35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Aged-related Changes in Brain Activity Classification with respect to Age by means of Graph Indexes

Petti M, Toppi J, Pichiorri F, Cincotti F, Salinari S, Babiloni F, Astolfi L, Mattia D.

Abstract— **Recent studies have investigated changes in the human brain network organization during the normal aging. A reduction of the connectivity between brain areas was demonstrated by combining neuroimaging technologies and graph theory. Clustering, characteristic path length and small-worldness are key topological measures and they are widely used in literature. In this paper we propose a new methodology that combine advanced techniques of effective connectivity estimation, graph theoretical approach and classification by SVM method. EEG signals recording during rest condition from 20 young subjects and 20 mid-aged adults were studied. Partial Directed Coherence was computed by means of General Linear Kalman Filter and graph indexes were extracted from estimated patterns. At last small-worldness was used as feature for the SVM classifier. Results show that topological differences of brain networks exist between young and mid-aged adults: small-worldness is significantly different between the two populations and it can be used to classify the subjects with respect to age with an accuracy of 69%.**

I. INTRODUCTION

In the last years, several studies demonstrated that by combining a variety of different imaging technologies (functional MRI, EEG, MEG) with graph theoretical analysis it is possible to characterize the topological properties of human brain networks. Among all the available graph indexes, clustering (overall integration), characteristic path length (local structure) and small-worldness, defined as ratio between them, represented some key topological metrics [1] describing the organization of information flows in a network. These metrics allowed to distinguish different classes of network such as regular, small-world, and random networks. In fact, a small-world network has a shorter path length than a regular network (characterized by high clustering and long path lengths) but a greater local interconnectivity than a random network (characterized by low clustering coefficient and short path lengths).

Several anatomical, functional neuroimaging and electro-physiological studies demonstrated that networks inferred from healthy individuals are characterized by an optimal small-world organization [2, 3]. In contrast, brain pathology and changes in cerebral structures related to aging effects generally led to the estimation of connectivity patterns with characteristics more similar to ‘‘random’’ networks [4, 5, 6].

Different approaches based on graph theory have been used to describe the effects of ageing on cerebral processes at rest or during the execution of cognitive tasks. In particular, the transition from childhood (8-12 years) to adulthood (21-26 years) is suggested to be characterized by a reduction of overall connectivity (decreased clustering and increased path length) [5]. Furthermore a description of the resting state networks during three different ages (children, mid-aged, elderly) was proposed by Zhu et al. [7] who provided their characterization also in terms of hemispheres asymmetry.

The aim of this study was to propose a new methodology able to investigate differences in resting state networks related to subjects' age by means of state of the art graph theory indexes. In particular we extracted some salient indexes synthesizing the architecture of the connectivity networks elicited during the rest condition by two populations of different ages (young and mid-aged adults). The indexes resulting statistically different between the two groups were used as features for classifying resting state patterns in relation to the age of the subject. The real novelty in respect to what found previously was that the proposed method is able to classify age-related differences at single subject level only by using indexes extracted form 1 minutes of eyes-closed condition.

II. METHODS

A. Partial Directed Coherence

The PDC [8] is a full multivariate spectral measure, used to determine the directed influences between any given pair of signals in a multivariate data set. This estimator was demonstrated to be a frequency version of the concept of Granger causality [9].

It is possible to define PDC as:

$$\pi_{ij}(f) = \frac{\Lambda_{ij}(f)}{\sqrt{\sum_{k=1}^{N} \Lambda_{ki}(f)\Lambda_{ki}(f)}} \quad , \quad \sum_{n=1}^{N} \left|\pi_{ni}(f)\right|^2 = 1 \qquad (1)$$

where *Λ(f)* is a matrix containing the coefficients of associated Multivariate Autoregressive (MVAR) model.

In this study we used the square formulation of PDC due to its higher accuracy and stability.

The coefficients of MVAR model were estimated by means of General Linear Kalman Filter (GLKF), an algorithm developed for time-varying connectivity

*Research partially supported by the European ICT Program FP7-ICT-2009-4 Grant Agreement 287320 CONTRAST and by the Italian Ministry of Health (RF-2010-2319611).

L. Astolfi (Corresponding Author. Phone: +39 06 51501466; e-mail: laura.astolfi@uniroma1.it), M. Petti, J. Toppi, F. Cincotti, are with the Department of Computer, Control, and Management Engineering, Univ. of Rome “Sapienza” and with Fondazione Santa Lucia Hospital, Rome, Italy.

D. Mattia and F. Pichiorri are with Fondazione Santa Lucia, Rome, Italy.

S. Salinari is with the Department of Computer, Control, and Management Engineering, Univ. of Rome “Sapienza”, Rome, Italy.

F. Babiloni is with the Dept. of Physiology and Pharmacology of the University of Rome “Sapienza” and with Fondazione Santa Lucia, Rome, Italy

estimation but which can be also applied in the stationary case with an appropriate choice of the adaptation constants. We selected the GLKF approach for including all the electrodes in the estimation process without applying any a-priori selection of cerebral sources.

B. General Linear Kalman Filter

In the GLKF an adaptation of the Kalman Filter to the case of multi-trial time series is provided. In particular, the equations at the basis of the algorithm are:

$$\begin{aligned} Q_n &= G_{n-1} Q_{n-1} + V_n \\ O_n &= H_n Q_n + W_n \end{aligned} \tag{2}$$

where n denotes the time sample, O_n represents the observation, Q_n is the state process, H_n and G_n are the transition matrices and V_n and W_n are the additive noises. To obtain the connection with the time-varying MVAR it is necessary to make the following associations:

$$Q_n = \begin{bmatrix} \Lambda_1(n)^T \\ \vdots \\ \Lambda_p(n)^T \end{bmatrix}, \quad O_n = \begin{pmatrix} x_1^{(1)}(n) & \cdots & x_d^{(1)}(n) \\ \vdots & \ddots & \vdots \\ x_1^{(K)}(n) & \cdots & x_d^{(K)}(n) \end{pmatrix} = Y_n \tag{3}$$

$$G_{n-1} = I_{dp} \quad , \quad H_n = \left(O_{n-1}, \quad \cdots, \quad O_{n-p} \right) \tag{4}$$

where K denotes the number of trials, whereas d is the dimension of the measured process. The details of the algorithm are provided in [10]. For the stationary case we select the two adaptation constants as follows c_1=0.001 and c_2=0.001.

C. Graph Theory Approach

A graph is a mathematical object consisting in a set of vertices (or nodes) linked by means of edges (or connections) indicating the presence of some sort of interaction between the vertices. The structure of the investigated graph is described by means of an adjacency matrix G. When a directed edge exists from the node i to j, the corresponding entry of the adjacency matrix is $G_{ij} = 1$, otherwise $G_{ij} = 0$. Several indices based on the elements of such matrix can be extracted for the characterization of the main properties of investigated networks [11].

Characteristic Path Length. The characteristic path length is the average shortest path length in the network, where the shortest path length between two nodes is the minimum number of edges that must be traversed to get from one node to another. It can be defined as follows:

$$L = \frac{1}{n} \sum_{i \in N} L_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}}{n-1} \tag{5}$$

where L_i is the average distance between node i and all other nodes and d_{ij} is the distance between node i and node j [12].

Clustering Coefficient. The clustering coefficient describes the intensity of interconnections between the neighbors of a node [13]. It is defined as the fraction of triangles around a node or the fraction of node's neighbors that are neighbors of each other. The binary directed version of Clustering Coefficient is defined as follows [12]:

$$C = \frac{1}{n} \sum_{i \in N} C_i = \frac{1}{n} \sum_{i \in N} \frac{t_i}{(k_i^{out} + k_i^{in})(k_i^{out} + k_i^{in} - 1) - 2\sum_{j \in N} g_{ij} g_{ji}} \tag{6}$$

where t_i represents the number of triangles involving node i, k_i^{in} and k_i^{out} are the number of incoming and outcoming edges of nodes i respectively and g_{ij} is the entry ij of adjacency matrix.

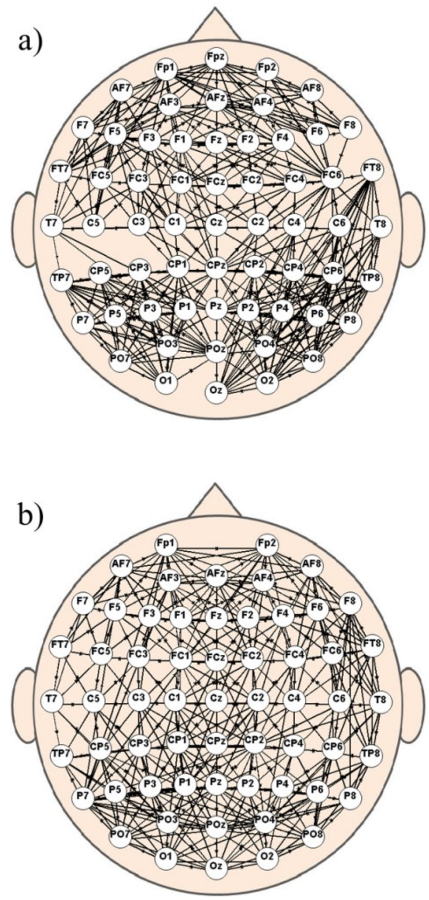

Figure 1. Connectivity patterns elicited during the eyes-closed condition by one representative young subject (a) and by one representative mid-aged adults (b) in the frequency range of 1-30 Hz. Patterns are represented on a scalp model seen from above with the nose pointing the upper part of the page. Each arrow codes the connection between two electrodes.

Small-Worldness. A network G is defined as small-world network if $L_G > L_{rand}$ and $C_G >> C_{rand}$ where L_G and C_G represent the characteristic path length and the clustering coefficient of a generic graph and L_{rand} and C_{rand} represent the correspondent quantities for a random graph [1]. On the basis of this definition, a measure of small-worldness of a network can be introduced as follows:

$$S = \frac{C_G / C_{rand}}{L_G / L_{rand}} \tag{7}$$

D.Support Vector Machine

The Support Vector Machine (SVM) was first proposed by Vapnik and has since attracted a high degree of interest in the machine learning research community [14]. SVMs are supervised learning models used for classification and regression analysis. In order to perform a binary classification (two separate classes), this method needs of training data, each marked as belonging to one of two categories, for introducing a separating hyperplane: this hyperplane must maximize the margin between the two classes and it is known as the optimum separating hyperplane. The details of the method are provided in [15].

E. Experimental design

20 young healthy subjects (age: 23.8 ± 1.05 years; 10 female) and 20 healthy mid-aged adults (age: 46.05 ± 5.27; 10 female) took part in the study. EEG signals were recorded during rest condition with eyes closed for one minute. A 64-channel system with a sampling frequency of 200 Hz (BrainAmp, Brainproducts GmbH, Germany) was used.

F. Connectivity Analysis and Graph Indexes Classification

After band-pass filtered (1-45 Hz + 50 Hz Notch filter), ocular artifacts were rejected by means of Independent Component Analysis. EEG traces were then segmented in epochs of 1s each in order to increase the robustness of methodologies applied in the following and residual artifacts were removed. Despite in this study we analyzed a stationary condition (resting state), we used GLKF for PDC computation to avoid the limitation on the number of signals included simultaneously in the stationary estimation approach. The achieved estimations were averaged over frequency range of 1-30 Hz. Then a binary directed adjacency matrix was extracted from the connectivity matrix estimated for each subject by applying a threshold able to maintain the 20% of the stronger connections.

In order to quantify the differences related to the network architecture of the two age-related groups of subjects, the graph indexes described above were extracted and the following steps were performed:

1. Statistical comparison (two-sample Student's t-test) for a significance level of 5% was computed between indexes from young and mid-aged subjects networks;
2. Classification of extracted features by means of SVM classifier with quadratic kernel.

The first step was performed for investigating which indexes were significantly different between the populations: this step was important for the right choose of the features to be used in classification process. For the second step a Leave One Out approach has been implemented to perform the classification. In particular, graph indexes extracted from 1 subject were used for classifying him in one of the two groups, using the indexes achieved by the other 38 subjects (19 young and 19 mid-aged adults) as training data for SVM classifier. Each subject was tested singularly for 50 times. During each iteration, the algorithm returned score equal to 1 for right classification and 0 otherwise. For each subject the classification performance was obtained among the iterations and then total, young and mid-aged performances were computed by performing the average of the performances among all subjects, young subjects and mid-aged adults respectively.

III. RESULTS

After the signal pre-processing, connectivity patterns were estimated by means advanced techniques of effective connectivity estimations and were averaged in the frequency range 1-30 Hz.

Connectivity networks for two representative subjects belonging to the two different groups (young, mid-aged adults) are reported in Figure 1. Patterns are represented on a scalp model seen from above with the nose pointing to the upper part of the page. Each arrow codes the connection between two electrodes.

Figure 1 shows that the network of young adult is characterized by a higher and more specific organization. Some electrodes (e.g. FT8, CP6, TP7, POz) showed an important role in the information flows of the network: they are characterized by a large number of short and long-range connections passing through them. Instead, mid-aged adult pattern was characterized by a more random topography: it is visible that the path length is greater than the one of young subject because there are few long-range connections and so the communication between spatially distant nodes goes through multiple edges.

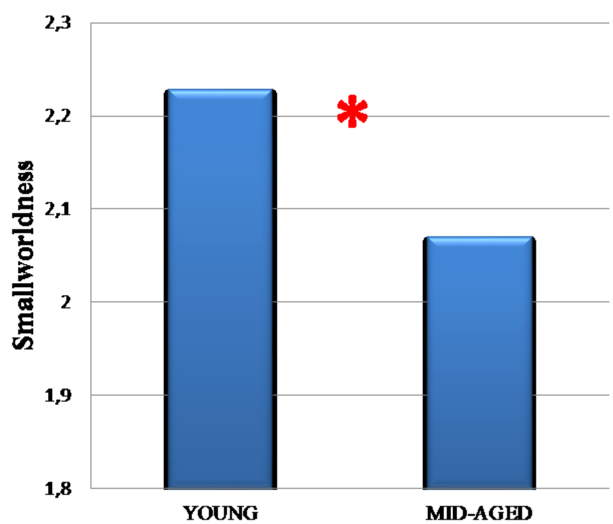

Figure 2. Bar diagram reporting the small-worldness mean value achieved for the two age-related groups. Two-sample t test for a significance level of 5% was performed between the two populations: the symbol (*) highlights a significant difference between the two groups (p=0.003).

Once extracted the correspondent adjacency matrix for each subject, graph indexes such as clustering, path length and small-worldness, were computed for extracting the topographic properties from the patterns.

Bar diagram in figure 2 shows mean values of the small-worldness achieved for the two groups. In order to evaluate if a significant difference exists between the two populations, a t-test for a significance level of 5% was performed on the small-worldness index. The statistical comparison reveals a significantly lower value of the index in the mid-aged group (p=0.003) when compared to the young group.

At last, small-worldness was used as a feature to classify each subject by means of SVM classifier. The total classification accuracy achieved by means of this process was 69%: in particular 73% for young subject and 64 % for mid-aged adults.

IV. DISCUSSION

In this paper we described a methodology able to combine advanced techniques of effective connectivity estimation, graph theoretical approach and classification by SVM method. The application of GLKF for the connectivity estimation to EEG signals showed different brain organizations in young and mid-aged adults in the range frequency 1-30 Hz and graph theoretical approach was useful to quantify these changes. In particular, after the indexes extraction from the estimated patterns, statistical comparison between the two studied populations shows that the small-worldness decrease significantly in mid-aged group: these results demonstrate that human brain networks are less well organized in middle age in respect to young age. The used algorithm for the classification is based on Leave One Out method and uses SVM classifier: with small-worldness as a feature for the classification, we obtained performance of 69%.

Results show that this approach allows to evaluate age-related differences from different populations and in particular the main novelty of the proposed method is the ability to classify these differences at single subject level only by using indexes extracted from 1 minute of eyes-closed condition. Furthermore the described results suggest that the normal ageing impacts the topological configuration of resting-state networks towards a decrease of small-worldness in according with literature in which the transition from childhood to adulthood is described with a reduction of overall connectivity (decreased clustering and increased path length) [5].

V.CONCLUSION

The methodological steps performed for the analysis purposed in this work, have been revealed as valid procedure for the description of differences related to the network architecture of the two age groups. Furthermore the implemented algorithm for the classification is characterized by good performance. However this methodology should be improved by means of the use of more features in the classification approach.

REFERENCES

[1] M.D. Humphries and K. Gurney, "Network 'small-world-ness': a quantitative method for determining canonical network equivalence," *PLoS ONE*, vol. 3, no. 4, p. e0002051, 2008.

[2] Y. He, Z.J. Chen, A.C. Evans : "Small-world anatomical networks in the human brain revealed by cortical thickness from MRI". Cereb Cortex 17:2409–2419, 2007.

[3] O. Sporns, G. Tononi, G.E. Edelman: "Connectivity and complexity:The relationship between neuroanatomy and brain dynamics". Neural Netw 13:909–922., 2000.

[4] S. Micheloyannis, E. Pachou, C.J. Stam, M. Breakspear, P. Bitsios, M. Vourkas, S. Erimaki, M. Zervakis: "Small-world networks and disturbed functional connectivity in schizophrenia". Schizophr Res 87:60–66, 2000.

[5] S. Micheloyannis, M. Vourkas, V. Tsirka, E. Karakonstantaki, K. Kanatsouli, and C. J. Stam, "The influence of ageing on complex brain networks: a graph theoretical analysis," Hum Brain Mapp, vol. 30, no. 1, pp. 200–208, Jan. 2009.

[6] C.J. Stam, B.F. Jones, G. Nolte , M. Breakspear, Scheltens Ph: "Small-world networks and functional connectivity in Alzheimer's disease". Cereb Cortex 17:92–99, 2007.

[7] C. Zhu, X. Guo, Z. Jin, J. Sun, Y. Qiu, Y. Zhu, and S. Tong, "Influences of brain development and ageing on cortical interactive networks," Clin Neurophysiol, vol. 122, no. 2, pp. 278–283, Feb. 2011.

[8] K. Sameshima and L. A. Baccalá, "Using partial directed coherence to describe neuronal ensemble interactions," *Journal of Neuroscience Methods*, vol. 94, no. 1, pp. 93–103, 1999.

[9] C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969.

[10] T. Milde, L. Leistritz, L. Astolfi, W. H. R. Miltner, T. Weiss, F. Babiloni, and H. Witte, "A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials," *Neuroimage*, vol. 50, no. 3, pp. 960–969, Apr. 2010.

[11] O. Sporns, D. R. Chialvo, M. Kaiser, e C. C. Hilgetag, «Organization, development and function of complex brain networks», *Trends in Cognitive Sciences*, vol. 8, n°. 9, p. 418–425, 2004.

[12] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.

[13] G. Fagiolo, "Clustering in complex directed networks," *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 76, no. 2 Pt 2, p. 026107, Aug. 2007.

[14] V. Vapnik, "The Nature of Statistical Learning Theory". NY: Springer-Verlag, 1995.

[15] Andrew R. Webb, "Statistical Pattern Recognition" ,Malvern, UK: John Wiley & Sons, LTD, 2002.