within variable;

---

[1] The scenarios and the instructions were pre-tested by means of a pilot study with 32 undergraduates (8 for each experimental condition).

gender (coded as dummy variable: 1 = male; 0 = female) was introduced as covariate. Means (with standard deviations) are reported in Table 1.

Table 1. Means (with standard deviations) of emotions as function of scenarios

	S1	S2	S3	S4
Relief	7.90 (1.83)	5.82 (3.08)	5.34 (2.73)	1.16 (0.61)
Anger toward oneself	5.76 (2.61)	6.20 (2.63)	4.90 (2.75)	6.44 (2.72)
Guilt	6.38 (2.57)	7.46 (2.28)	5.66 (3.26)	8.56 (1.05)
Sorrow	4.92 (2.41)	5.92 (2.63)	5.32 (3.08)	8.64 (.077)
Shame	5.44 (2.40)	6.34 (2.37)	4.50 (2.77)	4.36 (2.73)
Satisfaction	4.94 (2.57)	3.98 (2.59)	5.08 (2.69)	1.26 (0.98)

S1 = Norm transgression with self-advantage but without harm to others
S2 = Norm transgression with self-advantage and harm to others
S3 = Norm transgression with harm to others but without self-advantage
S4 = Unintended harm to others

Results showed two main effects**:** of scenario [$F(3, 195)$ = 7.91, p <.001, η^2p =.108] and of emotion [$F(5, 975)$ = 29.04, p <.001, η^2p =.130], and two interaction effects: emotion x scenario [$F(15, 975)$ = 28.38, p <.001, η^2p =.304], and emotion x gender [$F(5, 975)$ = 2.71, p <.01, η^2p =.014]. The latter, examined through parameter estimates, was due to the finding that males gave higher scores than females on relief and lower scores on anger toward oneself. As to the main effects, S1 & S2 elicited higher scores than S3 & S4; the rank order of emotion scores was the following: guilt > sorrow > anger toward oneself > relief & shame > satisfaction. The scenario x emotion interaction, examined through the simple effects analysis (with Bonferroni adjustment), revealed the following findings: S1 elicited mainly relief and guilt; relief obtained higher scores than the other 4 emotions, whilst guilt obtained higher scores only than sorrow and shame; anger toward oneself and satisfaction did not differ from guilt, sorrow and shame. S2 evoked principally guilt, followed by shame, anger toward oneself, sorrow and relief which had similar scores, while satisfaction obtained the lowest scores. As to S3, all emotions obtained analogous scores, except for guilt, which received higher scores than anger toward oneself and shame. The rank order of emotions evoked by S4 was the following: guilt & sorrow > anger toward oneself > shame > relief & satisfaction.

If the scenario x emotion interaction is scrutinized from the emotion perspective, then the results show that guilt was mainly evoked by S4 and S2; however, S2 produced only higher scores than S3 and analogous scores to S1 which, in turn, evoked similar scores to S3. Anger toward oneself has been evoked with similar intensity by all scenarios, except for S4, which elicited higher scores than S3. As to sorrow, S4 produced the highest scores; the other scenarios did not differ from one another. S2 elicited shame more than S3 and S4, whereas S1 did not differ from the other scenarios. Relief was produced mainly by S1, followed by S2 and S3 (whose scores did not differ significantly), and lastly by S4, which obtained the lowest scores. Finally, as to satisfaction, S4 caused the lowest scores, while the other scenarios did not differ reciprocally.

Successively, to test whether the emotions' intensity attributed to the protagonist at time1 would predict his future choice at time 2, directly or via the mediation of moral and egoistic thoughts, multiple logistic regression analyses were performed through

PROCESS macro[2], which allows to estimate direct and indirect effects in multiple mediator models with multiple predictors and/or binary dependent variable [24]. The six emotions were put in the analyses as predictors, the two thoughts (moral - vs. selfish) acted as mediators, choice (between moral vs. selfish options) was the criterion variable. Choice was coded as dummy variable (1 = moral choice; 0 = selfish choice). Gender was put in the analyses as covariate.

In mediational models [e.g. 19-21] the total effect of a predictor X on a dependent variable (or criterion) Y designates the overall effect that X exerts on Y both directly and through one or more mediators M; the direct effect of X on Y indicates the effect of X on Y independent of M's influence on Y; the indirect effect designates the effect that X exerts on Y through M(s). The indirect effect assesses the amount of mediation and is quantified as the product of the effects of X on M and of M on Y. The presence of the indirect effect reduces or annuls the direct effect of X on Y compared to the total effect, except in the case of inconsistent mediation, where the different signs of the relationship between X, M(s), and Y can produce opposing mediation effects that reduce or annul the total effect. Finally, if the effect of X does not vary after introducing M(s), then there is no mediation.

In Table 2 the effects of emotions on mediators are showed; in Table 3 total and direct effects of emotions on choice are reported, while the indirect effects of emotions on choice via mediators are reported in Table 4. Note that the total effects of emotions on choice – not calculated by the macro - was assessed through a multiple logistic regression analysis (enter method).

As regards the influence of the six emotions on moral thought [R^2 =.509; $F(7,192)$ = 28.47; $p<.001$], results showed that it increased as function of anger toward oneself and decreased as function of satisfaction. As to selfish thought [R^2 =.426; $F(7,192)$ = 20.32; $p<.001$], it decreased as function of anger toward oneself, guilt, and shame.

Total effects of emotions on choice [-2LL =160.09; Nagelkerke R^2 = .563] showed that moral choices augmented when the intensity of guilt, anger toward oneself, and relief augmented, and when the intensity of satisfaction decreased.

Table 2. Results of mediational regression analyses of predictors (emotions) and covariate (gender) on mediators (moral and selfish thoughts)

	Moral thought			Selfish thought		
	B	S.E.	p	B	S.E.	p
Constant	5.01	.663	.000	7.52	.821	.000
Relief	-.035	.054	.519	.058	.067	.394
Anger toward oneself	.190	.078	.015	-.236	.096	.015
Guilt	.153	.087	.081	-.246	.108	.024
Sorrow	.045	.070	.521	.162	.087	.065
Shame	.042	.068	.533	-.222	.084	.009
Satisfaction	-.233	.074	.002	.105	.092	.255
Gender	.038	.235	.870	.295	.291	.310

---

[2] The macro is available on `http://www.afhayes.com`

Table 3. Results of mediational logistic regression analyses on choice (1 = moral choice) with predictors, covariate and mediators: total and direct effects

	Total effects			Direct effects		
	B	S.E.	*p*	B	S.E.	*p*
Constant	-4.21	1.30	.001	-4.56	2.66	.086
Relief	.210	.104	.043	.380	.154	.014
Anger toward oneself	.268	.124	.030	.104	.174	.551
Guilt	.448	.160	,005	.206	.240	.392
Sorrow	.006	.129	.961	.014	.239	.951
Shame	.024	.117	.837	-.079	.171	.643
Satisfaction	-.301	.134	.024	-.378	.214	.077
Gender	-.251	.416	.545	.012	.612	.984
Moral thought				1.43	.337	.000
Selfish thought				-1.11	.262	.000

Table 4. Results of mediational logistic regression analyses on choice (1 = moral choice): indirect effects

	Indirect effects through moral thought			Indirect effects through selfish thought		
	Effect	S.E	*p*	Effect	S.E.	*p*
Relief	-.050	.081	.534	-.064	.078	.415
Anger toward oneself	.273	.131	.038	.261	.126	.038
Guilt	.220	.139	.113	.276	.139	.049
Sorrow	.065	.105	.536	-179	.108	.097
Shame	.061	.101	.547	.246	.112	.028
Satisfaction	-.334	.135	.013	-.116	.108	.283

After controlling for the mediational variables [-2LL =75.83; Nagelkerke R^2 = .836], the direct effect of emotions was no longer significant, except for relief. On the contrary, the two mediators exerted a robust effect on choice: moral thoughts increased moral choices whereas selfish thoughts decreased them. The indirect effect of emotions via mediators was the following: anger toward oneself affected choice through both mediators, guilt and shame affected choice via selfish thought, whereas satisfaction affected choice through moral thought. Anger toward oneself increased moral thought which, in turn, increased moral choices whereas it decreased selfish thought which, in turn, decreased moral choices: so this emotion had a positive indirect effect through both mediators. Guilt and shame decreased selfish thought, which, in turn, decreased selfish choices, and, consequently, they had a positive indirect effect on moral choices. The absence of total effect of shame on choice may be due to an inconsistent mediation. Satisfaction decreased moral thought, which increased moral choices: so, it had a negative indirect effect. Finally, no kind of effects was produced by sorrow and gender.

5 Discussion of the Results and Conclusion

This study had two goals: 1) to establish the relative importance of the violation of a moral norm and of the damage done to another person in the genesis of guilt and other (moral or) post-action emotions; 2) to investigate if the post-action emotions are

reliable predictors of the future behavior in conditions similar to the ones that elicited them at time 1.

As to the first goal, our results showed that the four scenarios we had built were able to evoke the six envisaged emotions, principally guilt – a finding that proves their adequateness. Guilt intensity was especially high in the situations of unintended harm to a friend (S4) and of trick and damage (S2 = violating a moral norm to gain an unmerited advantage, and thereby also damaging another person). More specifically, the unintended harm elicited more guilt than the conditions in which there was only a norm violation without harm to others (S1) or in which the harm was intentionally caused (S3), whereas the" trick and damage" scenario (S2) produced only more guilt than the one of deliberate harm (S3). As regard the debate on the intrapsychic or interpersonal nature of guilt, at first sight these results seem to support the latter position, according to which this emotion arises from an empathic concern for the others: unintentional harm, in which no norm is violated (S4), evokes more guilt than the scenario in which the norm violation does not involve harm to others (S1). However, the strength of this conclusion is weakened by the fact that the scenario in which an intentional harm is caused to others without self-advantage (S3) produces similar scores to the one in which there is only the violation of the norm on cheating (S1). A plausible explanation of these findings may be that S4 is the only situation in which the protagonist is not faced with a moral dilemma where two conflicting options – each of which involving opposite (moral vs. selfish) goals – are contrasted. So, guilt (and sorrow which reached the same intensity) was not mitigated by the positive emotions (relief and satisfaction) elicited by the fulfillment of the selfish goal. The finding that relief and satisfaction obtained the lowest scores on this scenario supports our hypothesis. In our opinion, our results advocate the idea that guilt has two main sources, reciprocally independent – the other-oriented empathy and the internalization of moral norms – and that it can only arise from one of these sources. On the contrary, shame seems to be a more self-centered emotion than guilt and less affected by the empathic concern toward another person. Although in this study shame did not reach high intensity, the "trick and damage" scenario elicited higher scores than harming (intentionally or not) another person. Note that also the norm violation without damaging another (S1) obtained higher scores than S3 and S4, even if this difference was not statistically significant. As concerns the debate about the nature of guilt and shame [for a review see 9,10], our results suggest that guilt is a stronger and more intense emotion than shame, since in all scenarios it obtained higher scores. It is noteworthy to recall that in all scenarios implying a norm violation and/or a deliberate damage, such actions remained hidden (and thus the private nature of transgression was hold constant) so that the possibility of being punished was excluded from the factors that could have increased the intensity of the emotions. Only the unintended damage was discovered but the friend did not exert any retaliation toward the protagonist, even if his presence may be considered a powerful source of emotional reactions. In the other situations the emotions were felt only in relation to oneself.

As regards the second goal of this study, our results corroborated the largely accepted idea that emotion affects decision process and future behavior [e.g. 1,3, 22]. More precisely, our findings showed that four emotions influenced the future choice

through cognitive mediation; only relief had a direct effect on the choice, by increasing the number of moral choices, whereas sorrow did not have any effect. Moral emotions, such as guilt, shame, and anger toward oneself, predicted an opposite behavior from the one that had caused their genesis at time 1, whereas satisfaction predicted an analogous behavior to that at time 1. It should be noted that guilt and shame influenced choice by decreasing the agreement with selfish thought, while satisfaction affected choice by decreasing the agreement with moral thought; only anger toward oneself affected choice through the mediation of both thoughts. Such a finding suggests that this emotion is a basic and crucial component of shame and guilt, representing the self-critical and self-punitive aspects of both emotions.

As to the cognitive mediation through which emotions predicted future behavior, in our study choice at time 2 was presented as a dilemma between two options: maybe this circumstance increased the importance of thought in the decision-making process, as it is recognized in the literature on moral judgment, even by the authors who emphasize the role of emotion in this domain [e.g. 23,24].

Finally, we have to remark that our results also revealed some unexpected findings: contrarily to our predictions, sorrow did not exert any effect on choice and relief increased moral, instead of selfish, choices. Probably sorrow is too broad and vague an emotion to be able to influence or predict specific future behavior. As to relief, it should be noted that this emotion arises when a worrying, uncertain or risky situation ends: in the three moral dilemmas at time 1 of our study, relief was related to the positive outcome of the situation the protagonist faced, despite his undiscovered misconduct. So, we speculate that at time 2 relief may have increased moral choices based on the superstitious belief that one cannot tempt fate again and that misconduct may not be always successful.

References

1. Damasio, A.: Descartes' error: Emotion, reason, and the human brain. Grosset/Putnam, New York (1994)
2. Loewenstein, G., Lerner, J.S.: The role of affect in decision making. In: Davidson, R., Scherer, K., Goldsmith, H. (eds.) Handbook of Affective Science, pp. 619–642. Oxford University Press, New York (2003)
3. Zeelenberg, M., Nelissen, R.M.A., Breugelmans, S.M., Pieters, R.: On emotion specificity in decision making: Why feeling is for doing. Judgment Decis. Making 3, 18–27 (2008)
4. Baumeister, R.F., Vohs, K.D., DeWall, C.N., Zhang, L.: How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. Pers. Soc. Psychol. Rev. 11, 167–203 (2007)
5. Frijda, N.H.: The emotions. Cambridge University Press, Cambridge (1986)
6. Cosmides, L., Tooby, J.: Evolutionary psychology and the emotions. In: Lewis, M., Haviland-Jones, J.M. (eds.) Handbook of Emotions, 2nd edn., pp. 91–115. Guilford, New York (2000)

7. Haidt, J.: The moral emotions. In: Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (eds.) Handbook of Affective Sciences, pp. 852–870. Oxford University Press, New York (2003)

8. Tangney, J.P., Stuewig, J., Mashek, D.J.: Moral emotions and moral behavior. Annu. Rev. Psychol. 58, 345–372 (2007)

9. Ghorbani, M., Liao, Y., Caykoylu, S., Chand, M.: Guilt, Shame, and Reparative Behavior: The Effect of Psychological Proximity. J. Bus. Ethics. 114, 311–323 (2013)

10. Carnì, S., Petrocchi, N., Del Miglio, C., Mancini, F., Couyoumdjian, A.: Intrapsychic and interpersonal guilt: A critical review of the recent literature. Cognit. Processing 14, 333–346 (2013)

11. Freud, S.: The dissolution of the Oedipus complex. In: Strachey, J. (ed.) The standard Edition of the Complete Psychological Works of Sigmund Freud, pp. 173–182. Hogarth, London (1961) (original work 1924)

12. Lewis, H.B.: Shame and Guilt in Neurosis. Int. Univ. Press, New York (1971)

13. Baumeister, R.F., Stillwell, A.M., Heatherton, T.F.: Guilt: An interpersonal approach. Psychol. Bull. 115, 243–267 (1994)

14. Hoffman, M.L.: Varieties of empathy based guilt. In: Bybee, J. (ed.) Guilt and Children, vol. 4, pp. 91–112. Academic Press, New York (1998)

15. Tangney, J.P.: Shame and guilt in interpersonal relationships. In: Tangney, J.P., Fischer, K.W. (eds.) Self-Conscious Emotions: Shame, Guilt, Embarrassment, and Pride, pp. 114–139. Guilford Press, New York (1995)

16. Orth, U., Robins, R.W., Soto, C.J.: Tracking the trajectory of shame, guilt, and pride across the life span. J. Pers. Soc. Psychol. 99, 1061–1071 (2010)

17. De Hooge, I.E., Breugelmans, S.M., Zeelenberg, M.: Not so ugly after all: Endogenous shame acts as a commitment device. J. Pers. Soc. Psychol. 95, 933–943 (2008)

18. Smith, R.H., Webster, J.M., Parrott, W.G., Eyre, H.L.: The role of public exposure in moral and nonmoral shame and guilt. J. Pers. Soc. Psychol. 83, 138–159 (2002)

19. Hayes, A.F., Preacher, K.J.: Statistical mediation analysis with a multicategorical independent variable. Br. J. Math. Psychol. (2013), doi: 10.1111/bmsp.12028

20. Kenny, D.A.: Mediation, http://davidakenny.net/cm/mediate.htm (accessed March 21, 2014)

21. MacKinnon, D.P., Fairchild, A.J., Fritz, M.S.: Mediation analysis. Annu. Rev. Psychol. 58, 593–614 (2007)

22. Finucane, M.L., Alhakami, A., Slovic, P., Johnson, S.M.: The affect heuristic in judgments of risks and benefits. J. Behav. Making 13, 1–17 (2000)

23. Haidt, J.: The emotional dog and its rational tail: A social intuitionist approach to moral judgment. Psychol. Rev. 108, 814–834 (2001)

24. Paxton, J., Greene, J.: Moral reasoning: Hints and allegations. Topics in Cognitive Science 2, 511–527 (2010)

Negative Mood Effects on Decision Making among Potential Pathological Gamblers and Healthy Individuals

Ivana Baldassarre, Michele Carpentieri, and Olimpia Matarazzo

Department of Psychology, Second University of Naples, Naples, Italy
{ivana.baldassarre,olimpia.matarazzo}@unina2.it,
michele.carpentieri@studenti.unina2.it

Abstract. In this study we investigated the effects of negative mood on decision making among potential pathological gamblers and healthy individuals. More specifically, we examined whether the two groups exhibited the same or different pattern of choice when being in a negative emotional state. To that end, participants were induced with negative mood through the emotional event recall technique and subsequently presented with four scenarios about monetary decision making. For each scenario they were asked to choose an option among four possibilities: two options were cautious and two risky. Results showed that negative mood affected healthy individuals and potential pathological gamblers in the opposite way: the former made more cautious choices, while the latter made more risky choices.

Keywords: Gambling, mood effects, decision making, pathological gamblers.

1   Introduction

Gambling disorder has received increasing attention from researchers over the past three decades since gambling opportunities have expanded around the world. Nosologically, the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV) characterizes Pathological Gambling (PG) as persistent and recurrent maladaptive gambling behaviour [1]. In the fifth edition of the DSM (DSM-5), published in May 2013, the diagnosis of PG was changed to Gambling Disorder (GD) [2]. The disorder was reclassified from an "Impulse-Control Disorder Not Elsewhere Classified" to one of the "Substance-Related and Addictive Disorders" in an effort to clarify the diagnosis and treatment of GD. This change also reflects the similarities between PG behaviour and addiction to substances. The prevalence of problem gambling varies across countries and cultures, with Italian rates estimated at 2.3% for youths and 2.2% for adults [3].

In regards to psychological factors underlying GD, there is a large body of evidence [4,5] showing that problem gambling is associated with social isolation and subjective distress (including depression, anxiety, and stress) in both adult and adolescent populations. It has been assumed that problem gamblers may gamble to

decrease negative emotions such as boredom or loneliness. Furthermore, anxious or depressed people may also gamble to relieve their negative emotional states [6,7,8]. However, in the short term gambling seems to be effective in reducing negative mood, but in the long term it has a rebound effect by making the pathological gamblers more anxious and depressed. Dickerson and his colleagues [9] reported that in high-frequency gamblers, prior mood and erroneous cognitions significantly accounted for continuation of gambling despite successive losses. Raghunathan and Pham [10] found that in decision making related to gambling sad individuals have a bias towards high-risk/high-reward choices, whereas anxious people have a bias towards low-risk/low-reward choices. They suggested that anxiety and sadness communicate different types of information about decision making process and, thus, influence the production of decision-makers' goals. Anxiety leads to the uncertainty reduction goals, whereas sadness leads to the rewards replacement goals. Anyway, both anxious and sad people are attempting to make a decision regarding what would make them feel better.

From a more general point of view, in the literature about the influence of emotions on cognitive processes and on individual behaviours, there are contrasting findings concerning the role of negative mood on performances [see 11 for review]. A number of studies have found that negative mood can result in more systematic, elaborate and analytical cognitive processing [12,13]. Individuals in a negative mood, compared to those in neutral or positive mood, seem to be especially likely to engage in systematic processing, to adhere to the data provided more consistently, [14] and to show less confidence in their assumptions. On the contrary, other studies showed that negative emotional states impair performances. For instance, most intense emotional states, which are accompanied by high levels of autonomic arousal, are known to impair working memory capacity [15]. This decrement in processing capacity has a variety of consequences that seem detrimental to sound reasoning. Compared to non-anxious participants, anxious participants tend to have lower ability to recall information and organize this information in memory [16], take longer to verify the validity of logical inferences [17] and select an option without considering every alternative [18].

More specifically, as far as problem gambling is concerned, trait-anxiety and stress have been linked to the development and maintenance of problem gambling. Moreover, negative mood states influence the type of gambling behaviours: in particular, depressed individuals are more likely to play games of skill, including cards, whereas stressed or anxious individuals tend to play low-skill games such as video lottery terminals [see 19 for review].

Empirical evidence is also contrasting with regard to the influence of emotional states on decision making [20,21]: emotions are shown to either increase one's attitude for risk or lead to risk avoidance. Several studies have shown that high-anxious individuals judge the risk of an event as being greater than non-anxious individuals do and choose safe options more often in tasks involving risk evaluation [22,23]. On the

contrary, when Miu *et al.* [24] investigated the effect of trait-anxiety on decision-making performances through the Iowa Gambling Task (IGT, [25]), they found that high-anxious participants showed increased selection of the risky (and disadvantageous) card decks, compared to low-anxious participants. Intense emotive and motivational states such as hunger, pain, sexual arousal, drug cravings and sleep deprivation are found to produce collapse in self-control and increase risk taking in order to alleviate those states [26]. Furthermore, in gambling literature several studies indicate that negative emotional states with strong arousal increase risk-seeking and, consequentially, risky choices [27-29]. For example, Leith and Baumeister [29] found that participants who were angry were more prone to choose economically inferior "long-shot" gambles over superior "safe-bet" gambles, whereas sad participants did not exhibit this bias. Moreover, Fessler *et al.* [27] found that anger triggered more risk-seeking in gambling, whereas Mano [28] found that intense emotional arousal increased the willingness to pay for lotteries and decreased the willingness to pay for insurance.

In sum, although research has been very interested in the influence of negative mood on the human decision-making processes, the recent data are controversial. If on the one hand several studies have shown that negative mood can result in more systematic, elaborate and analytic information processing, on the other hand other studies have shown that negative moods produce collapse in self-control and cause people to take risks in order to alleviate this state. However, as far as the gambling literature is concerned, findings are more homogeneous: negative mood seems to increase the tendency to gamble and to make risky choices.

2   Experiment

The present study has been envisaged in order to further investigate the effects of negative mood on decision making among both problem gamblers and healthy individuals. More specifically, we aimed to examine whether the two groups exhibited the same or a different pattern of choice when being in a negative emotional state. To that end, participants were induced with negative mood through the emotional event recall technique and, then, they were presented with four scenarios about monetary decision making. For each scenario, they were asked to choose an option among four possibilities: two options were cautious and two were risky. Such scenarios were built in order to depict some real life decision situations.

In order to divide participants in two groups - potential pathological gamblers and healthy individuals - the South Oaks Gambling Screen (SOGS, [30]) was used. This questionnaire consists of 20 items designed to assess the attitude toward gambling and to identify individuals who are either potential pathological gamblers or pathological gamblers.

2.1   Method

Participants. Sixty (n.= 60) individuals participated in this study, half male and half female, with mean age 32.28 (s.d.=14.26). Each participant gave informed consent and they were not paid for their participation.

Experimental design. The experimental design involved two between-subject variables – mood induction and gambling attitude – and one within-subjects variable, i.e., the four scenarios. Participants were divided into four groups on the basis of the two between-subjects variables: mood induction (negative mood and no mood induction) and attitude toward gambling (potential pathological gamblers and healthy individuals), assessed on the basis of SOGS scores.

Materials and Procedure. Participants were presented with a paper and pencil questionnaire. In the first page, a brief description of the study was presented followed by the consensus and the request to indicate their sex and age. In the second page, in negative mood condition, participants were asked to recall and describe a negative autobiographical event. Then they were asked to assess on a 7-point Likert scale (1= not at all/7= extremely) the extent to which they felt the following emotions: anger toward oneself, anger toward circumstances, regret, disappointment, fear, sadness, joy and relief. The control group participants, without mood induction, were only asked to evaluate the extent to which they felt the eight emotions.

In the third page, the four scenarios were presented, each with four choice options: two cautious and two risky options. The scenarios were the following:

-   *Casino scenario* described a situation in which the protagonist is playing roulette at a gambling casino. Red came out 4 times in a row. He has to decide whether to bet on the black and, if yes, the bet amount (100, 300 or 600 euro).

-   *Scratchcard scenario* depicted a situation in which the protagonist won 2 euro with a scratchcard. He is feeling lucky. He has to decide whether collect the winning, buy something else, or buy another scratchcard (of 2 or 5 euro).

-   *Final match scenario* described a situation in which the protagonist is going to watch the final match of his favourite team. He really desires his team to win. In that moment, his friends are betting for the winning of the team. He has to decide whether to bet and, if yes, the bet amount (5, 10 or 20 euro).

-   *Long play record selling scenario*: the protagonist is selling some of his stuff and he is forced to sell a long play record he loves. The estimated price is 200 euro and he has to establish the price at which he is willing to sell it (180, 200, 220 or 240 euro).

An example of the casino scenario with the four choice options was the following:

"It's your first time in a gambling casino. After having observed different gaming situations for about an hour, you are ready to start playing. You changed 500 euro and you have other 500 euro, which you preferred not to change in fiches. You start playing roulette and you win, now you have 600 euro in fiches. Red came out 4 times in a

row and people around suggest you to wager a large bet on black: "This is your chance!" they say. You know it's late, after this bet you have to come back home. At this point, you have to decide what to do."

The four options were: not to bet, bet 100 euro (cautious options); bet 300 euro, bet 600 euro (risky options).

Lastly, participants completed the SOGS questionnaire.

2.2   Results

In order to reduce the number of emotions, an explorative factor analysis (with principal component extraction method) was performed on the eight emotions. Varimax rotation was used after controlling the independence of the factors. Two factors were extracted with eigenvalue ≥ 1, explaining 63,02% of variance: negative emotions and positive emotions. Results are reported in Table 1. SOGS scores have been transformed in a dummy variable: potential pathological gamblers (SOGS$\geq$3; 27 participants) and healthy individuals (SOGS<3; 33 participants).

Table 1. Results of factor analysis performed on the emotions

Factor labels	Emotions	Loadings	Percent of variance	Cumulative percent of variance
Negative emotions	disappointment	.816	39.39	39.39
	anger toward circumstances	.762		
	sadness	.716		
	fear	.708		
	regret	.672		
	anger toward oneself	.648		
Positive emotions	joy	.886	23.63	63.02
	relief	.882		

In order to test whether the mood manipulation was efficacious, we conducted two ANOVA 2x2 on the two factors emerged from factor analysis (one for negative emotion and another for positive emotions) as manipulation check. Gender was put in the analysis as covariate. The independent variables were: attitude toward gambling (potential pathological gamblers vs. healthy group) and mood induction (negative vs. none). Results on negative emotions factor showed a main effect of the mood induction ($F_{1,55}$=18.02; p<.001; p-η^2=.25) and an interaction effect mood induction x gambling attitude ($F_{1,59}$=4.37; p<.05; p-η^2=.07) examined by simple effects analysis. Negative emotions were higher in negative mood than in no induction condition; in no induction condition, negative emotions were higher in gamblers than in healthy individuals, whereas no difference emerged in negative mood induction condition. Results on positive emotions factor showed an interaction effect of mood induction x gambling attitude ($F_{1,59}$=5.19; p<.05; p-η^2=.09) examined by simple effects analysis: among potential pathological gamblers, positive emotions were higher in no mood

induction than in mood induction condition; no difference emerged among healthy individuals.

In order to test the effects of the mood induction on the choices of the two participant's groups, a mixed ANOVA 2 (attitude toward gambling: potential pathological gamblers vs. healthy individuals) x 2 (mood induction: negative vs. none) x 4 (type of scenario situation: casino, scratchcard, final match, long play record) was conducted. The first two variables were between-subjects, the last was within subjects. The dependent variable was the participants' choice. Note that we dichotomized the four choices by coding as 1 the two risky option choices and as 0 the two cautious option choices. Gender (coded as dummy variable: 1 = male and 0 = female) was put in the analysis as covariate.

Mean percentages of risky choices are reported in fig. 1.
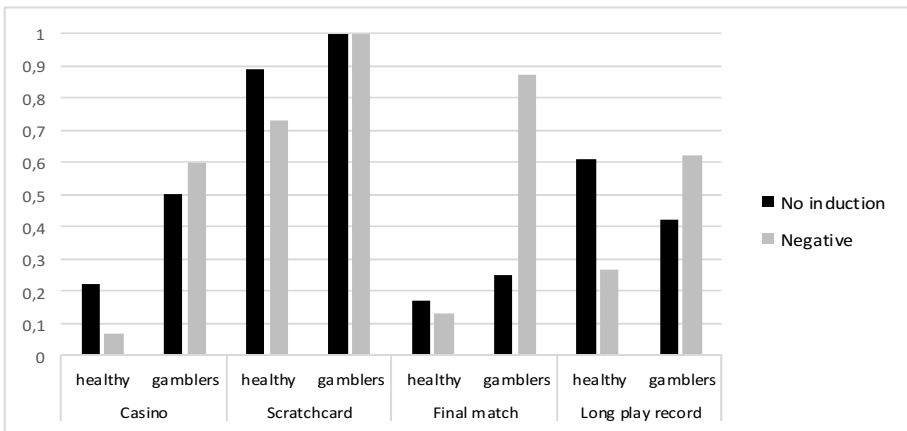
Fig. 1. Mean percentages of risky choices as function of gambling attitude (potential pathological gamblers; healthy individuals) and mood conditions

Results showed two main effects: type of scenario situations ($F_{3,165}=9.73$; $p<.001$; p-$\eta^2=.15$) and attitude toward gambling ($F_{1,55}=21.18$; $p<.001$; p-$\eta^2=.28$). Participants chose more risky options with scratchcard scenario (M=0.91; s.e.=0.4) than in the other situations. Indeed, casino scenario (M=0.35; s.e.=0.6), final match scenario (M=0.35; s.e.=0.5) and long play record selling scenario (M=0.49; s.e.=0.6) did not display any significant difference from each other. As far as the effect due to the attitude toward gambling is concerned, results showed that the participants classified as potential pathological gamblers chose more risky options (M=0.66; s.e.=0.4) than those classified as healthy individuals (M=0.39; s.e.=0.4).

Results also showed an interaction effect between mood condition and attitude toward gambling ($F_{1,55}=11.36$; $p=.001$; p-$\eta^2=.17$). The interaction effect was examined by means of simple effects analyses followed by pairwise comparisons with Bonferroni adjustment for multiple tests. The data showed that healthy individuals ($p<.05$) chose more risky options in "no mood induction" (M=0.47; s.e.=0.5) than in

"negative mood" (M=0.3; s.e.=0.6). The opposite occurred with potential pathological gamblers (p<.01), which chose more risky options in "negative mood" (M=0.78; s.e.=0.6) than in "no mood induction" (M=0.54; s.e.=0.6) condition (see fig.2).

Gender showed no significant effect.
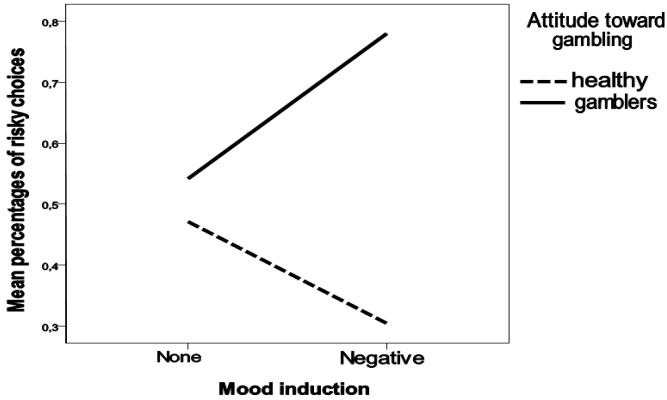
Fig. 2. Interaction effect between the attitude toward gambling (potential pathological gamblers; healthy individuals) and the mood induction (none induction; negative mood induction)

3   Conclusion

The aim of this study was to investigate the effects of induced negative mood on decision-making of both potential pathological gamblers and healthy individuals.

Results on mood manipulation showed, as expected, that negative emotions were higher in negative mood induction condition compared to the control condition, without mood induction. Furthermore, results showed that, in absence of mood induction, potential pathological gamblers felt higher levels of both negative and positive emotions than healthy individuals. This finding suggests that potential pathological gamblers are more sensitive to emotional stimuli rather than individuals without gambling problem. Williams and her colleagues [31], for example, found a positive relationship between gambling behaviour and negative emotional states; furthermore, they highlighted that individuals with gambling disorders reported a greater lack of awareness about their high levels of arousal and their emotions' intensity in comparison with healthy individuals.

Results on monetary choices showed, as expected, that potential pathological gamblers made more risky choices compared to healthy group. The only situation in which both groups of participants chose more risky options was the scratchcard scenario, a finding that is plausibly due to the scenario content effect. On the one hand, the availability of scratchcards has considerably increased in the last time period and, consequently, this scenario describes a more familiar situation as compared with the

situations of the other three scenarios. On the other hand, the monetary value of the scratchcard and of the winning (2 euro of winning and 2 or 5 euro for buying another scratchcard) was very small. So, most participants preferred to grope their luck rather than requiring the payoff.

The most interesting result obtained in this study concerns the difference in risky choices between potential pathological gamblers and healthy individuals in function of their mood condition. Healthy individuals made more risky choices in the control condition than in the negative mood condition. In contrast, potential pathological gamblers chose more risky options in the negative mood condition rather than in the control condition.

This finding suggests that healthy individuals when being in a negative mood envisage a more conservative and cautious behaviour, in conformity with the idea that negative mood can be associated with a more elaborative and analytic cognitive processing [13,14]. On the contrary, potential pathological gamblers' choices are in line with the findings that negative affective states increase risk seeking and, consequently, risky choices [31,32,33]. Perhaps such result could support the idea that some individuals gamble in order to avoid or to reduce negative emotional feelings [8]; however, in this study we did not directly investigate this hypothesis.

On the whole, the results of this study seem to be promising but more findings are required in order to better understand the relationship between emotional states and attitude toward gambling. More specifically, future research could investigate the effect of both positive and negative emotional states on both gamblers' and healthy people's decision making. In particular, the effects of definite emotions, such as anger, happiness, satisfaction, regret, sadness or guilt, should be examined in order to establish whether there is a relationship between specific emotions and the decision making outcomes.

References

1. American Psychiatric Association: Diagnostic and statistical manual of mental disorders, 4th edn. American Psychiatric Association, Washington DC (2000)
2. American Psychiatric Association: Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association, Washington DC (2013)
3. Bastiani, L., Gori, M., Colasante, E., Siciliano, V., Capitanucci, D., Jarre, P., Molinaro, S.: Complex factors and behaviors in the gambling population in Italy. J. Gambling Stud. 29, 1–13 (2013)
4. Battersby, M., Tolchard, B., Scurrah, M., Thomas, L.: Suicide ideation and behaviour in people with pathological gambling attending a treatment service. Int. J. Ment. Health Addict. 4, 233–246 (2006)
5. Petry, N.M.: Pathological gambling: etiology, comorbity and treatment. American Psychological Association, Washington DC (2005)
6. Blaszczynski, A.P., Wilson, A.C., McConaghy, N.: Sensation seeking and pathological gambling. Brit. J. Addict. 81, 113–117 (1986)
7. Blaszczynski, A.P., McConaghy, N., Frankova, A.: Boredom proneness in pathological gambling. Psychol. Rep. 67, 35–42 (1990)

8. Parke, J., Griffiths, M.: The role of structural characteristics in gambling. In: Smith, G., Hodgins, D.C., Willians, R.J. (eds.) Research and Measurement Issues in Gambling Studies, pp. 217–249. Academic Press, New York (2007)

9. Dickerson, M.G., Cunningham, R., England, S.L., Hinchy, J.: On the determinants of persistent gambling: Personality, prior mood and poker machine play. Int. J. Addict. 26, 531–548 (1991)

10. Raghunathan, R., Pham, M.T.: All negative moods are not equal: motivational influences of anxiety and sadness on decision making. Organ. Behav. Hum. Dec. 79, 56–77 (1999)

11. Pham, M.T.: Emotion and rationality: A critical review and interpretation of empirical evidence. Rev. Gen. Psychol. 11, 155–178 (2007)

12. Clore, G.L., Schwarz, N., Conway, M.: Affective causes and consequences of social information processing. In: Wyer, R.S., Srull, T.K. (eds.) Handbook of Social Cognition, 2nd edn., pp. 323–418. Erlbaum, Hillsdale (1994)

13. Weary, G., Jacobsen, J.A.: Causal uncertainty beliefs and diagnostic information seeking. J. Pers. Soc. Psychol. 98, 150–153 (1997)

14. Gasper, K.: When necessity is the mother of invention: mood and problem solving. J. Exp. Soc. Psychol. 39, 248–262 (2003)

15. Humphryes, M.S., Revelle, W.: Personality, motivation and performance - a theory of the relationship between individual-differences and information-processing. Psychol. Rev. 91, 153–184 (1984)

16. Mueller, J.H.: Effects of individual-differences in test anxiety and type of orienting task on levels of organization in free-recall. J. Res. Pers. 12, 100–116 (1978)

17. Darke, S.: Anxiety and working memory capacity. Cognition Emotion 2, 145–154 (1988)

18. Keinan, G.: Decision-making under stress: Screening of alternative under controllable and under controllable threats. J. Pers. Soc. Psychol. 52, 639–644 (1987)

19. Bagby, M.R., Vachon, D.D., Bulmash, E., Quilty, L.C.: Personality disorders and pathological gambling: a review and re-examination of prevalence rates. J. Pers. Disord. 22, 191–207 (2008)

20. Bechara, A., Damasio, H., Damasio, D.A.R.: Emotion, decision making and the orbital cortex. Cereb. Cortex 10, 295–307 (2000)

21. Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N.: Risk as feelings, Psychol. Psychol. Bull. 127, 267–286 (2001)

22. Hockey, G.R.J., Maule, A.J., Clough, P.J., Bdzola, L.: Effects of negative mood states on risk in everyday decision making. Cognition Emotion 14, 823–855 (2000)

23. Maner, J.K., Anthony, R., Cromer, K., Mallott, M., Lejuez, C.W., Joiner, T.E., Schmidt, N.B.: Dispositional anxiety and risk-avoidant decision-making. Pers. Indiv. Differ. 42, 665–675 (2006)

24. Miu, A.C., Heilman, R.M., Houser, D.: Anxiety impairs decision-making: Psychophysiological evidence from an Iowa Gambling Task. Biol. Psychol. 77, 353–358 (2008)

25. Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W.: Insensivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7–15 (1994)

26. Loewenstein, G.: Out of control: Visceral influences on behavior. Organ. Behav. Hum. Dec. 65, 272–292 (1996)

27. Fessler, D.M.T., Pillsworth, E.G., Flamson, T.J.: Angry men and disgusted women: an evolutionary approach to the influence of emotions on risk taking. Organ. Behav. Hum. Dec. 95, 107–123 (2004)

28. Mano, H.: Risk-taking, framing effects and affect. Organ. Behav. Hum. Dec. 57, 38–58 (1994)
29. Leith, K.P., Baumeister, R.F.: Why do bad moods increase self-defeating behavior? Emotion, risk taking and self regulation. J. Pers. Soc. Psychol. 71, 1250–1267 (1996)
30. Lesieur, H.R., Blume, S.B.: The South Oaks Gambling Screen (SOGS): A new instrument for the identification of pathological gamblers. Am. J. Psychiat. 144, 1184–1188 (1987)
31. Williams, A.D., Grisham, J.R., Erskine, A., Cassedy, E.: Deficits in emotion regulation associated with pathological gambling. Brit. J. Clin. Psychol. 51, 223–238 (2012)

Deep Learning Our Everyday Emotions
A Short Overview

Björn Schuller

Imperial College London, Department of Computing, SW7 2AZ London U. K.
bjoern.schuller@imperial.ac.uk

Abstract. Emotion is omnipresent in our daily lives and has a significant influence on our functional activities. Thus, computer-based recognising and monitoring of affective cues can be of interest such as when interacting with intelligent systems, or for health-care and security reasons. In this light, this short overview focuses on audio/visual and textual cues as input feature modality for automatic emotion recognition. In particular, it shows how these can best be modelled in a Neural Network context. This includes deep learning, and sparse auto-encoders for transfer learning of a compact task and population representation. It further shows avenues towards massively autonomous rich multitask-learning and required confidence estimation as is needed to prepare such technology for real-life application.

Keywords: Deep Learning, Neural Networks, Emotion Recognition, Affective Computing.

1 Introduction

Emotion is omnipresent in our daily lives and has a significant influence on our functional activities. Thus, computer-based recognising and monitoring of affective cues can be of interest such as when interacting with intelligent systems, or for health-care and security reasons.

Machine Learning aspects have been one of the major factors to boost the performance of automatic emotion recognition ever since along side search for the 'ultimate feature representation', general emotion representation such as by categories (such as one or several emotion labels or tags per instance of analysis) or dimensions, units of analysis such as frames, 'speaker turns' or longer sequences and finally data. In fact, any of the latter obviously impacts on the machine learning architecture. In this short overview, a discussion will be provided on recent trends related to machine learning aspects for automatic emotion recognition with a slight emphasis on neural approaches. Starting with deep learning in section 1 where some important 'tweaks' will be highlighted, section 2 will then focus on the title's 'everyday' aspect by touching upon lifelong learning and its implications. A short conclusion will be provided at the end.

2   Deep Learning

Recently, there is an increasing tendency in the field of machine learning to use deep learning methods for various kinds of applications owing to their great success in improving accuracies. A comprehensive recent overview on the topic is found in [[30]]. Having been impressively applied to various kinds of speech, video, and image analysis tasks, it is not surprising that deep learning has found its way also into automatic emotion recognition. A first example is emotion recognition from speech [[32]]. In the works of Brueckner et al. further related speaker states and traits from the ISCA Interspeech Computational Paralinguistics Challenges have been considered – often outperforming the best results obtained in those [[3,4,5,6]]. Further examples for emotional speech recognition include [[7,25,26,1,21,29]]. In a related way, deep learning has also been successfully applied to emotion recognition in music [[31]]. Further, a number of works combine video cues such as facial information with speech analysis in a deep learning paradigm, such as in [[24]] or in the winning contribution to the 2013 Emotion in the Wild Challenge held at ACM ICMI [[23]] which was able to raise the organisers' baseline of 27.5 % accuracy to 41.0 %. Manifold further articles report gains in facial expression analysis, including even lip shape analysis for emotion recognition [[28]]. Also physiological measurements such as by EEG data analysis [[22]] have recently been implemented successfully in deep learning architectures beyond further studies dealing with textual and other cues.

In fact, deep neural networks (DNNs) composed of multiple hidden layers were first suggested decades ago. However, their training was difficult. Neural networks are usually learnt by stochastic gradient descent (SGD) such as the well-known backpropagation algorithm. Yet, having large initial weights, the network parameters tend to converge towards poor local minima. On the other hand, having small initial weights tends to make the gradients in the lower layers vanish. Accordingly, training networks with many hidden layers becomes challenging. In addition, the paramter space for deep networks with many hidden layers and many hidden units in each layer is often large making it likely that the networks overfit the data sets. This is crucial in the field of emotion recognition where data sets are relatively small. Hinton et al. [[20]] helped to overcome these limitations by an efficient method to pre-train DNNs layer by layer. Originally, an undirected graphical model – the Restricted Boltzmann Machine (RBM) – was used. Interestingly, this pre-training is entirely unsupervised, i. e., without having target labels. Such pre-training moves the network parameters near to a local optimum in the parameter space. Later, the parameters can further be optimised by (supervised) iterations of SGD on the pre-trained network. This fine-tunes the network to the task at hand – emotion recognition in our case.

2.1   Deep Belief Networks

Once the weights of an RBM or suited alternative such as a denoising auto-encoder (DAE) network has been learnt, the outputs of the hidden nodes can be used as input data for training a 'next' RBM or similar which will learn

a more complex representation of the input data to step-wise establish a deep belief network (DBN) by stacking layers. Such layer-wise construction of a deep generative model is known to be considerably more efficient than learning all layers at once.

Importantly, in DBNs hidden states can be inferred very efficiently by a single bottom-up pass in which the top-down generative weights are used in the reverse direction. Further, with each added layer of learnt features added, the new DBN has a variational lower bound on the log probability of the training data that is better than the variational bound for the previous DBN, provided correct learning [[20]]. For many applications, discriminative fine-tuning of a such an initialised neural network leads to better results than the same neural network initialised with small random weights [[14]]. In fact, greedy layer-wise unsupervised pre-training is crucial in deep learning by introducing a useful prior to the supervised fine-tuning training procedure [[14]]. While a DBN is a generative model consisting of several RBM layers, it can be used to initialise the hidden layers of a standard feed-forward DNN. One then adds an output layer such as a *softmax* layer for (emotion) classification or a *linear* layer for (emotion) regression. Note that the terms DBN and DNN are often used interchangeably.

2.2 Dropout

To overcome overfitting, dropout was introduced [[19]] to prevent complex co-adaptations in which a hidden unit is only helpful in the context of several other specific hidden units by randomly omitting each hidden unit from the network. This is done with a given probability, such that a hidden unit cannot rely on 'the other' hidden units being present. This can be seen as equivalent to adding (particular) noise to the hidden unit activations during the forward pass in training, similar to [[33]]. However, dropout can be used in all hidden and input layers of a network and also during the final fine-tuning. While dropout strongly reduces overfitting, it increases the training time.

2.3 Rectified Linear Units

The key computational unit in a neural network is a linear projection followed by a point-wise non-linearity. The latter is often chosen as a logistic sigmoid or tanh function. Alternatively, the recently proposed *rectified linear unit* (ReLu) can improve generalisation and make training of deep networks faster and simpler [[27,35]]. It is linear when its input is positive and zero otherwise. If it is activated above zero, its partial derivative is one. Accordingly, vanishing gradients do not exist along paths of active hidden units in an arbitrarily deep network. Furher, they saturate at exactly zero. This can be useful when using hidden activations as input features for a classifier.

3 Life-Long Learning

Still to the present day, likely the major bottleneck in automatic emotion recognition systems is the lack of training data. This is likely to be overcome only

by using efficient manners of label acquisition. Moreover, ideally emotion recognition systems of the next generation will keep learning, e.g., in targeted and efficient interaction with their users or by exloiting sources of (rich) data such as the Internet or television, etc.

3.1   Transfer Learning

Even more efficiently, existing resources can be reused to 'transfer' knowledge across these different factors such as learning from adult material how to recognise emotions of children or elderly, etc., e..g., by DAE neural networks [[12,11,13]]. In fact, emotional data resources often come in very different labellings such as (often different) categories or dimensions – transferring by suited approaches of machine learning can also be of help in this respect. Further, it has been shown that this can go as far as transfering similarities of emotion as manifested across speech and music [[8]] or speech and sound [[34]].

3.2   Collaborative Learning

In fact, it is not the data that is usually sparse, but rather the labels. Thinking of the need to acquire data from different cultural backgrounds, in different languages, of different nature such as acted, elicited, masked, spontaneous, etc., of different parts of the population such as from children or aged persons, and covering also less researched states such as social emotions, one can barely imagine the effort that would need to go into a purely human-based data acquisition. It thus appears wise to include the computer systems in a 'cooperative learning' approach – ideally mixed with dynamic crowd sourcing to cover for experts of different cultures, languages, etc. Active learning [[37,17]] in combination with semi-supervised learning [[38,15]] provide a good basis to this end by allowing a machine to decide whether it can label new data itself, needs human aid, or can discard it as it seems not to be of sufficient interest. The first aspect, i.e., being sufficiently confident, can be based on the computation of suited confidence measures as will be touched upon in the next subsection. The second aspect, i.e., if the new data instance is of sufficient interest, can be decided based upon sparseness of the instance, where sparse examples naturally appear of particular interest. Sparseness can thereby be found in all sorts of different ways as mentioned at the beginning of this subsection, e.g., coming from a sparse emotion class, subject group, culture, etc. Further methods are of more technical nature such as the expected change in model parameters given the data instance would be used in learning. Ideally, such request for human aid could be scaled for most efficiency, such as making a decision on how many human opinions need to be casted, e.g., by crowd-sourcing. As an example, a machine could weight labels by itself and raters based on confidence measures and agreement 'so far' to decide on whether a further opinion will be needed.

3.3   Confidence Measures

For the above mentioned collaborative learning, meaningful measures of confidence in an emotion recognition engine's decision need to be established [[36]]. While there have been very few works addressing this topic in particular for the recognition of emotion by computer systems, there are some first successes reported. These base on predicting the human labelers' agreement (rather than the emotion) and good correlation between this estimated agreement of humans and the confidence one can have in the machine's prediction of the emotion per se where observed [[9]]. A second alternative trains systems on other emotional speech databases to learn in a semi-supervised manner to predict potential mistakes of the target emotion recognition system of interest [[10]].

3.4   Distributed Learning

In order to be able to collect large amounts of data from users while preserving privacy of these, distributed processing may become crucial. This allows to collect large amounts of data on a server for the update of recognition models. Seemingly, data can be reduced considerably without too big losses in recognition accuracy, but high gains in privacy protection due to reduced feature information by (split) vector quantisation [[18,2]]. In addition, distributed learning can allow for highly efficient processing.

3.5   Multitask Learning

Finally, training multiple emotion dimensions [[16]] or person state and trait informtion in parallel has been shown to boost performance for each individual task. This can be easily implemented in (deep) neural architectures by adding further output nodes to a network.

4   Conclusion

In this overview on selected recent machine learning trends in the automatic recognition of human emotions deep learning was discussed as a promising future direction in a field where unity on the best learning approach is entirely missing. Further, ways to reach a life-long learning approach have been touched upon. Overall, one can expect life-long deep learning emotion recognisers that exploit 'big' amounts of data as a likely future solution to make these systems ready for everyday usage.

References

1. Amer, M.R., Siddiquie, B., Richey, C., Divakaran, A.: Emotion Detection in Speech using Deep Networks. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013), Florence, Italy. IEEE (2013)
2. Bennett, I.: Emotion detection device and method for use in distributed systems. US Patent 8,214,214 (July 3, 2012)
3. Brückner, R., Schuller, B.: Likability Classification – A not so Deep Neural Network Approach. In: Proceedings of the INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association, ISCA, Portland, OR, 4 pages (September 2012)
4. Brückner, R., Schuller, B.: Hierarchical Neural Networks and Enhanced Class Posteriors for Social Signal Classification. In: Proceedings 13th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2013, Olomouc, Czech Republic, 6 pages. IEEE (December 2013)
5. Brückner, R., Schuller, B.: Being at Odds? – Deep and Hierarchical Neural Networks for Classification and Regression of Conflict in Speech. In: Poggi, I., D'Errico, F., Vinciarelli, A. (eds.) Conflict and Negotiation: Social Research and Machine Intelligence. Computational Social Sciences. Springer, Heidelberg (2014)
6. Brückner, R., Schuller, B.: Social Signal Classification Using Deep BLSTM Recurrent Neural Networks. In: Proceedings of the 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014, Florence, Italy, pp. 4856–4860. IEEE (May 2014)
7. Cibau, N.E., Albornoz, E.M., Rufiner, H.L.: Speech emotion recognition using a deep autoencoder. In: Proceedings of the XV Reunión de Trabajo en Procesamiento de la Información y Control (RPIC 2013), San Carlos de Bariloche (2013)
8. Coutinho, E., Deng, J., Schuller, B.: Transfer Learning Emotion Manifestation Across Music and Speech. In: Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN) as part of the IEEE World Congress on Computational Intelligence (IEEE WCCI), Beijing, China, p. 6. IEEE (July 2014)
9. Deng, J., Han, W., Schuller, B.: Confidence Measures for Speech Emotion Recognition: a Start. In: Fingscheidt, T., Kellermann, W. (eds.) Proceedings of Speech Communication; 10. ITG Symposium, Braunschweig, Germany, pp. 1–4. ITG, IEEE (2012)
10. Deng, J., Schuller, B.: Confidence Measures in Speech Emotion Recognition Based on Semi-supervised Learning. In: Proceedings of the INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association, ISCA, Portland, OR, 4 pages. ISCA (September 2012)
11. Deng, J., Xia, R., Zhang, Z., Liu, Y., Schuller, B.: Introducing Shared-Hidden-Layer Autoencoders for Transfer Learning and their Application in Acoustic Emotion Recognition. In: Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014, Florence, Italy, May 2014, pp. 4851–4855. IEEE (2014)
12. Deng, J., Zhang, Z., Marchi, E., Schuller, B.: Sparse Autoencoder-based Feature Transfer Learning for Speech Emotion Recognition. In: Proc. 5th Biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2013), Geneva, Switzerland, pp. 511–516. HUMAINE Association, IEEE (2013)

13. Deng, J., Zhang, Z., Schuller, B.: Linked Source and Target Domain Subspace Feature Transfer Learning – Exemplified by Speech Emotion Recognition. In: Proceedings 22nd International Conference on Pattern Recognition (ICPR 2014), Stockholm, Sweden, pp. 761–766. IAPR (August 2014)

14. Erhan, D., Bengio, Y., Courville, A., Vincent, P.-A.M.P., Bengio, S.: Why Does Unsupervised Pre-training Help Deep Learning? The Journal of Machine Learning Research 11, 625–660 (2010)

15. Esparza, J., Scherer, S., Schwenker, F.: Studying Self- and Active-Training Methods for Multi-feature Set Emotion Recognition. In: Schwenker, F., Trentin, E. (eds.) PSL 2011. LNCS, vol. 7081, pp. 19–31. Springer, Heidelberg (2012)

16. Eyben, F., Wöllmer, M., Schuller, B.: A Multi-Task Approach to Continuous Five-Dimensional Affect Sensing in Natural Speech. ACM Transactions on Interactive Intelligent Systems, Special Issue on Affective Interaction in Natural Environments 2(1), 29 (2012)

17. Han, W., Li, H., Ruan, H., Ma, L., Sun, J., Schuller, B.: Active Learning for Dimensional Speech Emotion Recognition. In: Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association, Lyon, France, pp. 2856–2859. ISCA (August 2013)

18. Han, W., Zhang, Z., Deng, J., Wöllmer, M., Weninger, F., Schuller, B.: Towards Distributed Recognition of Emotion in Speech. In: Proceedings 5th International Symposium on Communications, Control, and Signal Processing, ISCCSP 2012, Rome, Italy, pp. 1–4. IEEE (May 2012)

19. Hinton, G., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Improving neural networks by preventing co-adaptation of feature detectors. CoRR, abs/1207.0580 (2012)

20. Hinton, G.E., Osindero, S., Teh, Y.-W.: A fast learning algorithm for deep belief nets. Neural Computation 18(7), 1527–1554 (2006)

21. Huang, C., Gong, W., Fu, W., Feng, D.: A Research of Speech Emotion Recognition Based on Deep Belief Network and SVM. Mathematical Problems in Engineering, Article ID 749604, 7 (2014)

22. Jirayucharoensak, S., Pan-Ngum, S., Israsena, P.: EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation. The Scientific World Journal, Article ID 627892, 10 (2014)

23. Kahou, S.E., Pal, C., Bouthillier, X., Froumenty, Gülcehre, P., Memisevic, R., Vincent, P., Courville, A., Bengio, Y.: Combining Modality Specific Deep Neural Networks for Emotion Recognition in Video. In: Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI 2013), Sydney, Australia, pp. 543–550. ACM (2013)

24. Kim, Y., Lee, H., Provost, E.M.: Deep learning for robust feature generation in audio-visual emotion recognition. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013), Vancouver, Canada. IEEE (2013)

25. Le, D., Provost, E.: Emotion recognition from spontaneous speech using Hidden Markov models with deep belief networks. In: 2013 IEEE Workshop on Proceedings Automatic Speech Recognition and Understanding (ASRU), pp. 216–221. IEEE, Olomouc (2013)

26. Li, L., Zhao, Y., Jiang, D., Zhang, Y., Wang, F., Gonzalez, I., Valentin, E., Sahli, H.: Hybrid Deep Neural Network - Hidden Markov Model (DNN-HMM) Based Speech Emotion Recognition. In: Proceedings Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2013). IEEE, Geneva (2013)

27. Maas, A., Hannun, A., Ng, A.: Rectifier Nonlinearities Improve Neural Network Acoustic Models. In: Proc. of ICML Workshop on Deep Learning for Audio, Speech, and Language Processing, WDLASL, Atlanta, GA, USA (June 2013)

28. Popović, B., Ostrogonac, S., Delić, V., Janev, M., Stanković, I.: Deep Architectures for Automatic Emotion Recognition Based on Lip Shape. Infotech-Jahorina 12, 939–943 (2013)

29. Sánchez-Gutiérrez, M.E., Albornoz, E.M., Martinez-Licona, F., Rufiner, H.L., Goddard, J.: Deep Learning for Emotional Speech Recognition. In: Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., Olvera-Lopez, J.A., Salas-Rodríguez, J., Suen, C.Y. (eds.) MCPR 2014. LNCS, vol. 8495, pp. 311–320. Springer, Heidelberg (2014)

30. Schmidhuber, J.: Deep Learning in Neural Networks: An Overview. Technical Report IDSIA-03-14, IDSIA, Lugano, Switzerland (2014)

31. Schmidt, E.M., Kim, Y.E.: Learning Emotion-based Acoustic Features with Deep Belief Networks. In: Proceedings 2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, pp. 65–68. IEEE (2011)

32. Stuhlsatz, A., Meyer, C., Eyben, F., Zielke, T., Meier, G., Schuller, B.: Deep Neural Networks for Acoustic Emotion Recognition: Raising the Benchmarks. In: Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011, pp. 5688–5691. IEEE, Prague (2011)

33. Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.-A.: Extracting and composing robust features with denoising autoencoders. In: Proc. of ICML, New York, NY, USA, pp. 1096–1103 (2008)

34. Weninger, F., Eyben, F., Schuller, B.W., Mortillaro, M., Scherer, K.R.: On the Acoustics of Emotion in Audio: What Speech, Music and Sound have in Common. Frontiers in Psychology, Emotion Science, Special Issue on Expression of Emotion in Music and Vocal Communication 4(Article ID 292), 1–12 (2013)

35. Zeiler, M., Ranzato, M., Monga, R., Mao, M., Yang, K., Le, Q.V., Nguyen, P., Senior, A., Vanhoucke, V., Dean, J., Hinton, G.: On Rectified Linear Units for Speech Processing. In: ICASSP, Vancouver, Canada, May 2013, pp. 3517–3521. IEEE (2013)

36. Zhang, Z., Deng, J., Marchi, E., Schuller, B.: Active Learning by Label Uncertainty for Acoustic Emotion Recognition. In: Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association, Lyon, France, pp. 2841–2845. ISCA (August 2013)

37. Zhang, Z., Schuller, B.: Active Learning by Sparse Instance Tracking and Classifier Confidence in Acoustic Emotion Recognition. In: Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association, Portland, OR, p. 4. ISCA (September 2012)

38. Zhang, Z., Weninger, F., Wöllmer, M., Schuller, B.: Unsupervised Learning in Cross-Corpus Acoustic Emotion Recognition. In: Proceedings 12th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2011, pp. 523–528. IEEE, Big Island (2011)

Extracting Style and Emotion from Handwriting

Laurence Likforman-Sulem[1], Anna Esposito[2,3],
Marcos Faundez-Zanuy[4], and Stéphan Clémençon[1]

[1] Institut Mines-Telecom/Télécom ParisTech & CNRS LTCI, 46 rue Barrault,
75013 Paris, France
[2] Second University of Naples, Caserta, Naples, Italy
[3] International Institute for Advanced Scientific Studies-IIASS, Salerno, Italy
[4] Escola Universitaria Politecnica de Mataro, TecnoCampus Mataro-Maresme, Spain
laurence.likforman@telecom-paristech.fr

Abstract. Humans produce handwriting samples in their every-day life
activity. Handwriting can be collected on-line, thanks to digitizing tablets,
or off-line by scanning paper sheets. One question is whether such hand-
writing data can tell about individual style, health or emotional state of
the writer. We will try to answer this question by presenting related works
conducted on on-line and off-line data. We will show that pieces of off-line
handwriting may be useful for recognizing a writer or a handwriting style.
We will also show that anxiety can be recognized as an emotion from on-
line handwriting and drawings using a non parametric approach such as
random forests.

Keywords: Anxiety recognition, Style recognition, Emotion recogni-
tion, Random forests.

1 Introduction

Handwriting is still a popular means of communications. Handwriting data can
be collected either on-line [1], off-line [2] or both [3]. On-line handwriting consists
in recording the way how strokes are drawn including stroke order, direction and
speed (i.e the ductus), along with pen position and pressure. In addition in-air pen
positions are also recorded, i.e. when the pen is near the tablet though not touching
it. From such on-line data, writer gestures can be recovered as seen in Fig. 1-a. In
contrast, with off-line handwriting collection, writer gestures have been lost and
data consist in grey level or black and white images (see Fig. 1-b).

Since handwriting is easily produced, both on-line and off-line handwriting
have been extensively studied. Reading systems are able to process whole text-
lines with Hidden Markov Model (HMMs) and Recurrent Neural Network (RNN)
approaches [4][5]. Thus the recognition of handwritten envelopes, checks and
mail is possible for postal, banking and mail management applications. The
recognition of historical documents is also possible when the writing is regular
[6]. For irregular writings, keyword spotting approaches are preferable [7][8].
Besides recognition, writer authentication or identification can be achieved from
pieces of handwriting, using similarity measures and Kohonen Maps [9][10][11].
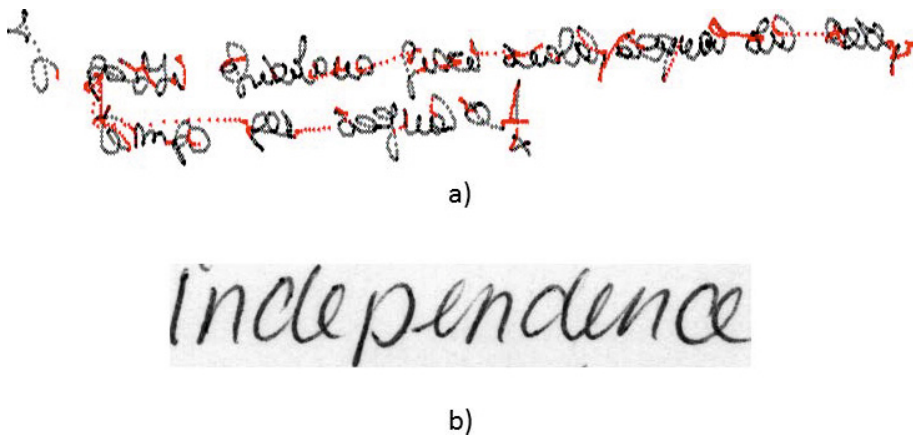
Fig. 1. a) On-line handwriting sample with on-paper (black) and in-air (red) points. b) off-line handwriting sample: grey level image (from [2]).

Other applications consist in separating handwriting from printed items within a document as in the MAURDOR challenge [12].

In another perspective, clinicians have observed that writing abilities can be modified with age, fatigue or illnesses [13]. To assist clinicians in their diagnosis, computerized platforms have been developed which collect and make measurements on handwriting. Among them,

- the ClockMe system [14] which automatically evaluates the score of the clock drawing test (CDT).
- The Compet system [15] which can predict age, and various health disorders such as major depression from measures extracted from handwriting tasks such as text copy.
- The 'strokes against stroke' system, which detects brain stroke risk from stroke measurement achieved on a tablet according to stimuli displayed on a screen [16] [17].

Other systems deal with detecting various motor disorders due to Parkinson (PD), Alzheimer (AD), sclerosis [18] or schizophrenia [19] deseases.

In the following we will detail research works related to style and emotion recognition.

2   Style Recognition

Handwriting style results from a combination of several factors such as references to glyphs learnt at school, personal habits and motor control abilities [20]. Handwriting style may refer to writer style since each handwriting is unique. Handwriting characteristics at the motor or ink level are for instance studied in

[21][22]. Handwriting style may also refer to broad classes of handwriting, including clusters of writings which look similar, but from different writers . This is the case for paleographical studies which mainly consist in comparing documents, looking for similarities and disimilarities. The comparison is based on features such as layout (spacings) and stroke angles. From an automatic processing point of view, documents or small pieces of writing, can be represented by such angle features. Among angle-based representations: directional histograms [23], matrix of curvatures occurences [24], and fractal dimension [25]. We have proposed [26] a document representation based on probability density functions (denoted as *pdfs* in the following). The probability density functions of a document is obtained by collecting angle values of pixel contours. The angles are obtained with a bank of Gaussian filters. The histogram of these angle values are smoothed and normalized as a probability density function. We propose several measures on these *pdf* curves in order to compare documents. The main ones are:

- Mode: refers to the principal mode and is indicative of the preferential slant direction of the script.
- Entropy: is the entropy of the pdf curve and varies mostly with curliness and stroke density. The higher the entropy, the flatter the pdf curve. A flat curve coresponds to a round writing. Conversely, a low entropy corresponds to a peaked curve and a writing with a dominant direction.
- Full amplitude: is the value of the pdf curve at the location of the mode. The higher the amplitude, the more linear is the writing.
- Modality: represents the number of modes in the pdf curve and relates perceptually to the presence of multiple strong writing directions.
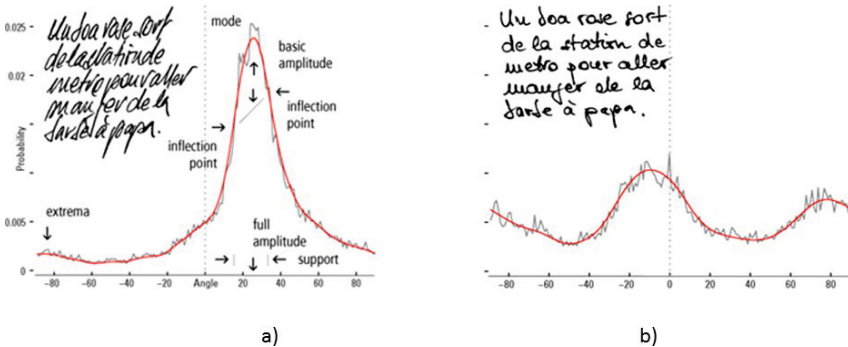
Fig. 2. Histogram of pixel angle values and corresponding probability density function curve (in red). Measurements on the curve (from [26]).

From one document, documents in a collection can be ranked according to each of these measures. Thus the more similar documents according to a chosen measure are retrieved by such comparison. This approach can be used to

retrieve documents from the same writer, or documents which share style characteristics with the query document, such as the dominant writing direction. For this task, the authors in [26] developed a demonstrator called REX, available at http://glyph.telecom-paristech.fr/.

3 Anxiety Recognition

As defined by Eysenck et al. [27](pp.336) *"Anxiety is an aversive emotional and motivational state occurring in threatening circumstances"*. It is a secondary emotion [28] which affects cognition and reduces the individual's effectiveness and efficiency in performing cognitive tasks. It is thus desirable to assess anxiety.

The Depression Anxiety Stress Scale (DASS) is a popular tool for assessing anxiety, along with other disorders such as depression and stress. It consists in filling a questionnaire including 42 questions such as "within the past 2 weeks, I had a feeling of shakiness (e.g. legs going to give away)". DASS provides a score for qualifying the anxiety level as normal, mild, moderate or severe.

Even though the DASS psychometric properties are well assessed [29], it is desirable to relate its properties to daily cognitive functional activities, such as handwriting, in order to automatize the discrimination process (supporting the clinician's diagnose and reducing health care's costs) between clinical and non-clinical indicators of the abovementioned disorders.

This can be down at low cost with non harmful device by exploiting handwriting as mentioned in Section 1. Thus we propose to build an automatic system which relates handwriting to emotion such as anxiety.

3.1 Data Collection

For building an emotion recognition system from handwriting, we need labeled data. This consist in collecting handwriting samples from subjects. DASS scores have been used for labeling the emotional state of each subject. In this study, we have collected samples from 50 subjects (from which two had to be discarded). The emotional states have been dichotomized into non anxious (normal DASS scale) and anxious (DASS moderate to severe).

The handwriting samples have been collected from an ensemble of five handwriting/drawing tasks, extracted from seven exercices completed on a sheet of paper laid on the digitizing tablet (see Fig. 3): pentagons, house drawings, handprint writing, circles (right hand, left hand), clock drawing, cursive writing. Circles were discarded because these exercises did not include in air strokes.

3.2 Extracting Features on Handwriting

The second step of a recognition system consists in extracting features for each subject. Twenty-five features f_1 to f_{25} (five tasks, five features per tasks), are extracted for each subject from his/her handwriting. The five features extracted from each task can be divided into time-based and ductus-based features, as listed below:

- f_1: time spent in-air while completing the task
- f_2: time spent on-paper while completing the task
- f_3: time to complete the whole task

- f_4: number of on-paper strokes [1] while completing the task
- f_5: number of in-air strokes while completing the task

Features f_1 to f_5 correspond to the pentagon task. Features f_6 to f_{25} correspond to the remaining tasks in the following order: house drawing, handprint, clock, cursive writing.

We thus extract both on-paper and in-air features per subject and we aim at recognizing subject's emotional state (anxious or not) from these features. In the following we use random forests both for recognizing anxiety from features but also to improve recognition by selecting the best features.
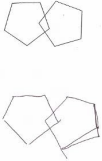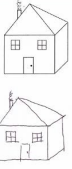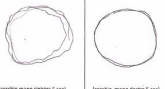
3.3 Selection of Best Features with Random Forests

We propose to recognize anxiety from these handwriting data with a non-parametric machine learning approach, namely random forests [30]. The main advantage with a machine learning approach is that measurements are not analysed in isolation but in combination. This is useful when classes overlap or when a variable has several modes. In addition, we need not to assume Gaussian distribution for the extracted measures as for parametric approaches. Moreover such non parametric approach is compliant with our scarce emotional data extracted from 48 subjects. Random forests also have the advantage of ranking the input features according to their importance. It thus provides cues for interpreting the recognition process in terms of relevant tasks and corresponding features.

Training a random forest consists in building N decision trees which are combined at decision level. Each tree is built from a subset of the training data (the remaining set includes the out-of-bag points or OOBs) and from a subset of the features. We use the four importance measures provided through random forest training. The ranks of the 25 features differ according to each measure. The automatic ranking process is the following. The ranks of each feature according to each importance measure are summed up. The lowest the sum, the best the feature.

For anxiety recognition, the seven most relevant features are the following:

- timing-based f_1 and f_6 features (time spent in-air) when completing pentagon and house drawing tasks.
- ductus-based f_9, f_{10} and f_{24} features which are the number of on paper and in air strokes when completing the house drawing and cursive writing tasks.
- timing-based f_2 feature (time spent on paper) when completing the pentagon drawing task.
- timing-based f_{13} feature (total time for a task) when completing the handprint writing task.

Fig. 3. top: whole set of tasks filled by one user. bottom: Sample of the pentagon drawing task with on paper (in black) and in-air points (in red). In blue, extrapolated lines between two in-air strokes.

Forests used in these experiments are built with $ntree = 10$ trees. But this number can largely be increased. Since we have $nfeat = 25$ features, $mtry = \sqrt{nfeat} = 5$ is a popular tuning for the number of variables considered at each tree node. We use the R-language and random forest package [31].

Our first finding is that a majority of features selected by the random forest are related to in air strokes (f_1, f_6, f_9, f_{24}). This shows the importance of in-air

---

[1] A stroke in this context corresponds to consecutive drawing points achieved without lifting the pen.

Table 1. Number of (best) features for each task

Tasks	Pentagons	House	Handprint	Clock	Cursive
Number of features	2	3	1	0	1

Table 2. Leave-one-out accuracies. Anxiety is recognized either using the whole set of features, or a subset of best features

Anxiety recognition	All features	Best features
Accuracy	62.5 %	68.7 %

motions, which cannot be observed in the ink trace. Drawing is a less automatic task than copying a text. Thus drawing tasks are more related to emotion affects as shown in Table 1. However both handprint or cursive writings are useful for complementing drawing tasks.

To provide results for anxiety recognition, we conduct one-leave-out cross validation experiments. At each run, a forest is built from all but one data point, and from the whole set of features. This forest is used to test the remaining data point. From the 48 runs, 48 classification results are collected and the corresponding accuracy is shown in Table 2 (column "All features"). The second cross validation experiment provides recognition accuracy when a subset of features is selected. At each run, the selection process provides the best features from 47 data points. A second random forest is then built from the selected features and tested on the remaining data point. Recognition accuracy is provided in Table 2 (column "Best features"). It can be noted that the subset of selected features slightly varies from one run to another.

The increase in accuracy when selecting features is greater than 6 % in absolute value. This shows the importance of selecting features within a classifier, and the efficiency of the selected features.

4 Conclusion

Writing and drawing can be easily collected through on-line device or scanned sheets. The ink trace can be characterized locally by angle values or globally by measurements on handwriting representations (such as probability density functions), and related to the style of the writing.

There is less control on motions necessary to produce a pattern than on the pattern itself. In a second part, we have assumed that emotion affects handwriting characteristics. We have proposed a machine-learning approach, namely random forests, in order to recognize secondary emotions such as anxiety. The assessment of anxiety was made by exploiting the DASS questionnaire. The collected database, called EMOTHAW for EMOTion recognition through HAndWriting will be made publicly available in the future.

Handwriting may be useful to recognize other emotions such as those evaluated through the DASS scale. Besides disorders, handwriting may be also useful for evaluating healing during physical rehabilitation, in an autonomous way [32].

References

1. Guyon, I., Schomaker, L., Plamondon, R., Liberman, R., Janet, S.: Unipen project of online data exchange and recognizer benchmarks. In: International Conference on Pattern Recognition ICPR, pp. 29–33 (1994)
2. Marti, U., Bunke, H.: A full english sentence database for off-line handwriting recognition. In: Proc. Int. Conf. on Doc. Analysis and Recognition, pp. 705–708 (1999)
3. Viard-Gaudin, C., Lallican, P.M., Binter, P., Knerr, S.: The ireste on/off (ironoff) dual handwriting database. In: ICDAR, pp. 455–458 (1999)
4. Bianne-Bernard, A.L., Menasri, F., El-Hajj, R., Mokbel, C., Kermorvant, C., Likforman-Sulem, L.: Dynamic and contextual information in HMM modeling for handwritten word recognition. IEEE PAMI 99(10), 2066–2080 (2011)
5. Morillot, O., Likforman-Sulem, L., Grosicki, E.: New baseline correction algorithm for text-line recognition with bidirectional recurrent neural networks. Journal of Electronic Imaging 22(2), 23028–23028 (2013)
6. Toselli, A.H.: Reconocimiento de Texto Manuscrito Continuo. PhD thesis, Departamento de Sistemas Informaticos y Computacion, Universidad Politecnica de Valencia (2004)
7. Vasilopoulos, N., Kavallieratou, E.: A classification-free word-spotting system. In: DRR (2013)
8. Leydier, Y., Lebourgeois, F., Emptoz, H.: Text search for medieval manuscript images. Pattern Recognition 40(12), 3552–3567 (2007)
9. Faúndez-Zanuy, M., Hussain, A., Mekyska, J., Sesa-Nogueras, E., Monte-Moreno, E., Esposito, A., Chetouani, M., Garre-Olmo, J., Abel, A., Smékal, Z., López-de-Ipiña, K.: Biometric applications related to human beings: There is life beyond security. Cognitive Computation 5(1), 136–151 (2013)
10. Brink, A., Smit, J., Bulacu, M., Schomaker, L.R.B.: Writer identification using directional ink-trace width measurements. Pattern Recognition 45, 162–171 (2012)
11. Marinai, S., Fujisawa, H. (eds.): Machine Learning in Document Analysis and Recognition. SCI, vol. 90. Springer, Heidelberg (2008)
12. Galibert, O., Kahn, J., Oparin, I.: The first MAURDOR campaign. In: NIST Open-Hart Workshop (2013)
13. Rosenblum, S., Parush, S., Weiss, P.: The in air phenomenon:temporal and spatial correlates of the handwriting process. Perceptual and Motor Skills, 933–954 (2003)
14. Kim, H.: The ClockMe system: computer-assisted screening tool for dementia. PhD thesis, Georgia Institute of Technology (2013)
15. Heinik, J., Werner, P., Dekel, T., Gurevitz, I., Rosenblum, S.: Computerized kinematic analysis of the clock drawing task in elderly people with mild major depressive disorder: an exploratory study. Int. Psychogeriatr, 479–488 (2010)
16. Plamondon, R., O'Reilly, C., Ouellet-Plamondon, C.: Strokes against stroke - strokes for strides. Pattern Recognition 47(3), 929–944 (2014)
17. O'Reilly, C., Plamondon, R.: Design of a neuromuscular disorders diagnostic system using human movement analysis. In: ISSPA, pp. 787–792 (2012)

18. Longstaff, M.G., Heath, R.A.: Spiral drawing performance as an indicator of fine motor function in people with multiple sclerosis. Human Movement Science 25, 474–491 (2006)
19. Caligiuri, M., Teulings, H., Filoteo, V., Song, D., Lohr, J.B.: Quantitative measurement of handwriting in the assessment of drug-induced parkinsonism (2006)
20. Sirat, C.: Ecriture et Civilisations. Editions du CNRS-IRHT (1976)
21. Marcelli, A., Parziale, A., Santoro, A.: Modeling handwriting style: A preliminary investigation. In: ICFHR, pp. 411–416 (2012)
22. Sriharia, S.N., Singerb, K.: Role of automation in the examination of handwritten items. Pattern Recognition 47, 1083–1095 (2014)
23. Bulacu, M., Schomaker, L.: Text-independent writer identification and verification using textural and allographic features. IEEE Trans. Pattern Anal. Mach. Intell. 29, 701–717 (2007)
24. Joutel, G., Eglin, V., Bres, S., Emptoz, H.: Curvelets based feature extraction of handwritten shapes for ancient manuscripts classification. In: DRR (2007)
25. Vincent, N., Bouletreau, V., Emptoz, H., Sabourin, R.: How to use fractal dimensions to qualify writings and writers. Fractals 8, 85–97 (2000)
26. Atanasiu, V., Likforman-Sulem, L., Vincent, N.: Writer retrieval - exploration of a novel biometric scenario using perceptual features derived from script orientation. In: ICDAR, pp. 628–632 (2011)
27. Eysenck, M.W., Derakshan, N., Santos, R., Calvo, M.G.: Anxiety and cognitive performance: attentional control theory. Emotion 7, 336 (2007)
28. Izard, C.: Basic emotions, relations among emotions, and emotions-cognition relations. Psychological Review 99, 561–565 (1992)
29. Crawford, J.R., Henry, J.D.: The depression anxiety stress scales (DASS): Normative data and latent structure in a large non-clinical sample. British Journal of Clinical Psychology 42, 111–131 (2003)
30. Breiman, L.: Random forests. Machine Learning 45, 5–32 (2001)
31. Team, R.C.: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2013)
32. Borghese, N.A., Lanzi, P.L., Mainetti, R., Pirovano, M., Surer, E.: Algorithms based on computational intelligence for autonomous physical rehabilitation at home. In: Bassis (ed.) Recent Advances of Neural Networks Models and Applications. SIST series. Springer (2014)

Part VII
Memristor and Complex Dynamics in Bio-inspired Networks

On the Use of Quantum-inspired Optimization Techniques for Training Spiking Neural Networks: A New Method Proposed

Maurizio Fiasché and Marco Taisch

Department of Management, Economics and Industrial Engineering,
Politecnico di Milano, Italy
`maurizio.fiasche@polimi.it`

Abstract. Spiking neural networks (SNN) are brain-like connectionist methods, where the output activation function is represented as a train of spikes and not as a potential. This and other reasons make SNN models biologically closer to brain principles than any of the alternative Artificial Neural networks (ANN) models proposed. In fact, they have great potential for solving complicated time-dependent pattern recognition problems defined by time series because of their inherent dynamical representation. A lot of works have been presented in the last decade about SNN which promote these models as third generation ANN. Nevertheless, several still open challenges have been reported in these studies. In this paper we analyze a particular type of SNN, the evolving SNN (eSNN), mainly focusing on their weights, parameters and features optimization using a new evolutionary strategy.

Keywords: Spiking Neural Network (SNN), evolving SNN (eSNN), Evolutionary Algorithms (EA), Quantum EA (QEA), Quantum Particle Swarm Optimization (QPSO).

1   Introduction

Spiking Neural Networks (SNN) represent a special class of artificial neural networks (ANN), where neurons communicate by train of spikes. Networks composed of spiking neurons are able to process the information using a relatively small number of spikes [1]. Being very similar to biological neurons on a functional point of view, they are used in powerful tools for the analysis of elementary processes in the brain, including neural information processing, plasticity and learning. At the same time spiking networks offer solutions to a broad range of specific problems in applied engineering, such as fast signal-processing, classification, time-series event prediction, pattern recognition, etc. It has been demonstrated that SNNs can be applied not only to all problems solvable by non-spiking neural networks, but that spiking models are computationally more powerful than perceptrons and sigmoidal gates [2]. Several improvements and variants of the spiking neuron model have been developed. An Evolving Spiking Neural Network (ESNN) is proposed by Wysoski et

al. [3] where the output neuron evolves based on the input patterns and weight similarity, and the neural model is trained by a fast one-pass learning algorithm [4]. Due to its evolving nature the model can be updated whenever new data becomes available, without requiring the retraining of earlier presented data samples. Some promising results could be obtained both on synthetic benchmarks and real world datasets. Like other neural network models, the correct combination of parameters would influence the performance of the network. On the other hand, increasing the number of used features does not necessarily translate into a higher classification accuracy. In some cases, having fewer significant features could help both to reduce the processing time and to produce good classifications. This study combines the ESNN architecture with a Quantum inspired evolutionary algorithm (QEA), so as to investigate the potential of ESNN when applied to Feature Subset Selection (FSS) problems following the wrapper approach.. The latter one is used to identify relevant feature subsets and simultaneously evolve an optimal parameter setting for the ESNN, while the ESNN itself operates as a quality measure for a given feature subset. By optimizing two search spaces in parallel, we expect to evolve an ESNN configuration specifically generated for the given dataset, providing in the meanwhile a specific feature subset maximizing classification accuracy. In particular, our aim is twofold: i) to describe the use of Evolutionary algorithms useful for training SNNs; and ii) to propose a clever algorithm based on Particle Swarm Optimization (PSO) joint with Quantum principles for simultaneous features and parameters optimization of an ESSN.

The paper is organized as follows. In Section 2 an overview of the main components of the proposed model is provided, starting from the adopted Spiking model and related optimization issues, and proposing next a brief review of PSO and its Quantum inspired generalization. In Section 3 the theoretical framework is described while experimental results are reported in Section 4. Finally, in Sections 5 and 6 results and conclusions are presented, giving some hints on possible future trends.

2   Spiking Neural Networks

SNNs represent information as trains of spikes, rather than as single scalars, thus allowing the use of features such as frequency, phase, incremental accumulation of input signals, time of activation, etc. Neuronal dynamics of a spiking neuron are based on the increase in the inner potential of a neuron (post synaptic potential, PSP), after every input spike arrival. When a PSP reaches a certain threshold, the neuron emits a spike at its output. In biological neural networks, neurons are connected at synapses and electrical signals (spikes) pass information from one neuron to another. SNN are biologically plausible and offer some means for representing time, frequency, phase and other features of the information being processed. A simplified diagram of a spiking neuron model is shown in Fig. 1(a). Fig. 1(b) shows the mode of operation of a spiking neuron, which emits an output spike when the total spiking input – Post Synaptic Potential ($u(t)$ in the figure), is larger than a spiking threshold.

The information in ESNN is represented as spikes; therefore, input information must be encoded in spike pulses. There are several information encoding methods in Spiking Neural Network (SNN). In this paper we have used a well-known encoding technique for ESNN: the Population Encoding [5]. Population Encoding distributes a single input value to multiple pre-synaptic neurons. Each pre-synaptic neuron generates a spike at firing time. The firing time is calculated using the intersection of Gaussian function. The centre of the Gaussian function is calculated using Equation (1) and the width is computed using Equation (2) with the variable interval of [I_{min}, I_{max}]. The parameter β controls the width of each Gaussian receptive field.

$$\mu = I_{min} + (2 * i\text{-}3) / 2 * (I_{max} - I_{min})/(M - 2) \tag{1}$$

$$\sigma = 1 / \beta \, (I_{max} - I_{min})/(M - 2) \text{ where } 1 \leq \beta \leq 2 \tag{2}$$
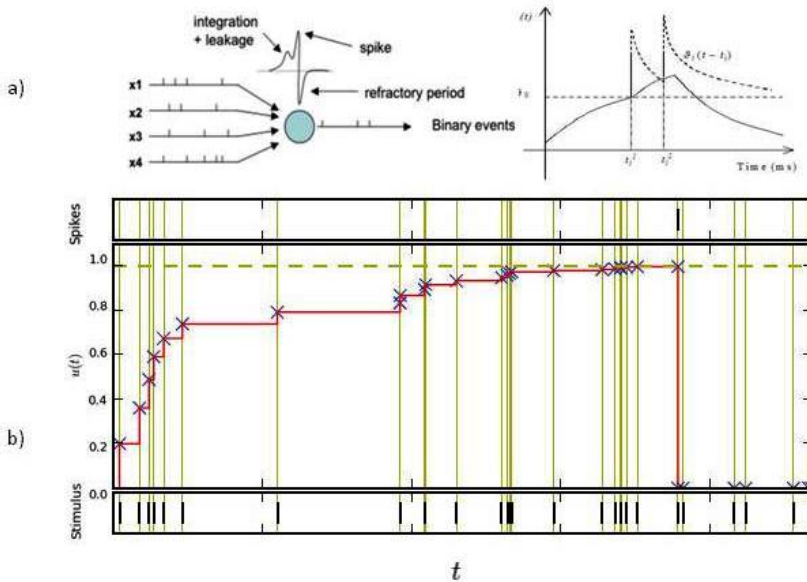
Fig. 1. (a) A short visualization of a simple spiking neuron model with its spike representation in time. (b) The representation of a spiking neuron emitting an output spike when the total spiking input – Post Synaptic Potential (*u(t)*representing the PSP), is larger than a spiking threshold [6].

2.1   Evolving Spiking Neural Networks (ESNNs)

ESNNs evolve/develop their structure and functionality in an incremental way from incoming data based on the following principles [6]: (i) new spiking neurons are created to accommodate new data, e.g. new patterns belonging to a class or new output classes, such as faces in a face recognition system; (ii) spiking neurons are merged if they represent the same concept (class) and have similar connection

weights (defined by a threshold of similarity). In [4,8] an ESNN architecture is proposed where the change in a synaptic weight is achieved through a simple spike time dependent plasticity (STDP) learning rule:

$$\Delta w_{j,i} = \text{mod}^{\,\text{order}(j)} \tag{1}$$

where: $w_{j,i}$ is the weight between neuron j and neuron i, mod $\in (0,1)$ is the modulation factor, order(j) is the order of arrival of a spike produced by neuron j to neuron i. For each training sample, the *winner-takes-all* approach has been used, where only the neuron that has the highest postysynaptic (PSP) value updates its weights. The postsynaptic threshold (*PSPTh*) of a neuron is calculated as a proportion $c \in [0, 1]$ of the maximum postsynaptic potential, max(*PSP*), generated by modulating the training sample with the updated weights, i.e.:

$$PSP_{Th} = c \max(PSP) \tag{2}$$

The One-Pass Algorithm is the learning algorithm for ESNN which follows both the SDTP learning rule and the time-to-first spike learning rule [7]. In this algorithm, each training sample creates a new output neuron. The trained threshold values and the weight pattern for that particular sample are stored in the neuron repository. However, if the weight pattern of the trained neuron greatly resembles a neuron in the repository, it will merge into the most similar one. The merging process involves modifying the weight pattern and the threshold of the merged neurons to the average value. Otherwise, it will be added to the repository as a newly trained neuron. The major advantage of this learning algorithm is the ability of the trained network to learn incrementally new samples without retraining.

Creating and merging neurons based on both localised incoming information and system performance are the main operations of the ESNN architecture that makes it continuously "evolvable".

2.2   Optimization Challenges

In order to provide an efficient and accurate solution to the simultaneous optimization task of features and parameters of an ESNN, for its interesting properties in terms of solution quality and convergence speed, a Versatile Quantum-inspired Evolutionary Algorithm (vQEA) [9] has been used in [10]. The method evolves in parallel a number of independent probability vectors, which interact at certain time intervals with each other, forming a multi-model Estimation of Distribution Algorithm (EDA) [11]. It has been shown that this approach performs well on epistatic problems, is very robust to noise, and needs only a minimal fine-tuning of its parameters. Moreover, the standard setting for vQEA is suitable for a large range of different problem sizes and classes, and in particular fits well to the feature selection problem under consideration. For the optimization of general numerical problems a QEA has been proposed in [12] as a clever modification of the technique introduced by Han and Kim [13]. In this paper we want to introduce another type of QEA for training ESNN, starting from the general optimization technique described in [12].

2.3   Quantum Inspired Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population based optimization technique developed by Eberhart and Kennedy in 1995 [14]. Individual particles work together to solve a given problem by responding to their own performance and to the performance of other particles in the swarm. Each particle computes its own fitness value during the optimization process and stores the best fitness value achieved so far, normally referred to as personal best or individual best (*pbest*). Likewise, the overall best fitness value obtained by any particle in the population is called global best (*gbest*). Each particle n moves to a new position $x_n(t)$ by computing, through the value of *pbest* and *gbest*, its velocity vector $v_n(t)$, according to the following formulas:

$$v_n(t) = w\,v_n(t{-}1) + c_1\,r_1\,(gbest_n - x_n(t)) + c_2\,r_2\,(pbest_n - x_n(t)) \tag{3}$$

$$x_n(t) = x_n(t{-}1) + v_n(t) \tag{4}$$

where: $c_1 > 0$ and $c_2 > 0$, called the cognitive and social parameters, control the particle acceleration towards the personal best or global best position, r_1 and r_2 are two uniform random realizations, and $w > 0$ is a constant called the inertia parameter. Moreover to create a swarm of N particles, at time t each particle i has:

1. a current position vector $x_i(t)$,
2. a current velocity vector $v_i(t)$,
3. a record of its own best positions $pbest = (pbest_1,...,pbest_t)$,
4. a record of the best global position $gbest = (pbest_1,...,pbest_t)$

As previously reported a QEA has been presented in 2002 [13] as inspired by the concept of quantum computing. According to the classical computing paradigm, information is represented in bits where each bit must hold either 0 or 1. However, in quantum computing, information is instead represented by a qubit in which a value of a single qubit could be 0, 1, or a superposition of both. Superposition allows the possible states to represent both 0 and 1 simultaneously based on their probability. The quantum state is modeled by the Hilbert function and is defined as $|\Psi> = \alpha\,|0>$ $+ \beta\,|1>$ where α and β are complex numbers defining probabilities of being in the corresponding state (when a qubit collapses, for instance, when reading or measuring). Probability fundamentals require

$$|\alpha|^2 + |\beta|^2 = 1$$

where $|\alpha|^2$ (resp. $|\beta|^2$) gives the probability that a qubit is in the OFF (0) (resp. ON (1))state. General notation for an individual with several qubits can be defined as

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \beta_2 & \beta_2 & \cdots & \beta_m \end{bmatrix}$$

The Quantum inspired PSO (QPSO) was first discussed in [15], and several variants of QPSO [16,17] have been developed thereafter. The main idea of QPSO is to use a standard PSO function to update the particle position represented in a qubit. In order

for PSO to update the probability of a qubit, the quantum angle θ is used, which can be represented as $\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$.

3   Theoretical Framework

This paper presents QPSO as a new optimizer for ESNN. From the well-known wrapper approach, QPSO interacts with an induction method (the ESNN), optimizes the ESNN parameters: modulation factor (mod), proportion factor (C), and neuron similarity (sim), and finally identifies relevant features.

   All particles are initialized with a random set of binary values and they subsequently interact with each other based on classification accuracy. Since there are two components to be optimized, each particle is divided into two parts. The first part, used for feature optimization, holds the feature mask values (value 1 represents a selected feature, 0 otherwise), while the second part holds binary strings for parameter optimization, where a set of qubits represent the parameters value. In fact, information held by each particle is in binary representation, therefore, conversion into real value is required. For this task, Gray code method is chosen since it is proven to be a simple and effective way in representing real value from binary representation. The proposed framework is depicted in Fig. 2.
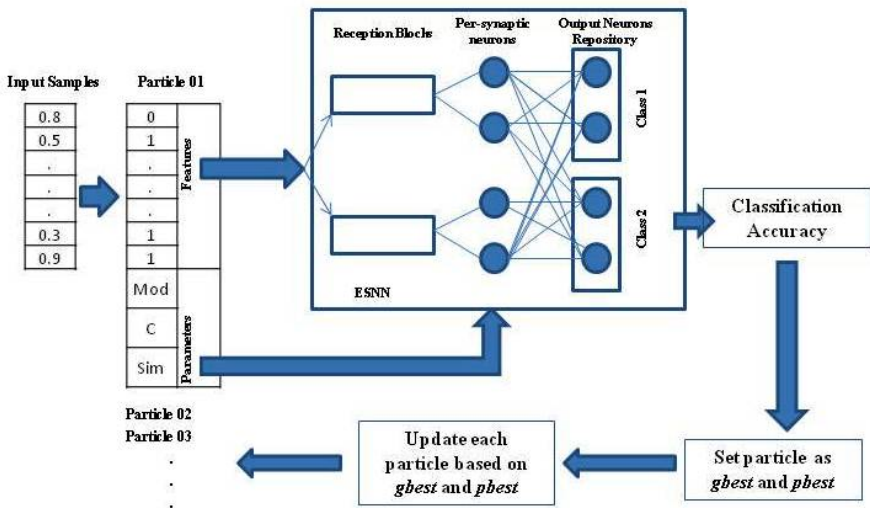
Fig. 2. A conceptual representation of the Framework QPSO-ESNN

4   Experimentations

The proposed QPSO-ESNN method is tested using the uniform hypercube dataset introduced in [18]. Only two of the ten features created, namely f_1 and f_2, are relevant to determine the output class. The problem consists of 600 samples grouped in two classes: 276 samples belong to class 1 and 324 to class 2; in particular, a sample belongs to class 1 when $f_i < \alpha * \gamma^{i-1}$ for i = 1 and 2, where γ = 0.8 and α = 0.5. The irrelevant features consist of four unit random values and four redundant copies of the two relevant features with the addition of a Gaussian noise having zero mean and standard deviation equal to 0.3. To compute performance measures, 10-folds data cross validation has been used.

From our preliminary investigation, we found that the location of the relevant features has a direct effect on the threshold value. In this experiment, features 1 to 4 are set as random, 5 to 8 as redundant, while 9 and 10 are defined as relevant. This is the reason to investigate the ability of QPSO in optimizing value of parameter C where it should be more than 0.5, since the relevant features are at the end of the ordered features. In reception Blocks described in Fig. 2, receptive fields were used to produce a weights pattern or weights vector of a particular sample for identifying the output class. During our experiments, different numbers of receptive fields for each dataset influenced the results accuracy. In our preliminary experimentation, 10 receptive fields were chosen, and 20 particles were used to explore the solution. The variables number to be optimized by QPSO are three ESNN parameters and 10 features. Since all three parameters ranged between 0 and 1, six qubits were sufficient to represent the real value. Cognitive and social parameters c_1 and c_2 were set to 0.05, leading to a balanced exploration between *gbest* and *pbest*, while the inertia weight w was set to equal to 2.0. We compared ESNN with feature optimization with and ESNN with all features; in both cases we performed parameter optimization for 15 continuous times iterations, computing the average results after 100 iterations. During our experimentation we also compared the ESNN considered with PSO (ESNN-PSO) as optimizer and the QPSO-ESNN proposed here. The ESNN-PSO algorithm, although with a high classification accuracy is highly dependent on the parameter optimization which has affected the results, giving the lowest accuracy compared with QPSO-ESNN. Accuracy obtained during the test phase for ESNN-PSO is in fact of 80% vs 90% of QSPO-ESNN.

5   Results

Results of our experimentations show that QPSO can optimize parameters and features in less than 80 iterations. The average accuracy for the ESNN with feature optimization is consistently above 90% compared to ESNN using all features whose mean accuracy attests toward 60%. In the latter case, the network was unable to react to the redundancy introduced by the irrelevant features. Moreover, it has been interesting to observe that the average accuracy of ESNN with feature optimization in the first iteration is extremely high, around 80%, thanks to the presence of a particle

able to reach an extremely good solution in the early stage of the algorithm. Due to the best global position *gbest*, the other particles then updated their position toward the best solution in a few iterations.

During the learning process, from the 10 features, the *gbest* particle is able to reduce number of features at the average of six features in the early iteration. Then, the algorithm keeps deleting irrelevant features in order to identify the most relevant ones. The two relevant features are always selected until the end of learning process in 15 runs. However, sometimes no significant features have been selected together with relevant features until the end of learning process and in this case, the proposed method was still able to identify the weights pattern to classify the output class correctly. In this experiment, QPSO managed to optimize binary string information, which represents ESNN parameter values as expected. Mod was used as a weight value with the objective to have different set of weights pattern to differentiate output class, therefore, normally the value must not be too low. If a lower value is selected, there would be only several connections with the weight value and this would make it difficult to have different set of weights pattern. In contrast, higher value means most of the weights will have the connection value, which can be translated into a well-presented weights pattern in accordance to their output class. Thus, in our preliminary investigation, we found that the Mod value should be between 0.6 and 1.0 and QPSO managed to come out with the Mod value within that range in this experiment. With the relevant features located at the last two features, result shows that the average C value found in this experiment is around 0.8, which is more acceptable. Finally, average Sim value found is 0.1 that is fairly adequate since the weights pattern is quite similar between input samples in the same class. Overall, all three parameters evolve steadily towards a certain optimal value, where the correct combination leads to the better classification accuracy.

6   Conclusions

This paper presents a new integration method for simultaneous optimization of features and parameters in an ESNN using QPSO starting from a more general QEA. The results have shown that QPSO is able to select the relevant features as well as to optimize ESNN parameters that generate higher classification accuracy for QPSO-ESNN with feature and parameter optimization compared to QPSO-ESNN using all features. Moreover also a comparison with an ESNN with a PSO as optimizer has been done highlighting that QPSO-ESNN performs better than the previous one both in terms of classification accuracy and for feature selected. Future work will focus on how to find a more effective method for eliminating the less relevant features, also proposing a general QEA for optimization [12] adapted to ESNN and compared it with the approach presented in this paper. The optimization also of other crucial aspects for a SNN, e.g. Connections and threshold to cause the spike of neurons will be analyzed. The proposed method will also be tested on different types of dataset such as string dataset as well as other real world dataset with comparison to the other classification algorithms like [19,20].

References

1. VanRullen, R., Guyonneau, R., Thorpe, S.: Spike times make sense. Trends Neurosci. 28, 1–4 (2005)
2. Maass, W.: Networks of spiking neurons: The third generation of neural network models. Neural Networks 10, 1659–1671 (1997)
3. Wysoski, S.G., Benuskova, L., Kasabov, N.: On-line learning with structural adaptation in a network of spiking neurons for visual pattern recognition. ICANN (1), 61–70 (2006)
4. Wysoski, S.G., Benuskova, L., Kasabov, N.: Brain-like evolving spiking neural networks for multimodal information processing. In: Hanazawa, A., Miki, T., Horio, K. (eds.) Brain-Inspired Information Technology. SCI, vol. 266, pp. 15–27. Springer, Heidelberg (2010)
5. Bohte, S.M., Kok, J.N., Poutre, H.L.: Error-Backpropagation in Temporally Encoded Networks of Spiking Neurons. Neurocomputing 48(1-4), 17–37 (2002)
6. Kasabov, N.: Evolving Connectionist Systems: The System Engineering Approach, vol. 2. Springer-Verlag New York Inc., Secaucus (2007)
7. Thorpe, S.J.: How Can the Human Visual System Process a Natural Scene in Under 150ms? Experiments and Neural Network Models. In: Verleysen, M. (ed.) Proceedings of European Symposium on Artificial Neural Networks, D-Facto public, ISBN 2-9600049-7-3, Bruges, Belgium (1997)
8. Soltic, S., Wysoski, S., Kasabov, N.: Evolving spiking neural networks for taste recognition. In: IEEE World Congress on Computational Intelligence (WCCI), Hong Kong (2008)
9. Defoin-Platel, M., Schliebs, S., Kasabov, N.: A versatile quantum-inspired evolutionary algorithm. In: IEEE Congress on Evolutionary Computation, CEC 2007, pp. 423–430 (2007)
10. Schliebs, S., Defoin-Platel, M., Worner, S., Kasabov, N.: Integrated Feature and Parameter Optimization for an Evolving Spiking Neural Network: Exploring Heterogeneous Probabilistic Models. Neural Networks 22, 623–632 (2009)
11. Platel, M.D., Schliebs, S., Kasabov, N.: Quantum-Inspired Evolutionary Algorithm: A Multimodel EDA. IEEE Transactions on Evolutionary Computation 13(6), 1218–1232 (2009)
12. Fiasché, M.: A Quantum-Inspired Evolutionary Algorithm for Optimization Numerical Problems. In: Huang, T., Zeng, Z., Li, C., Leung, C.S. (eds.) ICONIP 2012, Part III. LNCS, vol. 7665, pp. 686–693. Springer, Heidelberg (2012)
13. Han, K.H., Kim, J.H.: Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. IEEE Transactions on Evolutionary Computation, 580–593 (2002)
14. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: Proc. Sixth International Symposium on Micro Machine and Human Science, pp. 39–43. IEEE Press (1995)
15. Sun, J., Feng, B., Xu, W.B.: Particle swarm optimization with particles having quantum behavior. In: Proc. Congress on Evolutionary Computation, vol. 1, pp. 325–331 (2004)
16. Hamed, H.N.A., Kasabov, N., Shamsuddin, S.M.: Integrated feature selection and parameter optimization for evolving spiking neural networks using quantum inspired particle swarm optimization. In: Soft Computing and Pattern Recognition, SoCPaR 2009, pp. 695–698 (2009)

17. Hamed, H.N.A., Kasabov, N., Shamsuddin, S.M.: Quantum-Inspired Particle Swarm Optimization for Feature Selection and Parameter Optimization in Evolving Spiking Neural Networks for Classification Tasks. In: Kita, E. (ed.) Evolutionary Algorithms, InTech (2011)
18. Estavest, P., Tesmer, M., Perez, C., Zurada, J.: Normalized mutual information feature selection. Neural Networks 20(2), 189–201 (2009)
19. Kasabov, N.: Integrative probabilistic evolving spiking neural networks utilising quantum inspired evolutionary algorithm: A computational framework. In: Köppen, M., Kasabov, N., Coghill, G. (eds.) ICONIP 2008, Part I. LNCS, vol. 5506, pp. 3–13. Springer, Heidelberg (2009)
20. Kasabov, N.K.: NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. Neural Networks 52, 62–76 (2014)

Binary Synapse Circuitry for High Efficiency Learning Algorithm Using Generalized Boundary Condition Memristor Models

Jacopo Secco[1], Alessandro Vinassa[1], Valentina Pontrandolfo[1],
Carlo Baldassi[2], and Fernando Corinto[1]

[1] Politecnico di Torino, Department of Electronic and Telecommunication (DET),
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[2] Politecnico di Torino, Department of Applied Sciences and Technologies (DISAT),
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
{fernando.corinto}@polito.it

Abstract. Memristors are memory resistors that promise the efficient implementation of synaptic weights in artificial neural networks [1]. This kind of technology has permitted the implementation of a large number of real world data in an evolutionary learning artificial system. Human brain is capable of processing such data with standard always equal signals that are the synapsis. Our goal is to present a circuit which responds with binary outputs to the signal exiting from the memristors implemented in an artificial neural system that functions through a high efficiency learning algorithm.

Keywords: Memristor, Memristor–based Circuits, Binary Synapses, Neural Networks, Pattern Recognition.

1 Introduction and Background

1.1 Memristor Model

In 1971 Prof. Leon Chua introduced a theoretical model of the memristor [2]. The memristor is a two-terminal non–linear element capable of changing and maintaining its resistance (memristance), defined as $R(q) = \frac{d\varphi}{dq}$, is only function of q. The equation $\varphi(q)$ is the flux in the device and it is considered the constitutive equation of the memristor. Several model have been proposed to describe the nonlinear behavior of memristor devices. In this manuscript we exploit the Generalized Boundary Bondition Memristor model (GBCM), developed by Ascoli et.al, which preserves the features of the Boundary Condition Mermristor model (BCM) but allows the tuning of activation based dynamics and if ibidem data processing and storage capability [3,4,5]. Let D denote the length of the nano–film and $x = (w/D) \in [0,1]$ represent the longitudinal extension of the conductive part of the nano–film. The model equations are [3]

$$\frac{d\,x(t)}{dt} = ki(t)f_p(x(t), v(t)) \tag{1}$$

$$i(t) = W(x(t))v(t) \tag{2}$$

where $k \in \mathbf{R}$ is a constant depending on physical properties of the memristor (its dimension is C^{-1}, thus it is also referred to as memristor charge normalization factor), and $W(x(t))$ describes the state-dependent memory-conductance. On the other hand $f_p(x(t), v(t)) \in [0, 1]$ is a parametrizable *window–function* that describes the activation dynamics of the device under suitable boundary conditions. To perform learning essays the GBCM model was built on PSpice as shown in the referring article (see [3] for more details).

1.2   Learning Algorithm

Baldassi et al. in 2007 developed a learning algorithm called *Stochastic Belief–Propagation-Inspired* (SBPI) [6]. This kind of algorithm was studied to simulate biological learning dynamics with binary synapses. Here, we preliminary describe a simplified version of that algorithm (called CP in [6]), and we present a circuit which implements it. The CP algorithm has a reduced performance with respect to the SBPI algorithm, but our scheme can be extended rather straightforwardly to SBPI, or a variant thereof (such as the one proposed in [7]), since it is already able to model the most crucial quantities used in the algorithm; the extension to complete SBPI algorithm will be the subject of a future work, currently in preparation. A set ξ of binary patterns is presented to a network of N. The current outflowing from each memristor is function of the synaptic weight

$$w = \frac{1}{2}\left(sign(h) + 1\right)$$

where $h \in [-1, 1]$ is the normalized memristance value ($h = 0$ when $R_{mem} = (R_{on} + R_{off})/2$). At the same time the networks response (binary) is compared with the desired binary output σ. Given a set of $p = \alpha N$ (where $\alpha \in [0, 1]$) patternsxi and computed the stability parameter $\Delta = (2\sigma^\tau - 1)(I - \theta)$ where $I = \sum_i w_i^\tau \xi_i^\tau$ and θ is the threshold to obtain the desired binary response σ, at the time τ then (see [6] for a detailed presentatin of the SBPI algorithm)

(R1) If $\Delta \geq 0$, then all $w_i^\tau = w_i^{(\tau+1)}$
(R2) If $\Delta < 0$, then all $h_i^{(\tau+1)} = h_i^\tau + 2\xi_i^\tau(2\sigma^{(\tau+1)})$

with i ranging from 1 to N. All h_i can assume a number K of values (intermediate hidden states) supra- and under-threshold. A greater number of K states can ensure a better performance of the algorithm enabling the enlargement of the set of p patterns enhancing the uniqueness of the solution.

2   Methods

2.1   Electronic Circuit

In order to implement a circuit able to perform the tasks required for the CP algorithm, it was necessary to generate a binary response to the ξ patterns presented to the system. Figure 1 shows the Binary Synapse Circuitry (BSC) model that was designed to enable a binary response for each memristor implemented in the system [8]. In order to obtain a binary response by means of the memristance a CMOS system was designed as a logical NOT port.
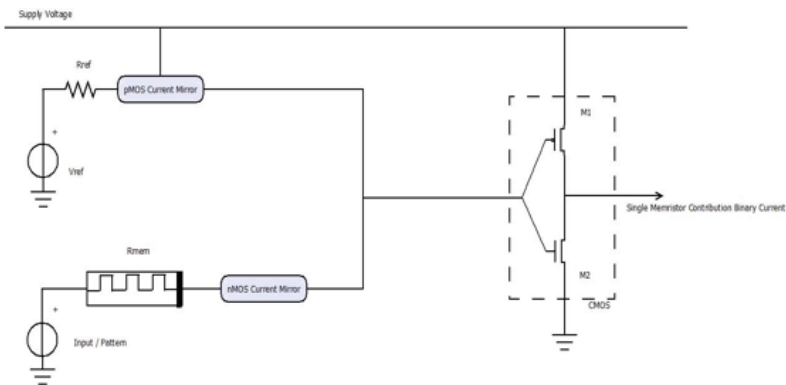
Fig. 1. Binary Synapse Circuitry model
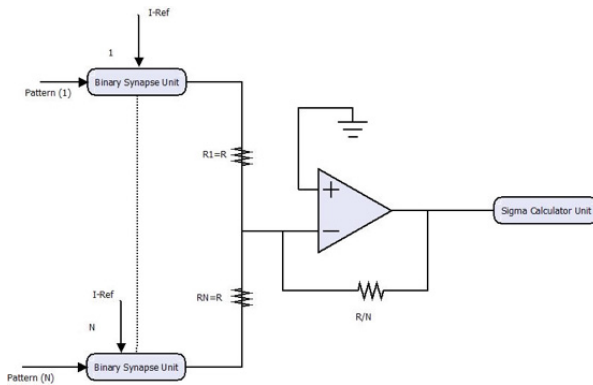
Fig. 2. Memristor Contribution Sum Unit and the Sigma Calculator Unit circuitry scheme
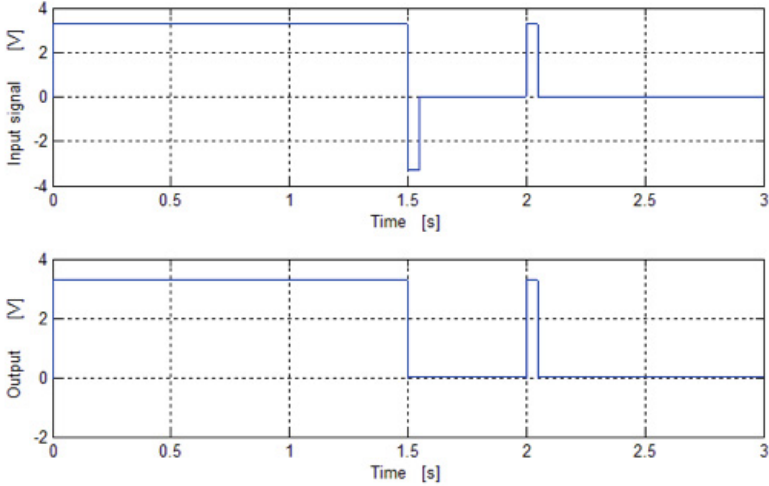
Fig. 3. Input and Output signal of the single BSC system

In order to perform the CP algorithm all the outputs of each binary memristor unit must be somehow summed and compared with the threshold θ in order to compare the outputs of the system with the desired response σ. Figure 2 shows the Memristor Contribution Sum Unit (MCSU) which was implemented with an op–amp. This solution permits to sum the binary voltages instead of the currents which results to be more efficient on the circuitry design level. MCSU was combined with the Sigma Calculator Unit (SCU) designed with the use of a double current mirror comparator. SCU has the scope of comparing the current outflowing from the MCSU with the θ threshold given by a voltage source connected to the nMOS mirror composing the SCU. Both MCSU and SCU are to be connected to a control unit able to calculate and switch from the "reading mode" to the "writing mode" of the memristors.

3   Results

The GBCM model, the BSC, the MCSU and the SCU were built on PSpice, while a model of the control unit performing the required actions of the CP algorithm was built with the use of Simulink. The two models were then coupled with the use of the Cadence tool SLPS. Several essays were performed to evaluate the activation and deactivation (R_{off} to R_{on} and R_{on} to R_{off} respectively) dynamics of the memristors, their required binary responses and the learning efficiency of the algorithm varying the θ threshold. As proof of concepts it was fixed a minimal Iref in order to better evaluate the binary response of the only BSC system. By applying a $3.3\,V$ and a $-3.3\,V$ potential to the memristor, as shown in Figure 3 it was possible to obtain binary responses.
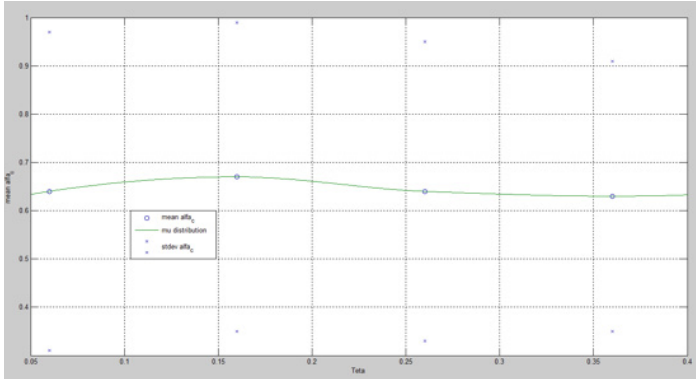
Fig. 4. Learning efficiency of the system (mean α_c) vs. θ

Since the supply voltage of the whole system was set to $3.3\,V$, for simplicity, also the ξ patterns and the writing inputs were set at the same amplitude. From the simulations that were ran on the system it was shown that the memristances did not vary with spikes that presented a period shorter than $1\mu s$. By setting this period to the ξ pattern spikes the system was able to perform the BSI algorithm without changing the synaptic weights at every pattern presentation. Figure 4 shows the learning efficiency of the system performing the BSI algorithm for θ values ranging from $0.6\,N$ to $0.36\,N$.

It is shown that for $\theta = 0.16$ there is a peak of efficiency of the system (mean $\alpha_c = 68\%$ of the total ξ patterns acquired). The final proof of the efficiency of this system composed by memristors was given by the analysis of the same system implemented with only CMOS technology. It was proven that a memristor which dynamics is divided in K states can be replaced with a number of transistors proportional to $N_t = N \log_2 K$. So for $K > 3$ the use of the memristor technology is preferable also in a design point of view.

4   Conclusion

Neural Networks trained with the BPI algorithm are able to learn a number of associations close to the theoretical limit in time that is sublinear in the number of input. Using binary synapses, implemented by a memristor, a single layer perceptron with BPI has been proposed and investigated.

References

1. Alibart, F., Zmanidoost, E., Strukov, D.B.: Pattern classification by memristive crossbar circuits using ex situ and in situ training. Nature Communications (2013)
2. Chua, L.O.: Memristor: the missing circuit element. IEEE Transactions on Circuit Theory 18(5), 507–519 (1971)
3. Corinto, F., Ascoli, A.: A boundary condition-based approach to the modeling of memristor nanostructures. IEEE Trans. Circuits Syst. I 59, 2713–2726 (2012), doi:10.1109/TCSI.2012.2190563
4. Corinto, F., Ascoli, A.: Memristive diode bridge with LCR filter. Electronics Letters 48(14), 824–825 (2012)
5. Batas, D., Fielder, H.: A memristor PSpice implementation and a new approach for magnetic flux-controlled memristor modeling. IEEE Transactions on Nanotechnology (2011)
6. Baldassi, C., Braunstein, A., Brunel, N., Zecchina, R.: Efficient supervised leaning in networks with binary synapses. PNAS (2007)
7. Baldassi, C.: Generalization Learning in a Perceptron with Binary Synapses. Journal of Statistical Physics (2009)
8. Manem, H., Rajedran, J., Rose, G.S.: Stochastic gradient descent inspired training technique for a CMOS/Nano memristive trainable threshold gate array. IEEE Trans. Circuits Syst. I (2012)

Analogic Realization of a Non-linear Network with Re-configurable Structure as Paradigm for Real Time Analysis of Complex Dynamics

Carlo Petrarca, Soudeh Yaghouti, Lorenza Corti, and Massimiliano de Magistris

Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione
University of Naples FEDERICO II - Via Claudio 21, I-80125 Napoli, Italy
```
{carlo.petrarca,soudeh.yaghouti,
lorenza.corti,m.demagistris}@unina.it
```

Abstract. A novel experimental set-up realized for the real time analysis of reconfigurable complex networks, with chaotic Chua's circuits as nodes, is considered. It has been designed to easily perform large scale experiments on networks of chaotic oscillators, possibly exploring in real time the parameters space in terms of topologies, coupling strengths, nodes' dynamics, with potential application in the area of neuro-computing and associative dynamic memories. A sample of the capabilities is given by considering diffusive coupling with a large range of coupling strengths in a set of topologies, and first experiments on a ring of 32 chaotic nodes are reported. Synchronization, chaotic waves and patterns are experimentally observed, and the potential of the realized set-up in terms of accuracy, flexibility and analysis time is fully revealed.

Keywords: Complex networks, nonlinear oscillatory networks, chaotic synchronization, chaotic waves.

1   Introduction

The exploration of complex networks, defined as eventually large ensemble of simple dynamical units with arbitrary, and possibly evolving topological structure, is strongly motivated as paradigmatic of very different phenomena, from neurobiology to social science [1,2]. Collective behavior (also referred as "emerging dynamics") are much richer than individual (stand alone) dynamics, with complexity arising from the combination of non linearity (local activity, [3]) and topological structure. Such studies embrace a vast range of applications; it is well known that Cellular Neural/nonlinear Networks (CNN) represent revolutionary paradigms for the information processing [4-5]. In particular modeling neurons as non linear oscillators has lead to bio-inspired paradigms for neurocomputers (see [6] and references therein).

In this context electronic systems have been clearly recognized as good candidates for modeling different real systems [7], mainly because of the availability of well developed simulation tools and, in principle, the possibility of realizing prototypes as integrated structures. A vast range of theoretical and numerical results is available in

literature, in particular for networks of interconnected non linear oscillators [7-10]. More rare is to find extensive experimental studies, especially when general purpose and configurable networks are considered.

We describe here the main features of a dedicated setup designed and realized as an analog simulator of configurable networks with high complexity nodes (Chua's circuits), and complete control on topology, type and strength of the interconnections. It allows real time experiments on complex dynamics with the double advantage of being more realistic than simulation and, at the same time, drastically reducing the time for getting results. In this work the structure and the potential of the realized setup is presented, and a sample of the experimentally observed dynamics is reported.

2   The Experimental Setup

The structure and realization of the experimental setup has already been described in quite detail in previous papers [11,12], and will not be repeated here, where only its major features will be briefly summarized. The network is based on a modular set of Chua's circuits, taken as paradigmatic case of complex and chaotic dynamic nodes. Stand-alone dynamics of individual nodes can be (individually) settled onto periodic or chaotic trajectories, by properly setting the values for the Chua's linear resistor R, allowing different regimes. The network nodes are interconnected via a fully configurable link network, with adjustable topology and settable coupling strength. Figure 1 depicts a schematic draw of the Chua's nodes and the interconnecting links, along with symbols and parameters used within the text. Figure 2 shows some of the topologies considered up to now.
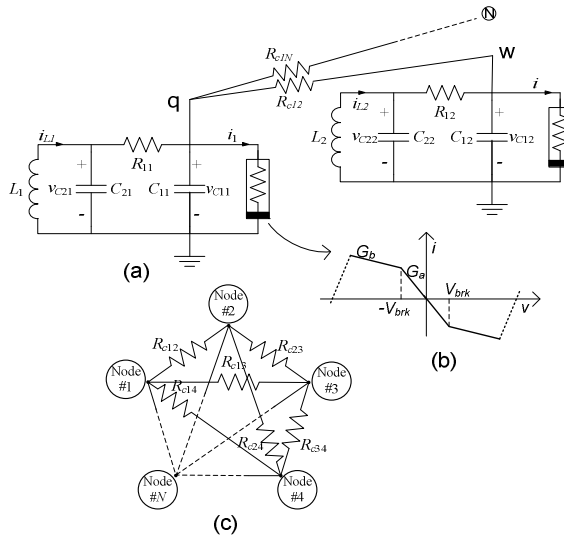
Fig. 1. a) Chua's circuit schematic, reference parameters, b) Chua's diode characteristic, c) general schematic of the link network

A modular National Instruments USB multi-channel data acquisition system is used to measure and monitor the variables of interest (nodes' states) in real time. Up to 64 state variables can be synchronously acquired by a single unit, at 60 k-sample/s rate for each channel. The whole network is controlled via a USB interface from a PC running LabVIEW. Two executive modes are possible: a "control mode", which allows to set all the network parameters, displaying in real time the waveforms of the signals; a "scan mode", which performs a scan of some preset values for network parameters (at present the topology is fixed at the beginning of each scan). Scan mode automatically sets, at each step, the defined parameter (link resistance value) in the predefined sequence, allows the system to reach regime, then acquires and saves all system variables in records up to 2500 samples per channel. A typical execution time for an accurate scan of 255 steps in the 32 nodes system is of about 45 minutes, mainly limited by the setting time for the link network values. Aside from experimental validation of theoretical conjectures and results, this allows very detailed analysis in a time by far shorter than any simulation system.
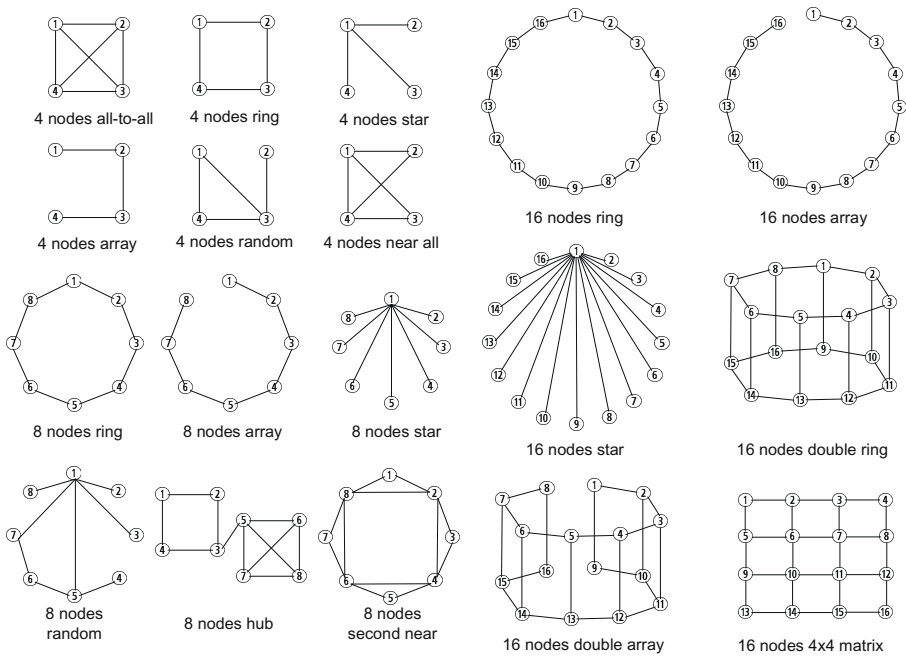
Fig. 2. A set of experimentally explored topologies

3   Experimental Results

The realized set-up allows in principle a wide exploration of link topologies and parameters space, with real time results. In particular the nodes oscillators can be set to different periodicity regimes as well as in chaotic behavior, and the link network can evolve in structure and link strength. Here we report briefly some results on chaotic

synchronization [13], spatio-temporal patterns and waves for the case of nominally identical nodes, settled in double scroll chaotic regime when uncoupled, and bilateral diffusive coupling, as a function of coupling strength (equally weighted on the links). Other remarkable results, obtained with different versions of the same assembly, are given in [14] (with reference to directed links and Pinning control) and [15] (with reference to links with dynamic elements).

Complete synchronization has been considered for all the network configurations as shown in figure 2 in the double scroll chaotic regime, and a proper defined RMS distance [11] in waveforms is calculated from measured data. Results are shown in fig. 3, where such RMS distance (in %) between corresponding variables are plotted as function of the resistance link values. It can be easily distinguished a strong threshold mechanism for the loss of synchronization, with an extremely good agreement to Master Stability Function (MSF) theoretical thresholds calculated as described in [16].
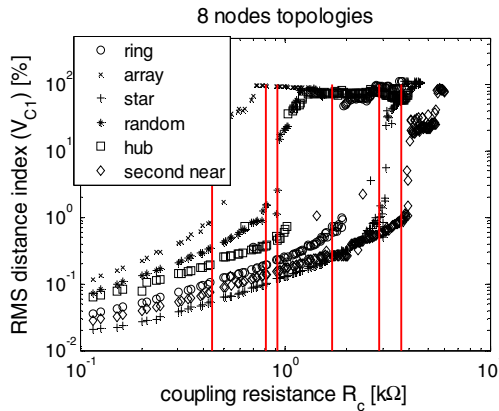
Fig. 3. Measured RMS distance for the 8 nodes topologies considered in figure2, as function of link resistance R_c; solid vertical lines correspond to theoretical thresholds calculated by MSF

A $N=32$ nodes ring structure has been realized with the same assembly, and we report here, for the first time, some results on patterns and waves. In figure 4 a typical output of the experiment is shown, for low value of link resistance ($R_c=198 \, \Omega$). Figure 4a illustrates the space state v_{C1k}-v_{C2k} dynamics for each system node, fig. 4b the synchronization diagrams of v_{C1k}, for $k=1..N$, fig. 4c the v_{C1k} waveforms for the node's system, fig. 4d an averaged RMS distance between node's state ($I_s(i,j)$). Dynamics of the nodes result in single scroll chaotic regime, and a space periodicity of N is recognized from the synchronization plots and the $I_s(i,j)$ distance matrix entries. A time-chaotic travelling (or "rotating") wave is clearly distinguishable from the waveform graphs. Exact space periodicity can be easily noted from the synchronization diagram $N=1$ vs. $N=32$ (bottom right in figure 4b)

A complete scan in the coupling resistance parameter range has been carried out, with dynamical behaviors in very good agreement to theoretical and numerical predictions given in [17].

In fig. 5 the dynamics corresponding to R_c=234 Ω link resistance is shown, in terms of v_{C1k} waveforms and the Poincaré diagrams referring to the v_{C2k}=0 condition.

(a)

(b)

(c)

(d)
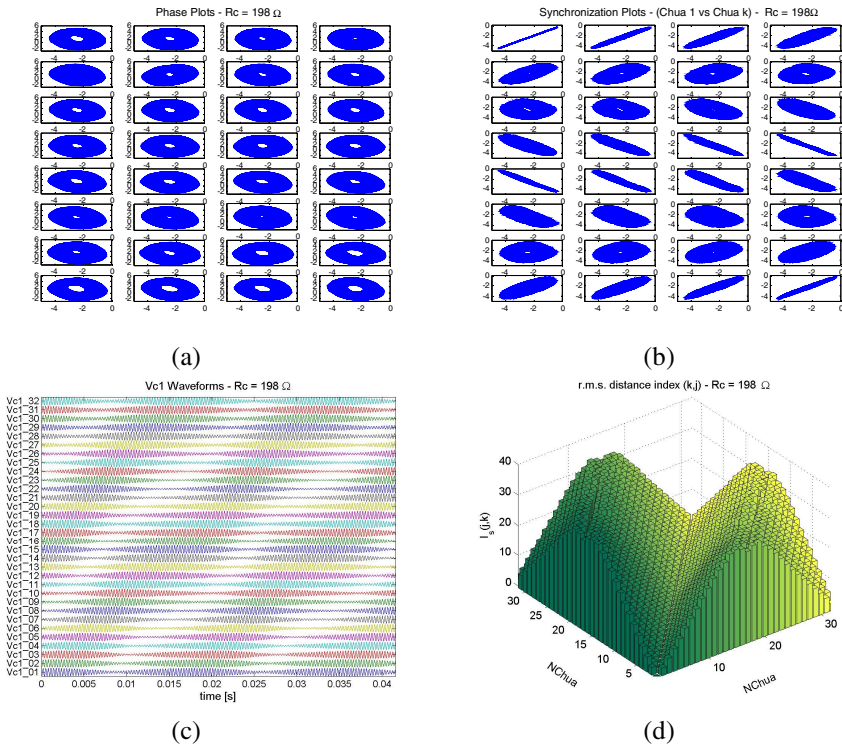
Fig. 4. results for the case N=32, ring, R_c =198 Ω: (a) phase plots for the nodes (v_{C1}- v_{C2}); (b) synchronization plots (v_{C11}- v_{C1k}); (c) waveforms (v_{C1k}); (d) distance index matrix (I_s(i,j))

(a)

(b)
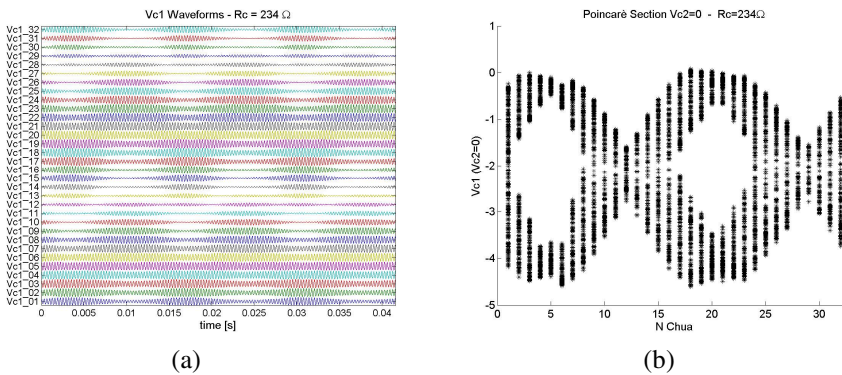
Fig. 5. results for the case N=32, ring, R_c =234 Ω: (a) waveforms (v_{C1k}); (b) Poincaré sections

Clock-wise and counter clock-wise rotating waves are distinguishable. As the link resistance increases, envelope oscillation frequency increase and loss of periodicity are observed (as shown in fig. 6). For higher values of R_C, $N/2$ wavelength chaotic waves appear (fig. 7).
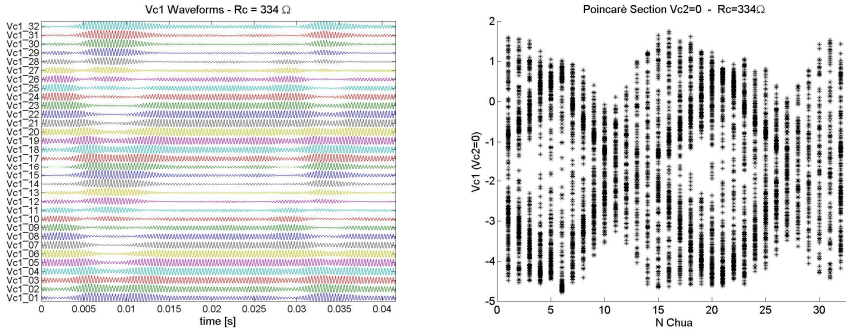
Fig. 6. Loss of periodicity (R_c=334 Ω).

(a)

(b)

(c)

(b)
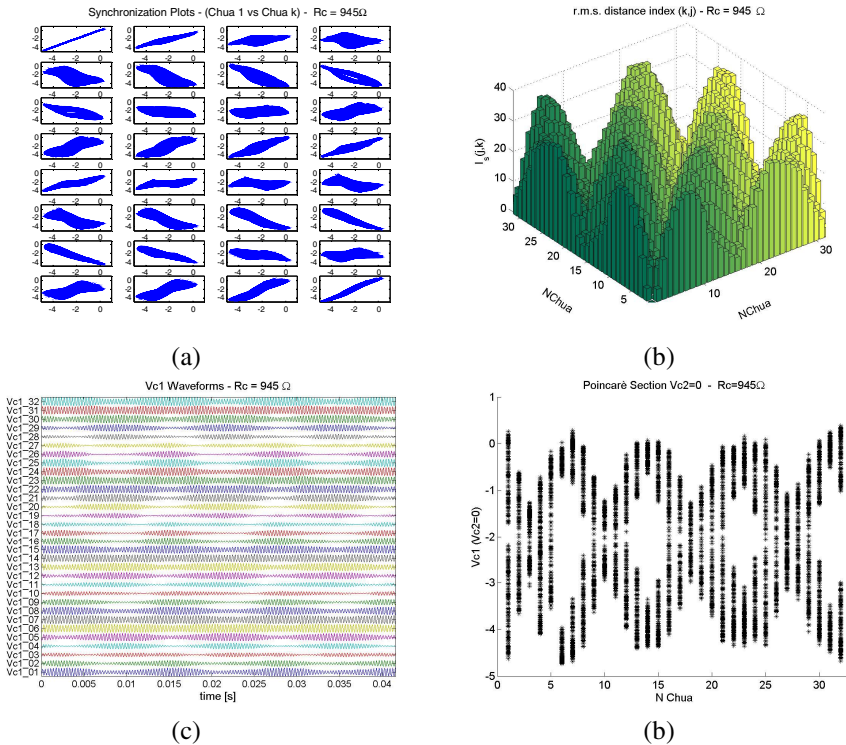
Fig. 7. Period $N/2$ rotating waves, (R_c=945 Ω); (a) synchronization plots; (b) relative distance index; (c) waveforms (v_{C1k}); (d) Poincaré sections;

4   Conclusions

The structure and the experimental potential of the analogical realization of a non linear network with re-configurable structure has been illustrated. It represents a flexible and accurate analogical simulator for real time analysis of dynamics in complex networks. It has been demonstrated as an effective validation tool of theoretical results, such as MSF, and a real system example for the observation of patterns and non linear waves. Moreover such analogical analysis, apart from being a realistic model of continuous systems, compete with simulation both in accuracy and granularity, with results by far better in terms of experiment time when the number of nodes grows to moderately large numbers. The great experimental potential of the considered set-up is at the moment still largely unused and will be explored in future works.

References

1. Strogatz, S.H.: Exploring complex networks. Nature 410, 268–276 (2001)
2. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex networks: Structure and dynamics. Physics Reports 424(4-5), 175–308 (2006)
3. Chua, L.O.: Local activity is the origin of complexity. International Journal of Bifurcation and Chaos 15(11), 3435–3456 (2005)
4. Chua, L.O., Roska, T.: Cellular Neural Networks and Visual Computing: Foundations and Applications. Cambridge University Press (2002)
5. Yalcin, M.E., Suykens, J.A.K., Vandewalle, J.P.L.: Cellular Neural Networks, Multi-scroll chaos and synchronization. World Scientific series on Nonlinear Science, series A, vol. 50 (2005)
6. Corinto, F., Bonnin, M., Gilli, M.: Weakly connected oscillatory network models for associative and dynamic memories. International Journal of Bifurcation and Chaos 17(12), 4365–4379 (2007)
7. Ogorzalek, M.J., Galias, Z., Dabrowski, A.M., Dabrowski, W.R.: Chaotic waves and spatio-temporal patterns in large arrays of doubly-coupled Chua's circuits. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 42(10), 706–714 (1995)
8. Dabrowski, A.M., Dabrowski, W.R., Ogorzalek, M.J.: Dynamic Phenomena in Chain Interconnections of Chua's Circuits. IEEE Transactions on Circuits and Systems 40(11), 868–871 (1993)
9. Nishio, Y., Ushida, A.: Spatio-temporal chaos in simple coupled chaotic circuits. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 42(10), 678–686 (1995)
10. Osipov, V.G., Shalfeev, V.D.: Chaos and structures in a chain of mutually-coupled Chua's circuits. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications 42(10), 693 (1995)
11. de Magistris, M., di Bernardo, M., Manfredi, S., Di Tucci, E.: Synchronization of Networks of Non-Identical Chua's Circuits: Analysis and Experiments. IEEE Transactions on Circuits and Systems I: Regular Papers 59(5), 1029–1041 (2012)

12. Colandrea, M., de Magistris, M., di Bernardo, M., Manfredi, S.: A Fully Reconfigurable Experimental Setup to Study Complex Networks of Chua's Circuits. In: Nonlinear Dynamics of Electronic Systems, Proceedings of NDES 2012, July 11-13, pp. 62–65 (2012)
13. Arenas, A., Diazguilera, A., Kurths, J., Moreno, Y., Zhou, C.: Synchronization in complex networks. Physics Reports 469(3), 93153 (2008)
14. DeLellis, P., de Magistris, M., di Bernardo, M., Manfredi, S.: Experimental Validation of Pinning Controllability in Networked Chua's Circuits. In: 2012 IEEE International Symposium on Circuits and Systems (ISCAS), May 20-23, pp. 616–619 (2012)
15. de Magistris, M., di Bernardo, M., Petrarca, C.: Experiments on synchronization in networks of nonlinear oscillators with dynamic coupling. Nonlinear Theory and Its Applications, IEICE 4(4), 462–472 (2013)
16. Pecora, L.M., Carroll, T.L.: Master stability functions for synchronized coupled systems. Physical Review Letters 80, 2109–2112 (1998)
17. Shabunin, A., Astakhov, V., Anishchenko, V.: Developing chaos on base of traveling waves in chain of coupled oscillators with period doubling synchronization and hierarchy of multi-stability formation. International Journal of Bifurcations and Chaos 12(8) (2002)

A Memristive System Based
on an Electrostatic Loudspeaker

Amedeo Troiano, Eugenio Balzanelli, Eros Pasero, and Luca Mesin

Politecnico di Torino, Department of Electronic and Telecommunication (DET),
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
`{amedeo.troiano,eugenio.balzanelli,eros.pasero,luca.mesin}@polito.it`

Abstract. The memristor (a memory–resistor) is a fundamental two-terminal circuit element with a nonlinear relationship between the integral of the voltage and the charge. In the literature, the research interest in the development of new memristive systems is growing, due to the potential applications as analog memories or as synapses in neuromorphic systems. In this paper, the possibility of using an electrostatic loudspeaker as a memristor-based system is explored. This kind of loudspeakers use a thin flat polarized diaphragm, usually consisting of a plastic sheet coated with a conductive material, between two electrically conductive plates, with a small air gap between them. When an electrostatic field is applied to the plates, a force is exerted on the charged diaphragm, and its resulting movement drives the air on either side of it. To get a memristor, the deformation of the diaphragm is here converted in a resistance value using a strain gauge attached over it. A mathematical model of the system is developed. Simulation results show that the device based on the combination of an electrostatic loudspeaker and a strain gauge has all the properties of the memristive systems.

Keywords: Memristor, Memristive system, Electrostatic Loudspeaker, Nonlinear circuit, Memory devices, Strain Gauge.

1 Introduction

The existence of a device that works as a resistor with memory capability was predicted by Prof L. O. Chua in 1971 [1]. Such a bipole was named memristor (memory-resistor). The ideal memristor is just one element from a large class of nonlinear dynamical systems, named memristive (memristor-based) systems [2]. Only more than 30 years later than the first theoretical studies, a nano-device with memristive properties based on titanium dioxide films was built at the HP labs [3].

Recently a great interest in this device is growing. Different simulators have been developed to study the titanium dioxide memristor developed at the HP labs. For example, a mathematical model is discussed in [4], where the dynamical behavior of such a system is shown to be highly nonlinear and asymmetric. In [5] and [6], SPICE models of the same memristor are also introduced. Another work

based on simulations proposed a memristive system with a current threshold and a nonlinear dependence on the charge [7]. In a recent paper [8], the dynamical behaviors of different memristor models are compared.

Other devices showing memristive properties were developed based on materials which are alternative to the titanium dioxide: polymeric materials [9], ferroelectric memristors [10], spintronic devices [11], spin-transfer torque magnetoresistances [12], spin memristive systems [13], superconductive materials [14], Mn doped ZnO films [15].

Finally, simple memristive systems including only passive components were developed, mainly for didactic purposes. For example, the electronic circuit proposed in [16] had the most important features of a memristor-based system [17], despite of its simplicity, as it included only a Graetz bridge with a RLC series filter. However, it had no memory capability, which was then added using a transfer charge circuit [18].

Memristors are founding fascinating potential applications. For example, a memristor based approach to artificial intelligence was introduced in [19], where a software that mimics how neurons process information was build based on memristive systems. This software can be used to describe neuromorphic systems. Moreover, memristors can be included as a readily used building block for bioelectrical data analysis and modeling [20].

In this paper, a new memristive system is proposed, based on an electrostatic loudspeaker and a strain gauge. The electrostatic loudspeaker has a membrane which is deformed by an external voltage; such a deformation induces a variation of the resistance of the strain gauge attached to the diaphragm. The electrostatic loudspeaker can maintain the deformation of its membrane due to the charge stored on its plates, showing memory properties.

2   Methods

2.1   The Electrostatic Loudspeaker

An electrostatic loudspeaker is composed by three main parts: the diaphragm (which is the moving part), the stators (plates that drive the diaphragm) and the spacers (separating stators and diaphragm) [21]. The diaphragm is a thin and lightweight membrane coated with a conductive material that is placed between the stators (dividing the room in two chambers) and polarized by applying a potential V_p (see Figure 1). When a voltage is applied to the stators, the membrane is subjected to an electrostatic force which deforms it, driving the air in the room and eventually producing a sound. Indeed, if an attractive electrostatic force acts between a rigid plate and the diaphragm, the membrane will be pulled closer to the plate [22].

We suppose to use a mylar membrane (Poisson's ratio $\nu = 0.35$, Young's modulus $E = 0.62$ GPa, and mechanical stress $T = 297.2$ N/m). We assume to neglect the effect of the mechanical resistance of the conductive material covering the mylar on the deformation of the membrane. The membrane and the stators are assumed to be square, with size 0.1 m. The thickness of the membrane is
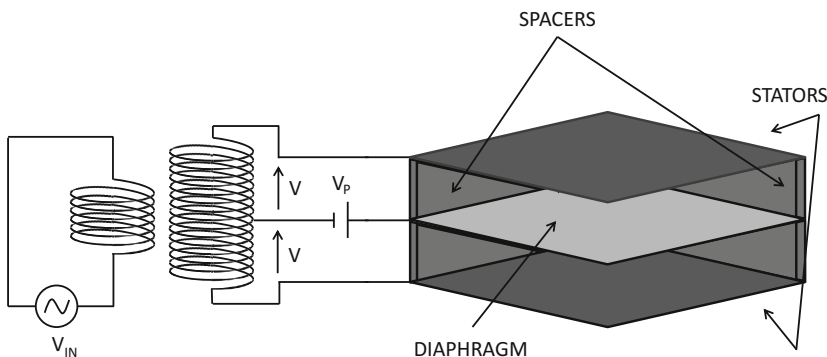
Fig. 1. Block diagram of the electrostatic loudspeaker

38 μm while the distance between the two plates is 760 μm (the membrane is supposed to be exactly in the middle of the plates).

2.2 Mathematical Model of the Electrostatic Loudspeaker

Referring to [23], the deformation of a thin diaphragm (or membrane) induced by an external pressure can be described as

$$c_2 \Delta^2 f - c_1 \Delta f = P \tag{1}$$

where $f = f(x, y)$ is the vertical deformation (which is a function of the x and y spatial variables), P is the pressure applied to the membrane, Δ and Δ^2 indicate the Laplacian and the Bilaplacian operators, respectively, c_1 and c_2 are constants which depend on the material

$$c_1 = T \tag{2}$$

$$c_2 = \frac{Ee^3}{12 \left(1 - \nu\right)} \tag{3}$$

where e is the thickness of the diaphragm.

Two boundary conditions are imposed [23]

– The membrane is locked along the edges, in which the deformation is null (Dirichlet homogeneous boundary condition)

$$f|_{\partial \Omega} = 0 \tag{4}$$

where $\partial \Omega$ is the boundary of the domain Ω in which the function $f(x, y)$ is defined.
– The membrane is elastically clamped (Neumann homogeneous condition)

$$\left. \frac{\partial f}{\partial n} \right|_{\partial \Omega} = 0 \tag{5}$$

where n is the versor orthogonal to $\partial \Omega$.

In the electrostatic loudspeaker, the pressure applied to the diaphragm is the difference between the two electrostatic forces applied by the plates

$$P = \frac{\epsilon}{2} \frac{(V - V_p)^2}{(\frac{d}{2} - f(x,y))^2} - \frac{\epsilon}{2} \frac{(-V - V_p)^2}{(\frac{d}{2} + f(x,y))^2} \tag{6}$$

where ϵ is the permittivity of the air, d is the distance between the plates, the membrane is assumed to be equidistant from them (in relaxed conditions), V_p is the polarization potential of the membrane and V is the potential between plates and membrane (see Figure 2). Introducing this expression in (1), the equation of a simply supported membrane under the effect of the electrostatic forces from the 2 plates is obtained [24] [25]

$$c_2 \Delta^2 f - c_1 \Delta f = \frac{\epsilon}{2} \frac{(V - V_p)^2}{(\frac{d}{2} - f(x,y))^2} - \frac{\epsilon}{2} \frac{(-V - V_p)^2}{(\frac{d}{2} + f(x,y))^2} \tag{7}$$

2.3   Mathematical Analysis of the Electrostatic Loudspeaker

Equation (7) was solved numerically, by the finite difference method [23]. Nodes were distributed uniformly (50 nodes for each space direction). Second and fourth order derivatives were discretized with a second and forth order precision, respectively. The nonlinear equation was solved iteratively, using the last estimated deformation to define the pressure for the subsequent iteration. As initial value, we used an approximated analytical solution, which is also useful to obtain some analytical estimates. Assuming that $f \ll d$, the right hand side of (7) was linearized as follows

$$\frac{\epsilon}{2} \frac{(V-V_p)^2}{(\frac{d}{2}-f)^2} - \frac{\epsilon}{2} \frac{(V+V_p)^2}{(\frac{d}{2}+f)^2} \approx \frac{2\epsilon}{d^2} \left((V - V_p)^2 \left(1 + 4\frac{f}{d}\right) + (-V - V_p)^2 \left(1 - 4\frac{f}{d}\right)\right) =$$
$$= \frac{2\epsilon}{d^2} \left(-4 V_p V + \frac{4}{d}(V^2 + V_p^2)f\right) \tag{8}$$

Equation (7) was then written as

$$\Delta^2 f - c\Delta f = \alpha + \beta f \tag{9}$$

where:

$$c = \frac{c_1}{c_2} \qquad \alpha = -\frac{8 V_p V \epsilon}{c_2 d^2} \qquad \beta = 16 \frac{\epsilon}{c_2 d^3}(V^2 + V_p^2) \tag{10}$$

Equation (9) is linear and can be solved introducing the following Fourier series

$$f(x,y) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} F_{mn} \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{L}$$
$$\alpha = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{mn} \sin \frac{n\pi x}{L} \sin \frac{m\pi y}{L} \quad \rightarrow \quad a_{mn} = \frac{16\alpha}{\pi^2 mn} \tag{11}$$

where the space variables are assumed to vary between 0 and L (size of the diaphragm and plates). Only the coefficients corresponding to odd values of the indexes m and n are different from zero (due to the symmetry of the problem). Note also that the sinusoids vanish on the boundary. The Fourier coefficients can

be obtained substituting the expressions (11) in (9) and using the orthogonality of the sinusoidal functions

$$F_{mn} = \frac{16\alpha L^4}{\pi^2 mn(\pi^4(n^2+m^2)^2 + c\pi^2 L^2(n^2+m^2) - \beta L^4)} \tag{12}$$

2.4   The Charge

The charge on the plates is the state variable of the memristor. We can compute it once we know the deformation of the membrane. The equation relating the charge on a surface and the potential is

$$-\nabla v \cdot n|_S = \frac{\sigma}{\epsilon} \tag{13}$$

where $v = v(x, y, z)$ is the potential, n is the versor normal to the surface S of the plate and σ is the surface density of the charge ($\int_S \sigma ds = q_\pm$, where q_+ and q_- are the total charges on the two plates). Thus, from equation (13), the total charge in a plate can be computed if we know the gradient of the electric potential along the direction normal to the surface of the plate. This requires to solve the following electrostatic problem in each of the two chambers in which the system is divided by the membrane

$$\begin{cases} -\Delta v = 0 \\ v|_{z=\pm L} = \pm V \\ v|_{z=f(x,y)} = V_p \end{cases} \tag{14}$$

For simplicity, as the deformations of the membrane are not so large, an approximate method was applied (instead of solving equation (14) and computing the integral of the surface charge density defined in (13)). Each chamber is considered as a capacitor, with capacitance

$$C_i = \epsilon \frac{S}{D_i} \,, \qquad i = 1, 2 \tag{15}$$

where D_i is the average distance between the ith plate and the membrane. As the membrane has an average position given by $\bar{f} = \frac{1}{S} \int_0^L \int_0^L f(x,y)dxdy$, we have the following distance between the ith plate and the membrane

$$D_i = \frac{d}{2} \mp \bar{f} \tag{16}$$

Thus, finally, from the definition of capacitance of a capacitor, the charges on the two plates can be computed

$$q_+ = C_1(V - V_p) \qquad q_- = -C_2(V + V_p) \tag{17}$$
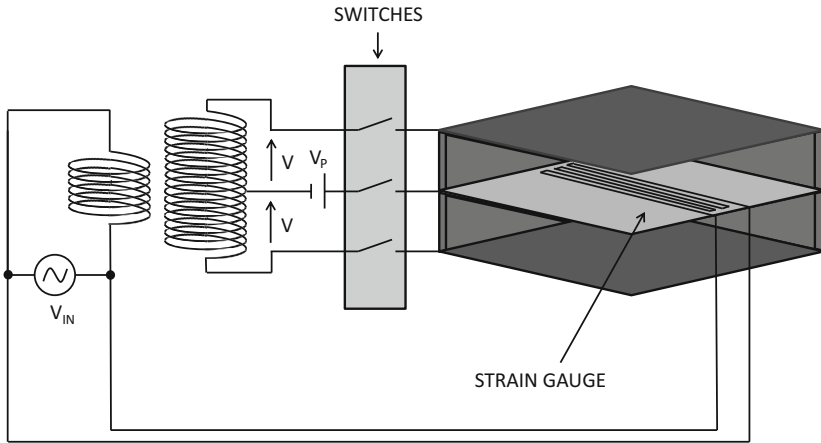
Fig. 2. Block diagram of the memristive system based on the electrostatic loudspeaker and the strain gauge

2.5   Electrostatic Loudspeaker Based Memristive System

A strain gauge is attached on the median line of the membrane (as shown in Figure 2). We suppose to neglect the effect of mechanical resistance of the strain gauge on the deformation of the membrane. When the membrane is deformed by the application of an electrostatic force, the strain gauge is consequently deformed, causing its electrical resistance to change. The resistance variation ΔR of the strain gauge is related to the deformation of the membrane by the following formula

$$\Delta R = G_F \frac{\Delta L}{L} R \tag{18}$$

where G_F is the gauge factor of the strain gauge, R is its rest resistance, L is its length at rest (equal to that of the membrane) and ΔL is the lengthening (induced by the deformation of the membrane).

From the electrical point of view, the strain gauge and the plates are connected together, according to the block diagram shown in Figure 2, obtaining a bipolar device (a required condition for being a memristive system).

3   Results

An example of simulation of the electrostatic loudspeaker is shown in Figure 3. The deformation of the membrane is shown for different applied input voltages (polarization voltage $V_p = 50$ V). In A), since the voltage applied to the top plate is positive with respect to that in the bottom plate, the membrane is deformed downward. In B), a positive voltage is again applied to the system (so that the membrane is deformed downward), but the deformation is smaller than that in A) since the applied voltage is lower. In C), no voltage is applied to the plates,
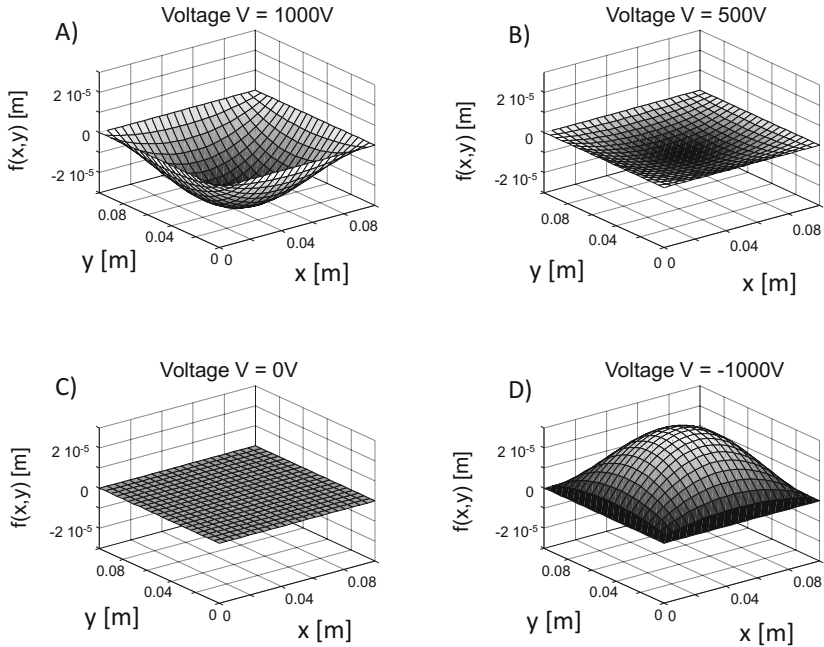
Fig. 3. Simulation of the electrostatic loudspeaker for different voltages: A) $V = 1000$ V, B) $V = 500$ V, C) $V = 0$ V, D) $V = -1000$ V

so that the membrane is not deformed. In D), a negative voltage is applied to the system so that the membrane is deformed upward.

Figure 4 shows a comparison between A) the numerical and B) the analytical solution of the problem for an applied voltage $V = 1000$ V and a polarization voltage $V_p = 50$ V. A few Fourier coefficients are enough to obtain a good approximation, as indicated in C).

Figure 5 shows the resistance variation of the device for different applied voltages in different conditions. The numerical solution of the mathematical model of the electrostatic loudspeaker was used. The input voltage V_{IN} of the system was a sine-wave with amplitude of 11.25 V and the winding turns ratio of the transformer was assumed to be equal to 200 (a sinewave with an amplitude of 2250 V is applied to the electrostatic loudspeaker). In A), different gauge factors are tested, keeping constant the polarization voltage applied to the membrane. In B), keeping constant the gauge factor of the strain gauge, different values of the membrane polarization voltage are tested.

The resistance variation of the strain gauge is very small, and can be read using a Wheatstone bridge, inserting the strain gauge in one of the four positions of the bridge, and resistors with a value equal to the rest resistance of the strain gauge in the others. Using the Wheatstone bridge, a differential voltage proportional to the resistance variation of the strain gauge and to the input voltage V_{IN} is
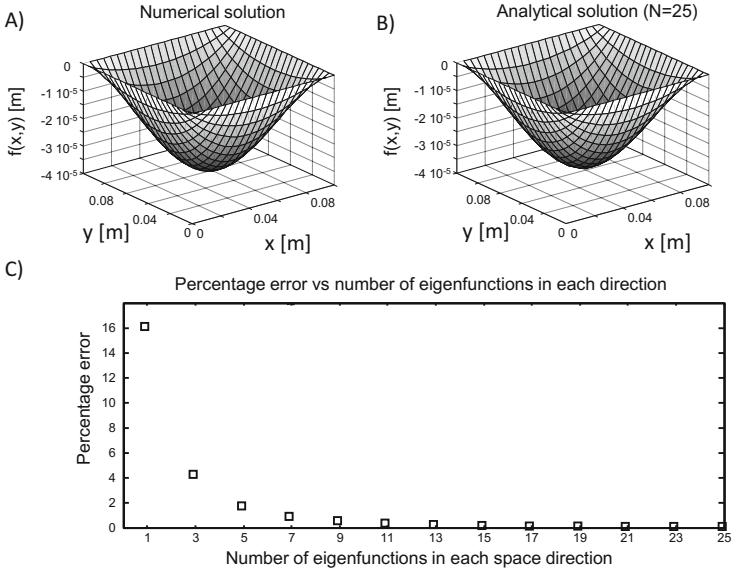
Fig. 4. Comparison between the numerical and analytical solution of the mathematical model of the electrostatic loudspeaker for an applied voltage equal to $V = 1000$ V: A) numerical solution, B) analytical solution, C) percentage error
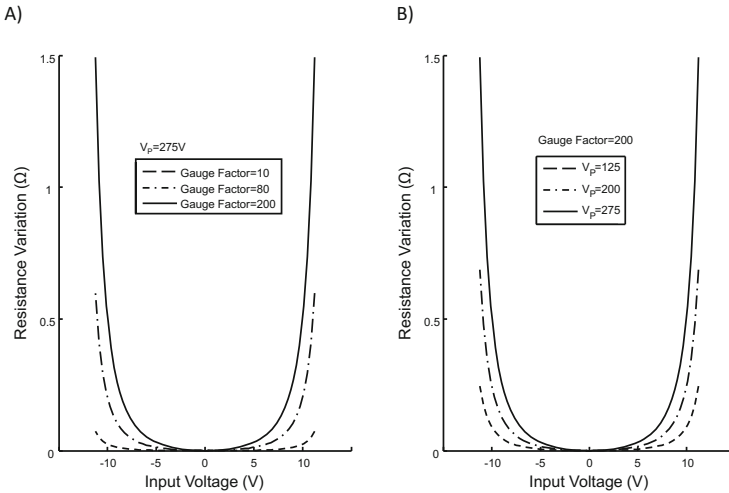
Fig. 5. Resistance variation of the memristive system in different conditions. A) Different gauge factors. B) Different values of the voltage applied to the membrane.
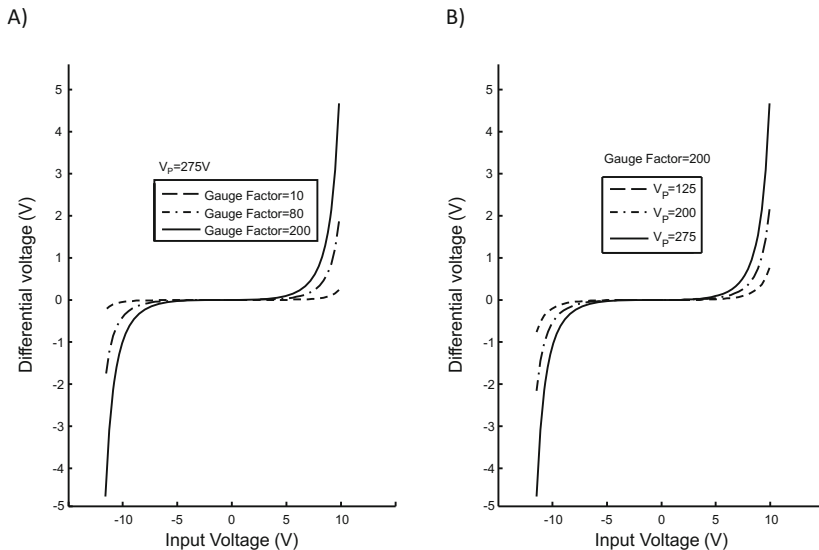
Fig. 6. Differential voltage of the Wheatstone bridge used to read the resistance variation in different conditions. A) Different gauge factors. B) Different values of the voltage applied to the membrane.

obtained. Figure 6 shows the differential voltage of the Wheatstone bridge for different applied voltages in different conditions. The applied voltage is equal to the input voltage V_{IN} of the system. Simulations are equal to those used for Figure 5. Different gauge factors are tested in A), with a constant value of polarization voltage V_P. Different values of the polarization voltage applied to the membrane are tested in B), keeping constant the gauge factor of the strain gauge. The device has the memory property of a memristive system, indeed the capacitance of the two plates of the electrostatic loudspeaker allows to store the last voltage value, when no voltage is applied to the device and the switches are open (see Figure 2).

4 Conclusions and Discussions

In this paper, a memristive system is proposed based on an electrostatic loudspeaker and a strain gauge. The system allows to store analog values of resistance. The proposed device can be used to add plasticity in analog hardware devices for spiking neural networks. A mathematical model was developed to provide quantitative indications in different conditions, reflecting different parameters of the system (regarding the mechanical properties of the membrane, the gauge factor of the strain gauge, the applied voltage, the geometry of the system, etc). This could help in the design of a system with the specific properties of interest. Simulations indicate the feasibility of the device as a memristive system. An experimental demonstrator is under development.

A)   **B)**
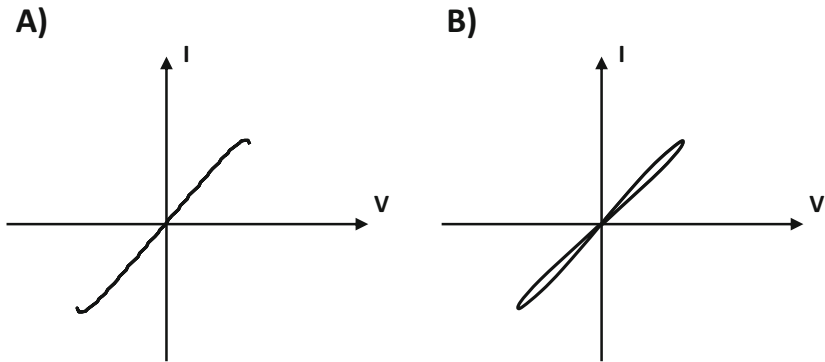
Fig. 7. Voltage-current characteristic of the electrostatic loudspeaker. A) Mathematical model with static behavior. B) Mathematical model with dynamic behavior.

Memristive systems should satisfy the directives of Prof. Chua [1]. In particular, they should show a pinched hysteresis loop in the Lissajoux figure (voltage-current characteristic), and the area of each lobe has to shrink as the frequency of the forcing signal increases. The system described in this paper has the memory capability typical of the memristive systems, it is a bipolar device and it is pinched in the Lissajoux figure; however, its voltage-current characteristic has no lobe for any frequency of the forcing signal, as shown in Figure 7A.

This is due to the approximations of the considered mathematical model of the electrostatic loudspeaker. For example, it does not consider the mechanical inertia of the membrane, which would introduce a dynamical behavior and the lobes, as shown in Figure 7B. In particular, the movement of the membrane would depend on the frequency of the forcing signal, and the area of the loops would decrease with the increasing of the frequency. Another approximation is included in our model, as the parasitic effects of the contacts are neglected. They include a resistance that is in series with the capacitor constituted by the stators, forming a low pass filter. A final consideration could be given on the possibility of miniaturizing the device. Mechanical devices have already been miniaturized in microelectromechanical systems (MEMS). For example, micro cantilevers have been widely used in sensing applications [26]. A similar technology could be exploited to miniaturize our membrane. Alternatively, a charged cantilever could replace our membrane, using again a capacitor to induce deformations and to store the information.

References

1. Chua, L.O.: Memristor: the missing circuit element. IEEE Transactions on Circuit Theory 18(5), 507–519 (1971)
2. Chua, L.O., Kang, S.M.: Memristive devices and systems. Proc. IEEE 64(2), 209–223 (1976)

3. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. Nature 14, 80–83 (2008)
4. Pickett, M.D., Strukov, D.B., Borghetti, J.L., Yang, J.J., Snider, G.S., Stewart, D.R., Williams, R.S.: Switching dynamics in titanium dioxide memristive devices. J. Appl. Phys. 106, 074508 (2009)
5. Biolek, Z., Biolek, D., Biolková, B.: Spice model of memristor with nonlinear dopant drift. Radio Eng. 18(2), 210–214 (2009)
6. Abdalla, H., Pickett, M.D.: SPICE modeling of Memristors. In: IEEE International Symposium on Circuits and Systems, pp. 1832–1835 (2011)
7. Kvatinsky, S., Friedman, E.G., Kolodny, A., Weiser, U.C.: TEAM - ThrEshold Adaptive Memristor Model. IEEE Trans. Circuits Syst. I 60(1), 211–221 (2013)
8. Ascoli, A., Corinto, F., Senger, V., Tetzlaff, R.: Memristor model comparison. IEEE Circuits and Systems Magazine 13(2), 89–105 (2013)
9. Krieger, J.H., Spitzer, S.M.: Non-traditional, non-volatile memory based on switching and retention phenomena in polymeric thin films. In: Proceedings of the 2004 Non-Volatile Memory Technology Symposium, p. 121 (2004)
10. Chanthbouala, A., Garcia, V., Cherifi, R.O., Bouzehouane, K., Fusil, S., Moya, X., Xavier, S., Yamada, H., Deranlot, C., Mathur, N.D., Bibes, M., Barthélémy, A., Grollier, J.: A ferroelectric memristor. Nature Materials 11(10), 860–864 (2012)
11. Wang, X., Chen, Y., Xi, H., Dimitrov, D.: Spintronic Memristor through Spin Torque Induced Magnetization Motion. IEEE Electron Device Letters 30(3), 294–297 (2009)
12. Huai, Y.: Spin-Transfer Torque MRAM (STT-MRAM): challenges and prospects. AAPPS Bulletin 18(6) (2008)
13. Pershin, Y.V., Di Ventra, M.: Spin memristive systems: spin memory effects in semiconductor spintronics. Physical Review B 78(11) (2008)
14. Di Ventra, M., Peotta, S.: Superconducting Memristors. Bulletin of the American Physical Society 59(1) (2014)
15. Wang, X.L., Shao, Q., Leung, C.W., Ruotolo, A.: Non-volatile, reversible switching of the magnetic moment in Mn-doped ZnO films. Journal of Applied Physics 113 (2013)
16. Corinto, F., Ascoli, A.: Memristive diode bridge with LCR filter. Electronics Letters 48(14), 824–825 (2012)
17. Chua, L.O.: Resistance switching memories are memristors. Appl. Phys. A 102(4), 765–783 (2011)
18. Troiano, A., Corinto, F., Pasero, E.: A Memristor circuit using basic elements with memory capability. In: Bassis, S., Esposito, A., Morabito, F.C. (eds.) Recent Advances of Neural Networks Models and Applications. SIST, vol. 26, pp. 117–124. Springer, Heidelberg (2014)
19. Versace, M., Chandler, B.: MoNETA: A Mind Made from Memristors. IEEE Spectrum (2010)
20. Johnsen, G.K.: An introduction to the memristor: a valuable circuit element in bioelectricity and bioimpedance. J. Electr. Bioimp. 3, 20–28 (2012)
21. Sanders, R.R.: The electrostatic loudspeaker design cookbook. Audio Amateur Press (1995)
22. Rangsten, P., Smith, L., Rosengren, L., Hök, B.: Electrostatically excited diaphragm driven as a loudspeaker. Sensors and Actuators A: Physical 52(1-3), 211–215 (1996)
23. Danaila, I., Joly, P., Kaber, S., Postel, M.: Elasticity: elastic deformation of a thin plate. In: An Introduction to Scientific Computing, pp. 151–164. Springer (2007)

24. Ventsel, E., Krauthammer, T.: Thin plates and shells: theory: analysis, and applications. CRC Press (2001)
25. Walter, D.P.: Formulas for stress, strain, and structural matrices. John Wiley & Sons Inc. (2004)
26. Vashist, S.K.: A Review of Microcantilevers for Sensing Applications. AZoJono – Journal of Nantechnology Online (2007)

Memristor Based Adaptive Coupling
for Synchronization of Two Rössler Systems

Mattia Frasca, Lucia Valentina Gambuzza,
Arturo Buscarino, and Luigi Fortuna

DIEEI, University of Catania,
Viale A. Doria 6, 95125 Catania, Italy
{mattia.frasca,lucia.gambuzza,
arturo.buscarino,luigi.fortuna}@dieei.unict.it

Abstract. Due to their intrinsic properties, memristors can be viewed as resistors in which the internal conductance is modulated by an external signal with the possibility to remember the previous state. This behavior is closely related to the functionality of the synapses and, for this reason, memristors have recently gained increasing attention also for their use as synapse in artificial neural networks for neuromorphic processing. As the most difficult step in the implementation of artificial neural networks is the realization of synapses, which often require a large number of transistors, the recent demonstration of the memory effect in memristors suggested a possible realization of synapses at the nanoscale with low power consumption and small size. In this work we explore the idea of using memristors as a synapse in a complex network to take advantage of the dynamics introduced by them and, in particular, propose a coupling scheme consisting of two HP memristors connected in antiparallel to achieve adaptive synchronization in two coupled Rössler systems.

Keywords: Memristor, adaptive synchronization, Rössler system.

1 Introduction

Recently, the discovery of resistive switching in thin films, with the understanding and demonstration of this mechanism in memristor devices, promoted the interest towards a variety of analog and digital applications based on these devices [1]. For its characteristics memristor has been widely explored as a component for applications in nonvolatile memories, in neuromorphic computing chips or to build simple Boolean logic gates [2]. Experimentally, memristive properties have been firstly discovered in films of titanium dioxide [3], where the resistance changes are due to an electric voltage bias applied to the device. The interest in memristors relies on the fact that it is able to remember its state even when the power is turned off and that can be easily realized with inexpensive fabrication techniques. In fact, besides the most widely used methods, such as nano-imprint lithography or atomic layer deposition, a breakthrough in the manufacturing of memristor was the fabrication of a device by spinning a titanium isopropoxide

solution on a flexible plastic substrate [4]. Memristor devices are not active components, so that, to use them in applications where they substitute a number of transistors by realizing a less dense architecture, is necessary to integrate it with a CMOS substrate that provides signal restoration and gain. Besides these applications, perhaps one of the most fascinating ways in which memristors can be used is to implement synapses for artificial neural networks.

In fact, the memristor characteristics such as its capability to remember the past state and to change the internal resistance, i.e., the weights, according to the external input, make it an ideal component to substitute the current CMOS synapses in neuromorphic chips. There are several examples in literature on the possible use of memristors as synapse in neural networks. In [5] the analogy between the memory behavior of memristor and the mechanism of biological memory that can occur in Physarum, led the authors to the implementation of memristive electronic circuits describing the ability of amoeba to learn. Jo et al. [6] demonstrate experimentally how important synaptic function, like spike timing dependent plasticity (STDP), can be achieved in a system made up of memristor synapses and complementary metal-oxide semiconductor (CMOS) neurons. An artificial synapse in which memristors are organized in a bridge-like structure capable of performing signed synaptic weights was proposed in [7,8]. Four identical memristors are placed in a Wheatstone bridge circuit, in this configuration the polarities of each memristor assure that, according to the sign of the input signal, the circuit can produce a positive or a negative output, and, therefore, positive or negative synaptic weight.

Our work proposes a new synapse based on memristors for solving the synchronization problem which arises in the coupling of two dynamical systems or in networks of nonlinear units [9]. In particular, the idea is to design a synapse that can be implemented at the circuit level with few components, and is effective for realizing an adaptive law for synchronization.

2 Model of Rössler Systems Coupled by Two Memristors in Antiparallel

The scheme adopted in this work is shown in Fig. 1 where two Rössler systems are coupled through a pair of memristors. We first discuss the model used to describe the memristor and, then, report the dynamical equations corresponding to the configuration of Fig. 1.

To simulate memristors, we use the model proposed by Strukov et al. [3] and derived on the basis of the first experimental observations of memristive behavior. The model refers to a memristor made by a TiO_2 thin film, of thickness D, doped with oxygen vacancies and sandwiched between two metal (Platinum) contacts. The memristor equations relating current through the device and voltage across it are:

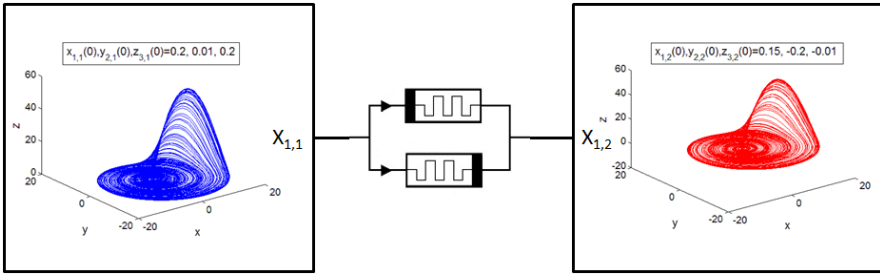$$v(t) = (R_{ON}w(t) + R_{OFF}(1 - w(t)))\, i(t) \tag{1}$$

Fig. 1. Scheme of two Rössler systems coupled through two memristors in antiparallel

where R_{ON} and R_{OFF} are technological parameters, and the variable $w(t)$ represents the width of the depletion zone of the memristor, normalized by the maximum width D, and therefore limited to values between 0 and 1. The doped region and the undoped one have different resistivity values (the doped region typically has a low value, while the undoped one an high value). When the width of the doped region is equal to the whole thickness (i.e., $w = 1$), the memristor has a resistance equal to R_{ON}, while, in the opposite case, when the undoped region covers the whole thickness of the device (i.e., $w = 0$), the memristor has a resistance equal to R_{OFF}. The memristor description is completed by the equation of the dynamics of $w(t)$:

$$\dot{w}(t) = \eta \frac{\mu_v R_{ON}}{D^2} F(w(t), i(t)) i(t) \tag{2}$$

where η ($\eta = 1$ or $\eta = -1$) accounts for the memristor polarity, that is, how it is connected in the circuit, and μ_v is the dopant mobility. $F(w, i)$ is the Biolek window function [10]: $F(x, i) = 1 - (x - stp(-i))^2$, where $stp(i) = 1$ when $i \geq 0$, and $stp(i) = 0$ when $i < 0$. For the sake of simplicity the term $M = (w(t) + \beta(1 - w(t)))$ with $\beta = R_{OFF}/R_{ON}$ which represents the normalized memristance of the device (that is, the value of the equivalent variable resistor, normalized by R_{ON}, associated with the memristor), is introduced.

In this paper we study the effect of the synapse made by two memristors in antiparallel on a pair of two Rössler systems as in Fig. 1. The study is carried out on the following dimensionless mathematical model:

$$\begin{cases} \frac{dx_{1,1}}{d\tau} = T_1[-x_{2,1} - x_{3,1} + k(\frac{x_{1,2}-x_{1,1}}{w_1+(1-w_1)\beta} + \frac{x_{1,2}-x_{1,1}}{w_2+(1-w_2)\beta}] \\ \frac{dx_{2,1}}{d\tau} = T_1[x_{1,1} + ax_{2,1}] \\ \frac{dx_{3,1}}{d\tau} = T_1[b + x_{3,1}(x_{1,1} - c)] \\ \frac{dx_{1,2}}{d\tau} = T_1[-x_{2,2} - x_{3,2} + k(\frac{x_{1,1}-x_{1,2}}{w_1+(1-w_1)\beta} + \frac{x_{1,1}-x_{1,2}}{w_2+(1-w_2)\beta})] \\ \frac{dx_{2,2}}{d\tau} = T_1[x_{1,2} + ax_{2,2}] \\ \frac{dx_{3,2}}{d\tau} = T_1[b + x_{3,2}(x_{1,2} - c)] \end{cases} \tag{3}$$

and

$$\begin{cases} \frac{dw_1}{d\tau} = \eta_1 F(w_1)\frac{x_{1,2}-x_{1,1}}{w_1+(1-w_1)\beta} \\ \frac{dw_2}{d\tau} = \eta_2 F(w_2)\frac{x_{1,2}-x_{1,1}}{w_2+(1-w_2)\beta} \end{cases} \tag{4}$$

where T_1 is a parameter characterizing the ratio of the time scale of the Rössler system as compared to that of the memristor, k is the coupling strength, and a, b, c are system parameters, ruling the dynamical behavior of each unit of the system.

3   Results

Numerical simulations have been performed on two identical Rössler system with parameters fixed such that chaotic behavior is assured, $a = 0.2$, $b = 0.2$ and $c = 9$. A standard Euler algorithm with integration step size equal to $\Delta t = 0.001$ has been used for numerical integration of Eqs. (3) and (4).

Fig. 2 shows that the system for an appropriate value of k and T_1 may be synchronized in an adaptive way. To illustrate this, we consider random initial conditions for the state variables of the Rössler systems and zero initial conditions for the memristor variables $w_1(0) = w_2(0) = 0$. When the circuits are uncoupled ($k = 0$), they are not synchronized and the memristor variables oscillate in the range from 0 to 1 as reported in Fig. 2(a)-(b). In correspondence of $k = 4$, instead, the two Rössler systems do synchronize and w_1 and w_2 reach a steady state that depends on their initial conditions (Fig. 2(c)-(d)). When a static coupling is used, synchronization depends on the coupling coefficient. For two Rössler systems, coupled as in Eqs. (3), there exists two constant α_1 and α_2 such that if the coupling coefficient belongs to the interval $[\alpha_1, \alpha_2]$, the two systems synchronize (for the system parameters indicated above $\alpha_1 = 0.093$ and $\alpha_2 = 2.307$ [11]). When the synapse of Eqs. (4) is used, the coupling strength, given by $k(\frac{1}{w_1+(1-w_1)\beta} + \frac{1}{w_2+(1-w_2)\beta})$, is not constant as w_1 and w_2 may evolve in time. In the case shown in Fig. 2 the system starts from a value of the coupling strength $k(\frac{1}{w_1(0)+(1-w_1(0))\beta} + \frac{1}{w_2(0)+(1-w_2(0))\beta})$ which is out of the synchronization range, so that synchronization without adaptation would not be possible. Thanks to the evolution of the memristor variables, the coupling strength evolves such that $\alpha_1 < k(\frac{1}{w_1(t)+(1-w_1(t))\beta} + \frac{1}{w_2(t)+(1-w_2(t))\beta}) < \alpha_2$; this is a condition ensuring that the synchronization error decreases until it reaches zero and w_1 and w_2 reach a steady state value.

The synapse allows the adaptation toward synchronization even when the coupling strength $k(\frac{1}{w_1(0)+(1-w_1(0))\beta} + \frac{1}{w_2(0)+(1-w_2(0))\beta})$ is initially beyond the second transition point α_2. The case study is illustrated in Fig. 3, where $k = 8$ and the initial condition of the memristor variables are $w_1(0) = w_2(0) = 1$. Under these conditions, $k(\frac{1}{w_1(0)+(1-w_1(0))\beta} + \frac{1}{w_2(0)+(1-w_2(0))\beta}) > \alpha_2$. Hence, also in this case synchronization without adaptation would not be possible. The memristor variables now evolve in such a way that the coupling strength decreases and, once again, $\alpha_1 < k(\frac{1}{w_1(t)+(1-w_1(t))\beta} + \frac{1}{w_2(t)+(1-w_2(t))\beta}) < \alpha_2$, when the steady state is reached.
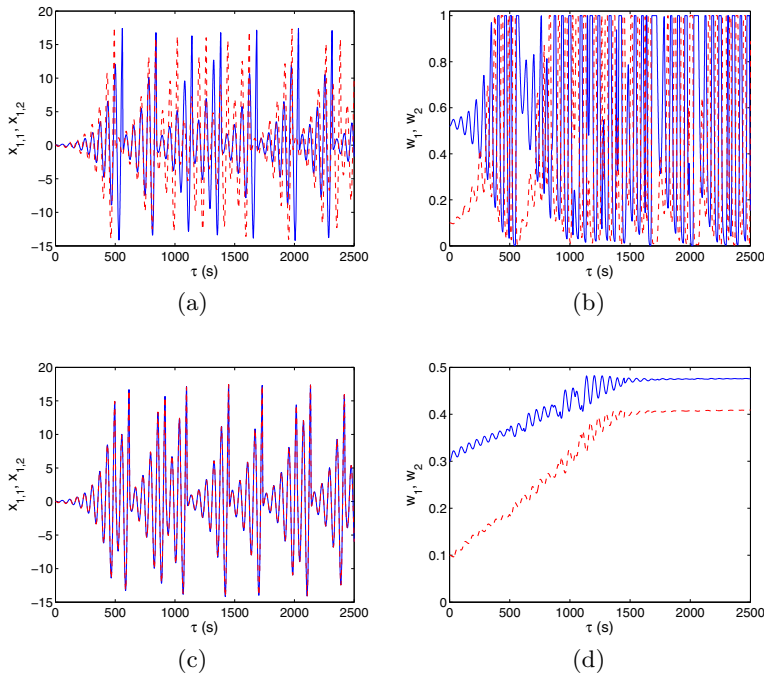
Fig. 2. Behavior of two Rössler systems coupled through memristors for $k = 0$ (a)-(b) and $k = 4$ (c)-(d): (a)- (c) state variables $x_{1,1}$ (continuous line) and $x_{1,2}$ (dashed line); (b)- (d) trend of memristor variables w_1 (continuous line) and w_2 (dashed line)
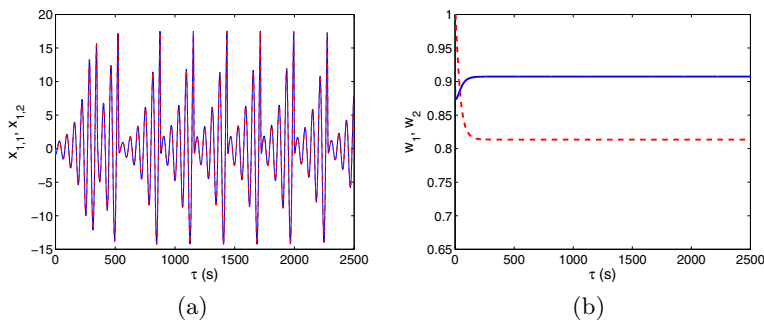

Fig. 3. Behavior of two Rössler systems coupled through memristors for $k = 8$: (a) state variables $x_{1,1}$ (continuous line) and $x_{1,2}$ (dashed line); (b) trend of memristor variables w_1 (continuous line) and w_2 (dashed line)

4   Conclusions

In this work a synapse based on two memristors connected in antiparallel has been used to study adaptive synchronization in a pair of Rössler systems. Although there exist models of memristive devices, able to accurately reproduce many of the physical phenomena appearing in these components (often strictly related to the fabrication technique and layout of the specific device), here a trade-off between the level of physical representation and computational efforts has led the choice of the memristor model: the so-called nonlinear drift model has been adopted in view of generalization of the approach to networks of N coupled oscillators, where the number of synapses would grow as $N(N-1)$.

In the depicted framework, adaptive synchronization has been shown. In particular, the synapse allows to reproduce the behavior typically found in other adaptive laws, where the weight of the connection is started from zero and grows up until it is able to guarantee synchronization. However, we have demonstrated that the synapse may also decrease the value of the weight, as the coupling scheme adopted for the two Rössler systems is such that, in the static case, either low or high values of the coupling strength do not guarantee synchronization.

References

1. Strukov, D.B., Kohlstedt, H.: Resistive switching phenomena in thin films: Materials, devices, and applications. MRS Bulletin 37(02), 108–114 (2012)
2. Gale, E., de Lacy Costello, B., Adamatzky, A.: Boolean logic gates from a single memristor via low-level sequential logic. In: Mauri, G., Dennunzio, A., Manzoni, L., Porreca, A.E. (eds.) UCNC 2013. LNCS, vol. 7956, pp. 79–89. Springer, Heidelberg (2013)
3. Strukov, D.B., Snider, G.S., Stewart, D.R., Williams, R.S.: The missing memristor found. Nature 453, 80–83 (2008)
4. Hackett, N.G., Hamadani, B., Dunlap, B., Suehle, J., Richter, C., Hacker, C., Gundlach, D.: A flexible solution-processed memristor. IEEE Electron Device Lett. 30(7), 706–708 (2009)
5. Pershin, Y.V., La Fontaine, S., Di Ventra, M.: Memristive model of amoeba learning. Physical Review E 80, 021926-1–6 (2009)
6. Jo, S.H., Chang, T., Ebong, I., Bhadviya, B.B., Mazumder, P., Lu, W.: Nanoscale memristor device as synapse in neuromorphic systems. Nano Lett. 10(4), 1297–1301 (2010)
7. Kim, H., Sah, M.P., Yang, C., Roska, T., Chua, L.O.: Memristor bridge synapse. Proc. IEEE 100(6), 2061–2070 (2012)
8. Adhikari, S.P., Yang, C., Kim, H., Chua, L.O.: Memristor bridge synapse-based neural network and its learning. IEEE Trans. Neural Netw. 23(9), 1426–1435 (2012)
9. De Lellis, P., di Bernardo, M., Garofalo, F.: Synchronization of complex networks through local adaptive coupling. Chaos: An Interdisciplinary Journal of Nonlinear Science 18(3), 037110–037110 (2008)
10. Biolek, Z., Biolek, D., Biolkova, V.: SPICE Model of Memristor with Nonlinear Dopant Drift. Radioengineering 18(2), 210–214 (2009)
11. Huang, L., Chen, Q., Lai, Y.C., Pecora, L.M.: Generic behavior of master-stability functions in coupled nonlinear dynamical systems. Physical Review E 80(3), 036204 (2009)

Author Index