Low-dimensional embedding of functional connectivity graphs for brain state decoding

Jonas Richiardi*†, Dimitri Van De Ville*†,Hamdi Eryilmaz‡
*Medical Image Processing Laboratory, Ecole Polytechnique Fédérale de Lausanne (EPFL)
†Medical Image Processing Laboratory, University of Geneva
‡Laboratory of Neurology and Imaging of Cognition, University of Geneva
{jonas.richiardi,dimitri.vandeville}@epfl.ch,hamdi.eryilmaz@unige.ch

***Abstract*—Functional connectivity graphs are fully defined by their weighted adjacency matrix. Beyond the computation of graph-theoretical measures, we propose to use these graphs for inter-subject classification. Since they form a class of graphs with undirected edges and fixed number and ordering of vertices, vector space graph embedding techniques can be used to provide good classification performance. We propose a method that represents connectivity graphs in low-dimensional spaces, and we show experimental results hinting that such low-dimensional projections are beneficial for classification performance.**

***Keywords*-fMRI; graph embedding; two-dimensional SVD; brain decoding; resting-state**

I. Introduction

In functional neuroimaging, brain decoding uses multivariate pattern recognition methods to classify brain images; the goal is to infer a discrete set of experimental conditions from an observed brain state, or to assign subjects to classes defined by a medical diagnosis. This is done either within a single subject, or between different subjects. While decoding is typically performed using activation patterns (maps) in regions of interest or throughout the brain, an interesting alternative is to perform brain decoding by using connectivity measures.

Functional connectivity between brain regions or voxels is classically represented by a "connectivity matrix", which in graph theoretical terms is the adjacency matrix of an unweighted, undirected graph [1], [2]. The elements of the matrix are in many cases obtained by computing temporal correlations between region-averaged or voxel time courses. Next, setting an (arbitrary) correlation threshold yields a binary matrix enabling the computation of properties such as hubbiness, mean path length, and others. However, graph-theoretical measures are largely dependent on the particulars of threshold settings. For example, in structural connectivity, it has been shown that these topological properties are very sensitive to the scale of parcellation, and to a smaller but significant extent to the algorithms used to compute anatomical connectivity through tractography [3].

It is therefore of interest to go beyond topological descriptions of connectivity graphs and obtain actionable models of connectivity. In previous work, we have demonstrated the feasibility of inter-subject brain state decoding on a simple cognitive task (resting versus movie watching) by using only functional connectivity graphs [4]. We used a vector-space graph embedding technique, enabling access to the vast repertoire of vector-space pattern recognition techniques. For pattern recognition applications, thresholding a connectivity matrix is equivalent to selecting feature subsets by values of the Fisher ratio, and is often sub-optimal with respect to the classifier being used later on. Foregoing this statistical thresholding in turn leads to a high-dimensional classification task. In the present paper, we extend our embedding technique by introducing a lower-dimensional representation of threshold-free functional connectivity graphs, and evaluate classification performance on an inter-subject brain state decoding task.

II. Direct connection label sequence embedding of functional connectivity graphs

Functional connectivity graphs can be formalised as complete, undirected (weighted) graphs with a fixed number of vertices M and a fixed vertex ordering. This is a family of graphs subsumed by the class of graphs with unique node labels [5]. This observation is important to brain decoding as it conditions the choice of graph matching algorithm: many traditional techniques are unlikely to produce good results for such graphs [6]. Therefore, vector-space embedding of graphs via transformation of their adjacency matrix offers very promising alternatives. In direct connection label sequence embedding (direct embedding or DE for short), we simply transform the upper-triangular part of the adjacency matrix into a vector. This is a lossless transformation, but generates high-dimensional spaces and high-dimensional learning problems (more dimensions than training samples[1]), as the vector generated thus lives in $\mathbb{R}^{\binom{M}{2}}$.

[1]The *curse of dimensionality* is endemic in fMRI neuroimaging, and is one of the factors behind the popularity of SVM classifiers in the field.

978-0-7695-4133-4/10 $26.00 © 2010 IEEE
DOI 10.1109/WBD.2010.15

III. Reducing embedding dimension with a two-dimensional SVD

Our goal, then, is to obtain a more compact representation of functional connectivity graphs, which will facilitate learning and inference.

Here, we propose to use a two-dimensional SVD, inspired largely by the work on 2DPCA in face recognition [7]. The goal is to overcome the limitations of the common method for spectral embedding of graphs, where eigenspace decomposition is typically performed on the covariance matrix of training vectors made from adjacency (or Laplacian) matrices [8]. In neuroimaging, where very little training data is available, the covariance matrix is susceptible to being very poorly estimated, yielding large reconstruction error and poor decomposition.

In related works, we note that [9] have used an SVD and other algorithms to obtain low-dimensional representations of spectral coherence matrices.

A. Learning the projection axes

The first step to make inter-subject decoding possible is to divide the subjects into two mutually exclusive subsets: the *projection training subset*, comprising N_p subjects, is used to learn the projection axis that will be used to reduce the dimensionality of adjacency matrices, and the *low-dimensional subset*, consisting of N_l subjects and C conditions whose adjacency matrices will be projected to a low-dimensional subspace, and in which brain decoding will be performed. Another possibility would be to learn the transform on all training data in a cross-validation fold.

We start by computing the average adjacency matrix across conditions and subjects from the projection training subset:

$$\overline{\mathbf{A}} = \frac{1}{N_p \times C - 1} \sum_{c=1}^{C} \sum_{n=1}^{N_p} \mathbf{A}_{n,c} + \mathbf{I}. \tag{1}$$

The principle behind the use of a linear combination is to obtain a matrix that represents some amount of variability due to various subjects and conditions.

We then compute a singular value decomposition of the average adjacency matrix. Using standard notation: $\overline{\mathbf{A}} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, where U and V are orthogonal matrices of singular vectors, and S is a diagonal matrix of singular values.

Since $\mathbf{A}_{n,c}$ are computed from a relatively large number of samples (volumes) and no data is missing, we can safely assume that they are symmetric (and positive semidefinite (SPD) when pre-conditioned with $\mathbf{I}$), entailing that $\overline{\mathbf{A}}$ is also SPD. Therefore, the decomposition simplifies to

$$\overline{\mathbf{A}} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T, \tag{2}$$

where $\mathbf{U} = \mathbf{V} = \mathbf{Q}$, and $\mathbf{S} = \mathbf{\Lambda}$. The columns of $\mathbf{Q}$ constitute the projection axes. We note again that this training procedure is performed on the held-out projection training subset.

B. Low-dimensional projections

Once the projection axes are known, we transform each of the $N_l \times C$ adjacency matrices in the low-dimensional subset by projecting it down to a low-dimensional subspace:

$$\mathbf{A}^{\downarrow}_{n,c} = \mathbf{Q}^T_{1:K}(\mathbf{A}_{n,c} + \mathbf{I})\mathbf{Q}_{1:K}, \tag{3}$$

where K is the number of principal components to use in the projection.

This reduced-dimension representation concentrates discriminatory power in the components corresponding to the largest singular values, and can offer interesting insight into brain states by expressing their connectivity patterns as a combination of several distributed functional networks. Beyond visualisation and interpretation, it can be used as a long-vector to train any vector-space classifier.

IV. Experiments

A. Dataset

$N = 15$ healthy young subjects (4 m, 11 f) contributed data. Scanning was performed on a Siemens 3T Tim Trio. Functional imaging data were acquired using a GR-EPI sequence (TR/TE/FA = 1.1s/27ms/90°, matrix = 64×64, voxel size = 3.75×3.75×4.2mm^3, 21 contiguous transverse slices, 1.05mm gap, 2598 volumes). Structural imaging data was acquired using a 3D MPRAGE sequence (192 slices, TR/TE/FA = 1.9s/2.32ms/9°, matrix = 256 × 256, voxel size = 0.90 × 0.90 × 0.90mm^3).

Stimulation was done in a block design of 9 alternating epochs of movie excerpts (50s) and resting periods (90s), for a total duration of 23 minutes. During rest, subjects were instructed to close their eyes, relax, and let their mind wander. The auditory stimulation was provided through MRI-compatible headphones. The 9 blocks for each of the $C = 2$ conditions are concatenated after linear detrending.

B. Data processing and functional connectivity computations

Standard preprocessing (realignment, coregistration, normalisation) is performed on the data using all default settings of SPM 5 (http://www.fil.ion.ucl.ac.uk/spm/). The structural data is atlased using the AAL template into $M = 90$ regions using the IBASPM toolbox [10], and subsequently warped to an individual functional atlas. Note that each of the N subjects is processed in this way independently of the others, to guarantee theoretically correct inter-subject decoding with no train/test separation issues. The timecourses within a region are averaged, yielding 90 time-series/subject/condition. Each time-series is filtered into four wavelet subbands using a non-subsampled orthogonal discrete wavelet transform (4 vanishing moments). Because of the stable properties of resting-state activity (consistently found across populations, genders, and imaging centers), we retain the 0.06-0.11 Hz

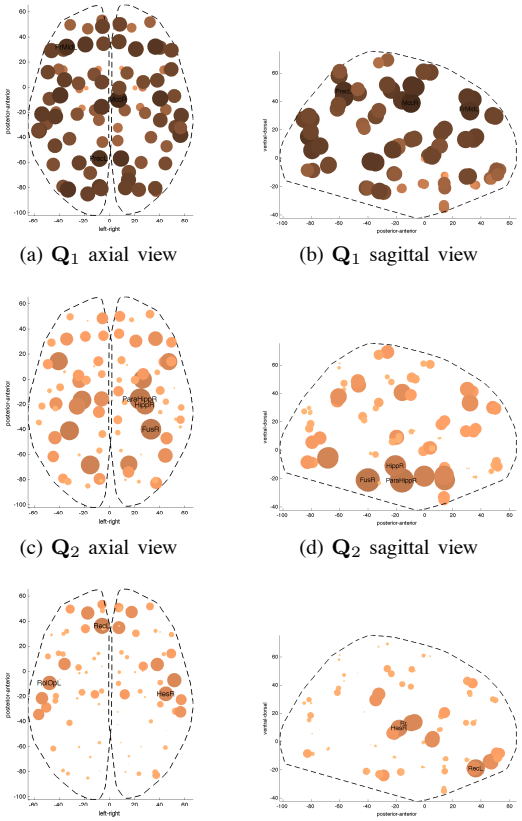
subband, yielding a matrix of filtered time courses $\mathbf{X}_{n,c} : \mathbb{R}^{M \times T}$.

For each $\mathbf{X}_{n,c}$, we compute all $\binom{90}{2}$ pairwise Pearson linear product-moment correlation coefficients between all regions of interest, yielding the correlation matrices $\mathbf{R}_{n,c} : \mathbb{R}^{M \times M}$.

We now consider the brain regions as a set of vertices V and the correlation coefficients as signed weights on the set of edges E, leading to an undirected complete weighted graph $\mathcal{G} = (V, E)$. The weighted graph adjacency matrices $\mathbf{A}_{n,c}$ can be defined as $\mathbf{A}_{n,c} = \mathbf{R}_{n,c} - \mathbf{I}$.

C. Principal components of functional connectivity

Figures 1 and 2 show a MNI space representation of the first few principal components corresponding to the largest singular values of the two-dimensional decomposition, computed on 5 arbitrary subjects, as well as the adjacency matrix extracted from the principal components by setting the matrix diagonal to 0 (to prevent self-loops in the graph). Since the graph is complete, the degree of a vertex is not informative, and the MNI space representation instead shows vertex size proportional to the sum of edge weights for all vertices.



(a) $\mathbf{Q}_1$ axial view (b) $\mathbf{Q}_1$ sagittal view

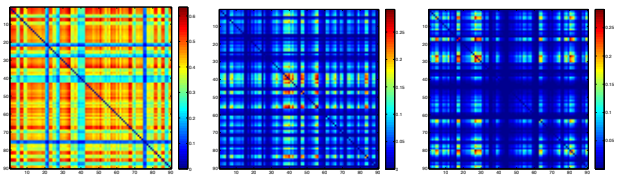
(c) $\mathbf{Q}_2$ axial view (d) $\mathbf{Q}_2$ sagittal view

(e) $\mathbf{Q}_3$ axial view (f) $\mathbf{Q}_3$ sagittal view

Figure 1. MNI space view of first 3 principal component of the two-dimensional SVD of $\overline{\mathbf{A}}$. Vertices with larger, darker spheres correspond to regions that have larger sums of edge weights, and are therefore more correlated with other regions. Scale is different for first component.

The first principal component has the highest correlations of all components, and connectivity is very spatially distributed throughout the whole brain, with slightly lower connectivity in ventral areas. The second principal component captures connectivity within the limbic areas (especially hippocampal and parahippocampal formations) and between limbic regions and ventral areas. The third principal component represents mostly connectivity within and between temporal areas and frontal ares, with notable bilateral symmetry.

Other components are omitted due to space constraints. Results differ depending on subjects included in the projection training subset, because of its the small size.



(a) $\mathbf{Q}_1$ (b) $\mathbf{Q}_2$ (c) $\mathbf{Q}_3$

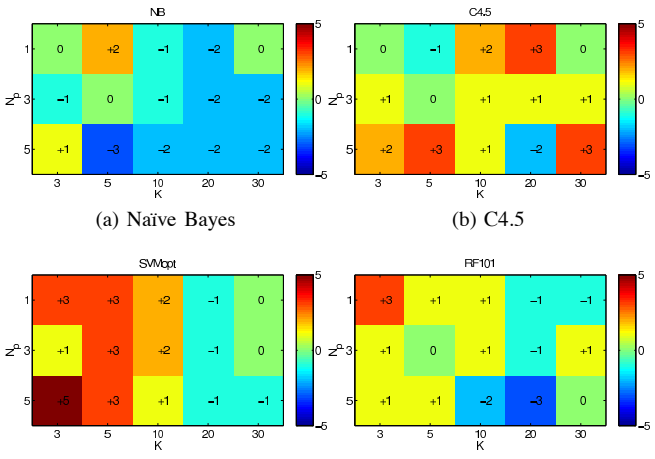
Figure 2. Adjacency matrices of first 3 principal component of the two-dimensional SVD of $\overline{\mathbf{A}}$. Note different colour scales.

D. Classification

The success of two-dimensional SVD projection for pattern recognition depends on two parameters: the number of subjects N_p used in training the decomposition, and the number of principal components used for the projection K. To ascertain the impact of these factors, we performed classification experiments using 1 generative classifier (Naïve Bayes with Gaussian distributions), 1 decision tree classifier (C4.5), 1 kernel-based classifier (SVM with linear kernel, cost parameter optimised in-fold), and 1 ensemble classifier (Random forest with 101 trees). For each $K \in \{3, 5, 10, 20, 30\}$ and $N_p \in \{1, 3, 5\}$, we performed 5 times the random splitting of the N subjects into the N_p subjects for projection training and $N_l = 10$ subjects for low-dimensional classification experiments, and report results averaged over the 5 repetitions. In every case, the direct connection label sequence embedding (DE) into $\mathbb{R}^{4005}$ is compared with the two-dimensional SVD embedding (2DSVD).

Figure 3 shows the differences in terms of number of classification errors between DE embedding and 2DSVD embedding: negative differences mean 2SVD outperforms DE. For all classifiers, it is possible to find a setting of embedding parameters (N_p, K) where the low-dimensional embedding improves on the direct embedding. The clear pattern (except for C4.5) is that $N_p = 5$ is the best choice, which can be expected given the large inter-subject variability typical in neuroimaging. The choice of the number of principal components K indicates that about 20 components are

sufficient to obtain an improvement over direct embedding. Note that a difference of 1 error corresponds to a $1/20 = 5\%$ absolute difference in classification performance. Given the very small size of the dataset, however, it is not possible to ascertain whether these improvements are statistically significant.



(a) Naïve Bayes (b) C4.5

(c) SVM (d) Random Forest

Figure 3. Difference in number of errors between direct embedding (DE) and two-dimensional SVD embedding (2DSVD) for 4 classifiers and different projection parameters. Negative numbers mean 2DSVD causes fewer errors than DE, positive numbers the opposite. K is the number of principal components used in the projection. N_p is the number of subjects held-out for the projection learning subset.

Table I shows classification accuracies (for each class) for the direct embedding and the 2DSVD embedding with best-case settings of (N_p, K) chosen post-hoc. Assuming a sufficient number of subjects is used for learning the projection, two-dimensional singular value decomposition seems a promising approach in terms of classification performance. Interestingly, it seems to enable the use of low-complexity, fast classifiers such as Naïve Bayes, and seems to also improve performance for a classifier such as Random Forest, which performs embedded feature selection. Caution is still warranted because of the small size of the dataset. The abysmal performance of the C4.5 classifier seems to indicate that discriminatory power is distributed amongs features, rather than sharply peaked.

Table I
CLASSIFICATION ACCURACY (CLASSES I:RESTING AND II:MOVIES) FOR DIFFERENT CLASSIFIERS AND EMBEDDING TYPES.

Classifier	Embedding	acc_I	acc_{II}
NB	DE	100%	60%
	2DSVD (N_p=5,K=20)	100 %	80%
C4.5	DE	90%	30%
	2DSVD (N_p=5,K=20)	100 %	40%
SVM	DE	100%	80%
	2DSVD (N_p=5,K=20)	100 %	90%
RF	DE	100%	60%
	2DSVD (N_p=5,K=20)	100 %	90%

V. CONCLUSION

We have proposed to apply a dimensionality reduction technique to adjacency matrices of connectivity graphs in a vector-space embedding setting for graph matching, and shown that improved classification performance is likely attainable. Future work will concentrate on applying the method to larger datasets for validation, and on other dimensionality reduction techniques that are applicable out-of-sample.

ACKNOWLEDGMENTS

We thank S. Schwartz and P. Vuilleumier for the fMRI data. This work was supported in part by the Swiss National Science Foundation (grants PP00P2-123438), in part by the Société Académique de Genève and the FOREMANE foundation, and in part by the Center for Biomedical Imaging (CIBM) of the Geneva and Lausanne Universities, EPFL, and the Leenaards and Louis-Jeantet foundations.

REFERENCES

[1] R. Salvador, J. Suckling, C. Schwarzbauer, and E. Bullmore, "Undirected graphs of frequency-dependent functional connectivity in whole brain networks," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 937–946, 2005.

[2] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *The Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, Jan. 2006.

[3] A. Zalesky, A. Fornito, I. H. Harding, L. Cocchi, M. Yücel, C. Pantelis, and E. T. Bullmore, "Whole-brain anatomical networks: does the choice of nodes matter?" *NeuroImage*, vol. 50, pp. 970–983, 2010.

[4] J. Richiardi, H. Eryilmaz, S. Schwartz, P. Vuilleumier, and D. Van De Ville, "Decoding brain states from fMRI connectivity graphs," *NeuroImage*, 2010, (in press).

[5] P. Dickinson, H. Bunke, A. Dadej, and M. Kraetzl, "Matching graphs with unique node labels," *Pattern Analysis & Applications*, vol. 7, no. 3, pp. 243–254, Sep. 2004.

[6] J. Richiardi, D. Van De Ville, K. Riesen, and H. Bunke, "Vector space embedding of undirected graphs with fixed-cardinality vertex sequences for classification," in *Proc. 20th Int. Conf. on Pattern Recognition (ICPR)*, 2010.

[7] J. Yang, D. Zhang, A. Frangi, and J. yu Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *EEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[8] R. Wilson, E. Hancock, and B. Luo, "Pattern vectors from algebraic graph theory," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 7, pp. 1112–1124, 2005.

[9] B. Thirion, S. Dodel, and J.-B. Poline, "Detection of signal synchronizations in resting-state fMRI datasets," *NeuroImage*, vol. 29, pp. 321—327, 2006.

[10] Y. Alemán-Gómez, L. Melie-García, and P. Valdés-Hernandez, "IBASPM: Toolbox for automatic parcellation of brain structures," in *Proc. 12th Annual Meeting of the Organization for Human Brain Mapping*, Florence, Italy, June 2006.