Combining Graph Analysis and Recurrence Plot on fMRI Data

Angela Lombardi, Pietro Guccione, Luigi Mascolo  
Department of Electrical and Information Engineering  
Technical University of Bari  
70125 Bari, Italy

Paolo Taurisano, Leonardo Fazio  
Department of Neurological and Psychiatric Sciences  
University of Bari  
Bari, Italy

Giovanni Nico  
Istituto per le Applicazioni del Calcolo  
National Research Council  
70124 Bari, Italy

Abstract—In this work we investigate on the nonlinear properties of the brain networks using Graph Analysis and Cross Recurrence Plot. The nonlinear dynamics of the brain is analyzed using time series coming from fMRI data. Two groups of human subjects, one affected by schizophrenia and the other of healthy controls, are imaged during the completion of a working memory task. To examine the spatio-temporal properties of the BOLD signal, nonlinear recurrence properties are extracted from the time series of the most relevant voxels, using the cross recurrence plots and the corresponding measures. Then, a graph is built using such measures as weights between different brain regions (the nodes). The purpose of the paper is to give a description of the most relevant functional areas activated during the task completion and to capture the differences between the groups. Results are promising, since the methodology is still to be fully developed and explored.

Keywords—Recurrence Plot; Graph Analysis; functional Magnetic Resonance Imaging; Schizophrenia

I. INTRODUCTION

The term "schizophrenia" was used for the first time by the psychiatrist Bleuler to indicate the cleavage of mental functions reflecting the onset of the disease [1]. It is a psychiatric disease characterized by the persistence of symptoms of abnormal thinking, behavior and affect with strong mismatch of the person [2]. Although genetic, environmental, psychological and social factors seem to contribute significantly to the development of schizophrenia, organic and biological causes of the disease have not yet been found. Several scientists have investigated the evolution of schizophrenia starting from the common idea of "disconnection": the behavioral and cognitive dysfunction of a patient may be explained in terms of disruption of several functional areas of the brain that work together to achieve a particular function [3,4]. The dynamic neural process underlying a particular function would be broken or damaged, changing the normal functional connectivity of the brain.

During the last few decades neuroimaging techniques have been improved significantly and, thanks to their high spatiotemporal resolution, it has been possible to obtain large datasets of anatomical or functional connection patterns. For this reason, complex networks methods [5] were adopted to explore the anatomical and functional relationships between the different elements of the networks of the whole human brain. The complex networks analysis is a graph-based method that allows to describe and quantify the topological properties of networks through a small number of easily computable measures [6]. Such analysis is also able to detect both changes in connectivity on a large scale and at the local level [7].

Among all the existing neuroimaging techniques, functional Magnetic Resonance Imaging (fMRI) is the most used to investigate the functional relationships between different brain networks, thanks to a good balance between spatial and temporal resolution [8]. Recently, there has been a rapid development in the brain networks analysis on fMRI data, to explore the functional organization of the human brain [9] and to carry out studies on topological configurations in the presence of psychiatric disorders [10] and neurodegenerative diseases [7,11]. In most of these studies, the network analysis is realized considering groups of voxels or regions of interest (ROI) as nodes of the brain graph, the link being defined by linear measurements of correlation or synchronization between the nodes. However, the links represent functional interactions between areas of a system that, in more accurate models, can be considered complex and nonlinear [12]. For this reason the use of nonlinear methods may be more effective to learn about the integration processes between the various brain anatomical locations.

As in [13], we have considered the fMRI data as projections of a nonlinear dynamic system evolving in time and then we have explored the evolution of this system through its manifold, a representation of the dynamics of hidden variables that shape the phenomenon in a defined hyperspace. To investigate the time-dependent behavior of two fMRI time series we use the Cross Recurrence Plots (CRP), an extension of the method of Recurrence Plots that allows the comparison of the phase space trajectories of two processes in the same phase space [14]. These plots are based on the properties of recurrence of nonlinear dynamical systems, whose basic concepts are derived from the Poincare ‘Recurrence theorem’ and can be qualitatively summarized in the following two
properties: (i) similar situations often evolves in a similar way, (ii) some situations occur over and over again. Hence we use measures of complexity, the Cross Recurrence Quantitative Analysis (CRQA) extracted from the CRP, to assess the similarity of the systems [14, 15]. Some of these measures are then used to build the brain graphs from which various complex network measures are extracted and compared.

In this paper we applied this very general approach to a set of schizophrenic subjects and a corresponding set of healthy subjects (used for control). The purpose of this preliminary analysis (not all the CRQA measures, nor all the graph measures were explored) is to show that regions of interest, that are considered relevant for the mental disorder under investigation, may be highlighted; moreover differences on measures have been found between the group of subjects.

II. FRAMEWORK

A. Clinical data

9 patients suffering from schizophrenia and 9 healthy control subjects have performed an alternating block task paradigm, consisting of a visual-motor condition and a working memory (WM) condition, the visual N-back WM task [16]. A gradient echo BOLD echo-planar imaging pulse sequence has been used to acquire 120 images, one every 2 sec. Each functional image consists of 20°6 mm thick axial slices covering the entire cerebrum and most of the cerebellum (matrix=64×64 pixel for each slice). Data have been preprocessed using the SPM8 software package (http://www.fil.ion.ucl.ac.uk/spm/). Images for each subject were aligned to the first volume in the time series to correct for head motion, spatially normalized into a standard space and smoothed by a 10 mm isotropic 3D Gaussian kernel (more details are in [17]).

B. Statistics and selection of Regions of Interest

First of all, we used a measure to check the degree of stochasticity of the time evolution of the system. This operation was necessary to ensure that the fMRI series were suitable for a statistical nonlinearity test in the successive steps. In the Lempel-Ziv complexity measure, the time series is transformed by quantization into a finite sequence of symbols and then the number of different sub-strings within the sequence is computed, \( c(n) \). This measure reflects the rate of new patterns arising with the increasing of the sequence length. The normalized complexity measure LZ is the current measure divided by its asymptotic trend.

\[
LZ = \frac{c(k)}{d(k)} = \lim_{k \to \infty} c(k) \frac{k}{\log_2 k}
\]  

where \( c(k) \) is the observed time series, \( \Phi = \max (0, r, s) \), \( \theta = \min (0, r, s) \) and \( n \) is the time series length. The proposed test [18] compares the estimated third order moment \( \hat{\mu}(r,s) \) of the series with a set of surrogates. Surrogates time series are constructed to have the same linear properties (as power spectrum) of the original time series but no other (nonlinear) structure. So, the comparison of high–order statistical measures performed on the original time series and the surrogates is significantly different and may be detected by a statistical test.

The proposed test compares \( \hat{\mu}(r,s) \) with a set of limits generated from linear stationary phase scrambled bootstrap data: large differences indicate nonlinearity. We collected, for each voxel, the C-value, i.e. 1 minus the \( p \)-value out of the test, and used it as a quantification of the rejection rate of the null hypothesis (i.e. the series is linear). For each of the two groups of subjects, a mean C-value map was computed and both maps were divided in 116 non-overlapping anatomical regions of interest (ROI) using the Automated Anatomical Labeling (AAL) atlas [19]. For each area a Significance Index (SI) is obtained by dividing the number of nonlinear voxels of each ROI by the total number of voxels in that ROI. This index is used to exclude, from the successive nonlinear analysis, the ROIs for which a too low (relative) number of voxel is nonlinearly activated.

C. Cross recurrence plots and CRQA

Of the 116 ROIs, those with the lowest SI were discarded. For each of the 105 survived ROI, a single time series was obtained by averaging the fMRI time courses over all the voxels within the ROI. The Takens theorem was used to reconstruct the trajectories in phase space of the dynamic system represented by each time series. The embedding dimension \( m = 6 \) (estimated with the False Nearest Neighbors algorithm [21]) and time delay \( \tau = 2 \) (estimated with First Local Minimum of Average Mutual Information algorithm [22]) were found for all the time series. All pairwise combinations of the time series were compared using cross recurrence plots:

\[
CR_{i,j}(\epsilon) = \theta(\epsilon - \| \hat{x}_i - \hat{y}_j \|) \quad i = 1, ..., M \quad j = 1, ..., N
\]  

with \( \hat{x}_i \) \( (i = 1, ..., M) \) and \( \hat{y}_j \) \( (j = 1, ..., N) \) being the trajectories of the two distinct systems, \( \theta \) the Heaviside function, \( \| - \| \) a norm and \( \epsilon \) the degree of neighborhood of the states, set as a few percent of the maximum phase state diameter. A cross recurrence plot is a \( M \times N \) matrix whose entries include information on the degree of closeness of each point of the first trajectory with each point of the second trajectory [14].

Even if a visual inspection of CRPs can reveal valuable information about the relationship between the two systems, the quantification of the structures in CRPs through CRQA has been used for making more objective the comparison between the time series.

\[
\hat{\mu}(r,s) = \frac{1}{n} \sum_{t=1}^{n-\Phi} X_t X_{t+r} X_{t+s}
\]  

where \( X_t \) is the observed time series, \( \Phi = \max (0, r, s) \), \( \theta = \min (0, r, s) \) and \( n \) is the time series length. The proposed test [18] compares the estimated third order moment \( \hat{\mu}(r,s) \) of the series with a set of surrogates. Surrogates time series are constructed to have the same linear properties (as power spectrum) of the original time series but no other (nonlinear) structure. So, the comparison of high–order statistical measures performed on the original time series and the surrogates is significantly different and may be detected by a statistical test. The proposed test compares \( \hat{\mu}(r,s) \) with a set of limits generated from linear stationary phase scrambled bootstrap data: large differences indicate nonlinearity. We collected, for each voxel, the C-value, i.e. 1 minus the \( p \)-value out of the test, and used it as a quantification of the rejection rate of the null hypothesis (i.e. the series is linear). For each of the two groups of subjects, a mean C-value map was computed and both maps were divided in 116 non-overlapping anatomical regions of interest (ROI) using the Automated Anatomical Labeling (AAL) atlas [19]. For each area a Significance Index (SI) is obtained by dividing the number of nonlinear voxels of each ROI by the total number of voxels in that ROI. This index is used to exclude, from the successive nonlinear analysis, the ROIs for which a too low (relative) number of voxel is nonlinearly activated.
Among all the available measures we focused our attention on:

- Measures based on recurrence density. The simplest one is the Recurrence Rate (RR) and is a measure of the density of the recurrence points in the CRP. It reveals the probability of occurrence of similar states in both systems.
- Measures based on diagonal lines. The lines which are diagonally oriented represent segments on both trajectories which run parallel for some time. We used the Determinism (DET) that is the proportion of recurrence points forming long diagonal structures with respect to all the recurrence points.
- Measures based on vertical lines. They refer to the time intervals in which the system remains in a specific state. We focused on the maximal length of vertical lines ($V_{max}$).

**D. Networks construction and graph analysis**

For each subject three adjacency matrices were built, whose entries $(i, j)$ represent the values of the CRQA measures for the pair $(i, j)$ of ROIs. To exclude self-loops, the main diagonals were reset resulting in weighted undirected graphs.

Thus, we explored the local and global properties of each graph through some complex networks measures. In particular we focused on:

- The Strength of a node that is defined as the sum of all the neighboring link weights. Nodes with high strength are essential for maintaining global connectivity.
- The Clustering coefficient of a node, which measures the density of connections among the neighbors of the node. It is a measure of local segregation. The average of the clustering coefficients for each individual node is the clustering coefficient of the graph.
- The Characteristic path length that is the average shortest path length between all pairs of nodes in the network and is the most commonly used measure of functional integration.
- The Betweenness centrality, an index of the importance of a node in the network: it is the fraction of all the shortest paths in the network that pass through the node.

**III. RESULT AND DISCUSSION**

The measures extracted from each graph have been averaged among subjects belonging to the same group, allowing a comparison between controls and schizophrenic patients for each selected complex measure. In the following, the results on the relation between the nonlinear areas and the metrics of the graphs are shown.

Figure 1 shows the bar plot of the ROIs with the highest difference of SI between the two groups of subjects (for the sake of visualization, not all the ROIs SI have been plotted). Control subjects show a higher number of nonlinearly activated voxels in correspondence of the left olfactory cortex, right middle occipital area and right angular gyrus. The right middle occipital area is the site of the secondary visual cortex and its main functions include the sustained attention to color and shape, the processing of visual-spatial information and the response to emotion/attention in visual processing. The angular gyrus, instead, is the part of the brain associated with complex language functions; in particular, the right angular gyrus has been associated with spatio-visual attention toward salient features. For schizophrenic subjects, instead, we found a high number of nonlinearly activated voxels in all the orbital frontal lobe, involved in sensory integration, in decision-making and expectation, and in the right parahippocampal gyrus, whose main function involves memory creation and recall of visual scenes. Functional abnormalities in this area have been already related to schizophrenia [20]. Other significant areas with different number of activated voxels in patients are the left pallidum, a structure involved in the regulation of voluntary movement, the thalamus that connects different subcortical areas and the putamen, that controls the motor learning, the motor performance and task. Significantly, the putamen is also devoted to regulate the dopamine, a neurotransmitter whose altered level is hypothesized to be one of the main causes of schizophrenia.

The meaning of the cross recurrence quantitative measures are exemplified in Figure 2. Figure 2a shows the CRP of similar time series. Here we can notice diagonal lines longer than those present in Figure 2b, which refers instead to a CRP of two time series belonging to very different anatomical areas; moreover, in the latter there is an almost complete absence of vertical lines.

![Figure 2](image-url)
Figure 1. Comparison of averaged Significance Index for schizophrenic (SZ) and control subjects (NC) for the most different 28 ROIs.

Figure 3: The graph measures for the three CRQA described in Section II-C. (a)-(c): the graph strength for the recurrence rate, the determinism, the maximal length of vertical lines; (d)-(f): the clustering coefficient for the recurrence rate, the determinism, the
maximal length of vertical lines; (g)-(i) the betweenness for the recurrence rate, the determinism, the maximal length of vertical lines. For all the plots a significative number of ROIs is represented grouped for patients and controls.

Figure 3 shows the average values of strength, clustering coefficient and betweenness in controls and patients, obtained from recurrence rate, determinism and maximal length of vertical lines relative to the areas with highest differences between the two groups of subjects. In all the plots, the most significative areas are reported.

The networks obtained from with recurrence rate and determinism measures show the highest differences on almost the same areas for the value of strength between the two groups of subjects (mainly orbital frontal lobe, hippocampus). The networks obtained from the maximal length of vertical lines measure exhibit the highest difference in other areas such as the temporal and occipital lobe. Figure 3c reveals the highest difference of strength between controls and patients with respect to 3a and 3b.

Similar considerations may be done for clustering coefficients: recurrence rate and determinism graphs show higher values of the clustering coefficient at the orbital frontal lobe and left amygdala in the group of control subjects compared to schizophrenics; instead, the maximal length of vertical lines graphs exhibits higher values of the clustering coefficient in the frontal orbital lobe for the group of patients.

The values of betweenness give interesting information about the centrality and importance of a node. A node with a high value of betweenness centrality may be seen as a "controller" of the information flow traveling over a network because it is located at the intersection of many short paths. In particular, the values of betweennes in graphs obtained from recurrence rate and determinism show the amygdala, the orbital frontal lobe and the hippocampus in the control subjects as "hubs" of the network while the inferior occipital lobe is the most important node in subjects with schizophrenia.

Table I shows the average values of the clustering coefficients and characteristic path length ($\lambda$) for the two groups of subjects. As it can be seen, $\lambda$ is always comparable between healthy and schizophrenic patients, regardless of the adopted complexity measure. This means that the graphs that describe the brain connectivity have a characteristic path length that is slightly variable between the two groups of subjects and therefore the average number of links that separates a ROI from another one in the network is similar in the two groups. On the other hand, the average value of the clustering coefficient found in the control group is always higher than that exhibited by schizophrenics in all the measures considered; in addition schizophrenic patients show less variability around the mean value of the clustering coefficient.

<table>
<thead>
<tr>
<th></th>
<th>Recurrence rate</th>
<th>Determinism</th>
<th>Length of the longest vertical line</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NC</td>
<td>SZ</td>
<td>NC</td>
</tr>
<tr>
<td><strong>Cluster</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>0.10</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.01)</td>
<td>(0.14)</td>
</tr>
<tr>
<td><strong>Lambda</strong></td>
<td>9.50</td>
<td>9.99</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>(2.91)</td>
<td>(1.50)</td>
<td>(1.52)</td>
</tr>
</tbody>
</table>

Table I. Clustering coefficient and characteristic path length average value and standard deviation in brackets for patients and controls divided for RR, DET, Vmax.

IV. CONCLUSIONS

In this paper we applied a composite method of analysis of fMRI data based on complex networks and concepts derived from the theory of nonlinear dynamic systems. A statistical nonlinearity test and an anatomical parcellization scheme were applied to choose the most nonlinear functional areas. From the time series representing each area, trajectories of dynamic systems interacting in brain networks were computed. The CRP and the CRQA were used to obtain the indices of similarity of these systems. These indices have been employed in the construction of graphs of functional connectivity of the whole brain volume for schizophrenics and healthy subjects. The comparison of some complex networks measures extracted from graphs shows significant differences between the areas of activation in the two groups and their characterization in terms of connection nodes in the functional integration process.

As successive step in this research, we will try to find a systematic scheme to identify the complex measures that may become effective for the classification of diseases, disorders, mental or emotional states and will seek to optimize their use in the clinical setting.

REFERENCES


