

Task-Specific Functional Brain Geometry from Model Maps

Georg Langs^{1,2}, Dimitris Samaras^{1,3}, Nikos Paragios^{1,2}, Jean Honorio³,
Nelly Alia-Klein⁴, Dardo Tomasi⁴, Nora D. Volkow⁴, and Rita Z. Goldstein⁴

¹ Laboratoire de Mathématiques Appliquées aux Systèmes,
Ecole Centrale de Paris, France

`georg.langs@ecp.fr`, `nikos.paragios@ecp.fr`*

² Equipe GALEN, INRIA Saclay, Île-de-France, France

³ Image Analysis Laboratory, Stony Brook University, USA
`samaras@cs.sunysb.edu`, `jhonorio@cs.sunysb.edu`

⁴ Medical Department, Brookhaven National Laboratory, USA
`rgoldstein@bnl.gov`

⁵ National Institute on Drug Abuse - National Institutes of Health, USA

Abstract. In this paper we propose *model maps* to derive and represent the intrinsic functional geometry of a brain from functional magnetic resonance imaging (fMRI) data for a specific task. Model maps represent the coherence of behavior of individual fMRI-measurements for a set of observations, or a time sequence. The maps establish a relation between individual positions in the brain by encoding the blood oxygen level dependent (BOLD) signal over a time period in a Markov chain. They represent this relation by mapping spatial positions to a new metric space, the *model map*. In this map the Euclidean distance between two points relates to the joint modeling behavior of their signals and thus the co-dependencies of the corresponding signals. The map reflects the functional as opposed to the anatomical geometry of the brain. It provides a quantitative tool to explore and study global and local patterns of resource allocation in the brain. To demonstrate the merit of this representation, we report quantitative experimental results on 29 fMRI time sequences, each with sub-sequences corresponding to 4 different conditions for two groups of individuals. We demonstrate that drug abusers exhibit lower differentiation in brain interactivity between baseline and reward related tasks, which could not be quantified until now.

1 Introduction

Despite the tremendous progress in studying the brain, little is known about its function and flexibility for specific activities. One way to understand these

* This work has been partially supported by the Region Île-de-France; by NIDA 1 R01 DA020949-01; by grants from the National Institute on Drug Abuse (RZG: 1R01DA023579 and R21DA02062); Laboratory Directed Research and Development from U.S. Department of Energy (OBER; to RZG: LDRD #07-055), National Institute on Alcohol Abuse and Alcoholism (NDV: AA/ODO9481-04), and General Clinical Research Center (5-MO1-RR-10710).

processes is through a statistical analysis of functional brain imaging datasets. Neural activity can be captured by functional magnetic resonance imaging (fMRI). It takes advantage of the hemodynamics response caused by active nerve cells that consume oxygen. The resulting fMRI signal is different depending on the level of oxygenation, and can be detected by using blood-oxygen-level dependent (BOLD) contrast.

Recently the extraction of information from fMRI was approached by various methodologies: In [1] the correlation of the BOLD signal *between* individuals while watching a movie was utilized to search for regions relevant for memory. Inter-subject synchronization was also studied in [2]. In [3] the strength of fMRI signals and its change for certain activities, and preconditions was assessed, in order to differentiate between drug addicts and control subjects. In [4], and [5] fluctuations within the brain were connected with the variability in subsequent fMRI observations. Even-though these works refer to the interdependency of brain regions, a comprehensive method to capture and represent this functional brain structure, and its dynamics, is still missing.

In this paper we address the question of synchronization within the brain. That is, we search for brain regions that exhibit highly coherent behavior as a strong indication of cooperation during an activity. We expect the BOLD signals corresponding to these regions to be observations stemming from a single source - the cooperative work caused by a certain condition. We propose a method to capture and represent these relationships in a transparent manner. It allows for data exploration, and for quantitative measurements of relationships between different regions of the brain, which are dynamic and task-specific. We call the set of these relationships the *functional geometry*. In addition to knowledge about functional regions [6] it offers information about their subtle mutual interaction patterns. The distance between the signals is derived from the joint model description length [7] of groups of signals, describing their joint modeling behavior [8]. In contrast to standard correlation, it can capture more complex or hidden relations, that go beyond mere synchronization. The method can be applied to other modalities like EEG/ERP or PET.

We use diffusion maps [9] to retrieve and encode the functional geometry. A diffusion map is a space constructed by the eigenfunctions of a Markov matrix. The Euclidean distance within this space offers a geometry that can capture complex relations between nodes in a Markov chain, and is related to spectral segmentation approaches [10]. In [11] diffusion maps were used to perform dimensionality reduction by parameterizing entire brain states, to represent relations between brains. In [12] they were used to segment activated regions. A related line of work addresses correlation analysis with regard to seed points, [13], while in [14] the brain is partitioned, and representative BOLD signals are searched for by a correlation and expectation maximization approach. The method presented in this paper tackles a different question: instead of differentiating between activated and non-activated voxels in the fMRI volume, or between subjects, we ask for a continuous functional relation between brain regions. We do not classify brain regions, but quantify their - often subtle - relations continuously.

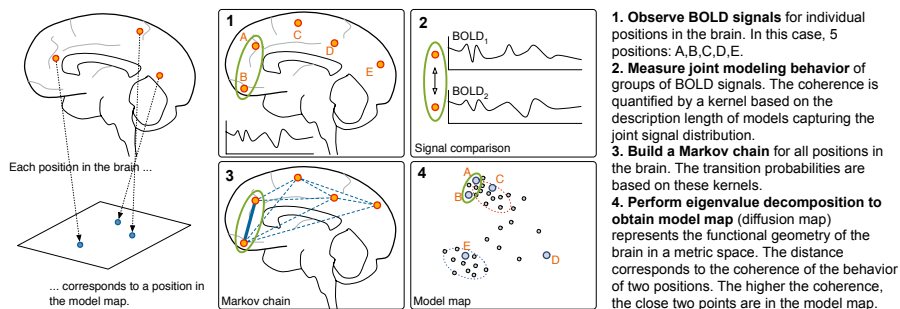


Fig. 1. Generating a model map generation from fMRI data

We propose to use *model maps* to explore the functional geometry of a brain for a certain task or time period. Each spatial position in the brain is mapped to a position in the map that is governed by the functional coherence of the corresponding observed BOLD signals in the brain (Fig. 1). The model map is built by calculating a Markov chain with nodes representing the positions in the brain, and transition probabilities defined by the description lengths [7] of models, that encode the joint density of the signals. The resulting model map captures joint modeling behavior of signals acquired at different positions, and reflects this functional geometry. It has several interesting properties: functional relations are translated to Euclidean distance, therefore groups of voxels, that have a high probability to stem from the same model, form clusters in the map. The density for positions in the map provides information about how *connected* a point is to any other region in the brain. High density indicates high coherence with many other signals, while low density indicates relatively independent behavior. These properties are essential for data exploration of complex fMRI sequences. Instead of a parcellation of the brain, they present the entire functional geometry including subtle dependencies. The unique position of points in the map makes a comparison between subjects, and between time-points for the same subject possible. These properties have considerable diagnostic value (as reported in the experiments), and we believe that they are an important tool, to explore and assess the changing distribution patterns of individual brain activities, that are not captured by the BOLD signal strength at individual positions. We evaluate the method on a challenging data set, that exhibits subtle cognitive changes regarding reward processing. Experiments show that the method is able to capture subtle differences, and interactions for different tasks.

2 Model Maps to Find Geometry in Functional Brain Data

We aim at a representation that maps measurements to positions in a space, so that low distance between two points indicates high compactness for a model

that encodes both of them, or high temporal coherence of their signals. We derive model maps from a set of signals $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$, where $\mathbf{x}_i \in \mathbb{R}^n$. The signals are mapped to a new manifold $\{\Psi_i, \dots, \Psi_m\}$ in the model map, where the Euclidean distance relates to a similarity between signals. In our case, each \mathbf{x}_i is the BOLD signal observed at one position in the brain for n time points. In this section we will first discuss how to define a similarity function, that captures relations between BOLD signals based on a multivariate Gaussian model. Then we will describe how to construct a Markov chain, and the corresponding model map with new positions Ψ_i for each signal \mathbf{x}_i . Finally we will explain how to perform measurements in the model map. The approach is closely related to *shape maps*, introduced in [8].

2.1 Comparing Signals: Compactness of a Gaussian Model

To establish a relation between a group of signals $\{\mathbf{x}_1, \dots, \mathbf{x}_l\}$ we find the principal axes of the joint distribution, and approximate it by a multivariate Gaussian distribution with Gaussians along the principal axes. For each dimension of the eigenspace used to encode the data we can apply Shannons theorem [15] to the according one-dimensional distribution. That is we can estimate the description length or complexity of a model, that encodes the data. The description length comprises the cost L of communicating a model \mathcal{M} itself (the parameters of the Gaussians) and the data D (i.e. BOLD signals) encoded with the model: $L(D, \mathcal{M}) = L(\mathcal{M}) + L(D|\mathcal{M})$. An extensive derivation of the description length calculation for Gaussian models is given in [16]. The description length reflects the complexity of the representation, and thus the plausibility of data stemming from a certain model. The method can be applied to models other than Gaussian straight-forwardly, and an optimal choice is subject of current research.

2.2 A Markov Chain That Describes Dependencies

Given a set of n examples, each consisting of BOLD signal observations for m points, we derive a metric on the set of points, that reflects their joint modeling behavior. The construction of such a diffusion map is explained in detail in [9]. The Markov chain consists of m nodes X which correspond to the individual signals \mathbf{x}_i , and pair-wise relations $d(i, j)$. We define $d_k(i, j)$ for two signals and *kernel size* k based on the minimum description length of models encompassing the two signals i and j and $k - 2$ others: $d_k : \{1, 2, \dots, n\}^2 \rightarrow \mathbb{R}$: $d_k(i, j) = \min_{\mathcal{M}}(\mathcal{L}(\mathcal{M})|i, j \subseteq \mathcal{M} \text{ and } \#\mathcal{M} = k)$, where the model \mathcal{M} with cardinality $\#\mathcal{M} := \#\{h_1, \dots, h_k\} = k$ represents k signals. $\mathcal{L}(\mathcal{M})$ denotes the description length $\mathcal{L}(\mathbf{x}_{\{h_1, \dots, h_k\}})$ as defined in Sec. 2.1, and $i, j \subseteq \mathcal{M} \Leftrightarrow i, j \in \{h_1, \dots, h_k\}$ for \mathcal{M} . That is, $d(i, j)$ is the minimum of the description lengths of all models representing i and j , and arbitrary $k - 2$ other entries of the observations. Note that for $k > 2$, $d_k(i, j)$ not only depends on the behavior of the two signals i and j but on the behavior of a larger sub set that is most *affine* to the two, i.e., that can be described by the least complex model. With increasing k only larger coherent sub-sets will benefit in terms of the distance in the model map,

allowing for a multi-resolution approach. d_k is non-negative, and symmetric, and the nodes X and edges weighted according to d_k between the nodes build a symmetric graph (X, d_k) . In practice we can estimate d_k by randomly choosing sub-sets of the data, calculating the according value $d_k(i, j)$ for all pairs in the sub-set and keeping the minimum of all samples for (i, j) . Using this set of pairwise relations, one can proceed to form a Markov chain that encompasses the notion of compactness in the entire set of fMRI signals.

From this, the normalized graph Laplacian construction [17] generates a reversible Markov chain, resulting in a diffusion operator P and its powers P^t . P is the Markov matrix with the entries $p(i, j)$, which are the transition probabilities in the Markov chain. P^t allows to propagate information through the Markov chain. The probability of the transition from any point i to another point j in t timesteps is given by the according kernel $p_t(i, j)$. By increasing t we can analyze the data at multiple *scales* i.e., propagate the relations between pairs of nodes. This operator P defines the geometry on the set of signals we are looking for, which can be mapped to an Euclidean geometry by an eigenvalue decomposition of P . Corresponding to P^t we can define a family of *diffusion distances* parameterized by t on the set of fMRI signals:

$$D_t(i, j) = \sum_{l=1, \dots, m} \frac{(p_t(i, l) - p_t(j, l))^2}{\pi(l)} \quad \text{where} \quad \pi(i) = \frac{d(i)}{\sum_j d(j)} \quad (1)$$

is the probability of i in the unique stationary distribution (the uniqueness is fulfilled if the graph is connected). D_t is an L^2 distance and captures the connectivity in the Markov chain, summing over all possible paths from i to j . The distance D_t is low if there is a large number of paths of length t with high transition probabilities between the nodes i and j .

2.3 Constructing Model Maps of Functional Data

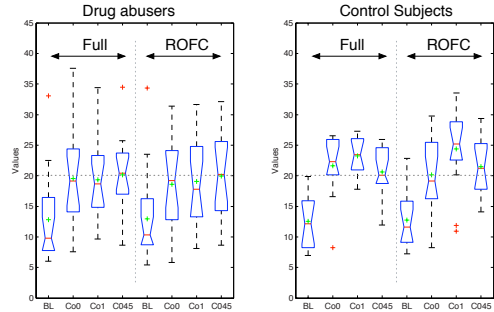
We now construct the space that encodes the functional relations of all fMRI signals from this chain. An eigenvalue decomposition of the operator P results in a sequence of eigen values $\lambda_1, \lambda_2 \dots$ and corresponding eigen functions Ψ_1, Ψ_2, \dots that fulfill $P\Psi_i = \lambda_i\Psi_i$. In [9] the authors explain how a *diffusion map*, the space spanned by the eigenfunctions of a Markov chain relates to the geometry determined by a diffusion distance D_t . We use this to create a metric space, a *model map* $\Psi_t : X \rightarrow \mathbb{R}^w$, that embeds the nodes (fMRI signals) $i = 1, \dots, m$ which are represented in the Markov chain into a Euclidean space where the diffusion distance in Eq. 1 corresponds to the euclidean distance in the eigen space:

$$\Psi_t(i) \triangleq \begin{pmatrix} \lambda_1^t \Psi_1(i) \\ \vdots \\ \lambda_w^t \Psi_w(i) \end{pmatrix}, \quad \text{and} \quad \|\Psi_t(i) - \Psi_t(j)\| = D_t(i, j). \quad (2)$$

Thereby the *functional* relations between fMRI signals are translated into spatial distances in the model map. The local density in the map corresponds to

Table 1. Relative increase of model map density from baseline. Left: entire brain, right: right orbitofrontal cortex. Right: corresponding plots for the absolute values.

	Full		ROFC	
	Abuser	Control	Abuser	Control
BL	0	0	0	0
C0	0.68	0.94	0.63	0.77
C1	0.70	1.04	0.68	1.14
C45	0.74	0.77	0.82	0.91



the amount of other signals, that are closely related. We can perform standard density estimation, clustering, or the definition of neighborhoods in this space. Closest neighbors indicate high coherence, and the scatter of specific brain regions in the map indicates the level of co-dependency of their sets of signals. Such a process captures the relations between the fMRI signals during the observation time.

3 Experiments

Data and set-up: We conducted experiments with previously acquired fMRI data from 29 individuals, 16 of which were cocaine abusers (group A), and 13 were control subjects (group B) provided by the authors of [3]. The MRI scans were acquired by a 4-T whole body MRI scanner (Varian), and BOLD responses were measured for 91 time points. The dataset focused on the way different subjects process monetary rewards. The overall neuropsychological experiment design included six blocks each consisting of three monetary reward conditions: 0 cent, 1 cent and 45 cents. This results in 3 baseline conditions (*BL*) without stimulus and the stimulus conditions (*Co0/1/45*). To assess if the model maps provide quantitative diagnostically relevant information, we assess their ability to differentiate between the subject groups, and between tasks. We measure the model map density for each point. We evaluate the mean density for the entire map, and the densities for the map positions corresponding to anatomical regions, e.g. the orbitofrontal cortex. We performed three comparisons: 1. We expect functionally connected regions within the brain to be in close proximity in the model map. Thus the density should increase if a specific task is performed, and we compared the model map density for *Co0*, *Co1*, and *Co45* with *BL1*. 2. For the drug abusers the relative increase of interaction between *BL1* and *Co0*, *Co1*, and *Co45* is expected to be lower, thus we compared the relative increase of density for the two subject groups. 3. We observed the model map scatter for the cortical region of *orbitofrontal cortex*, *lateral prefrontal cortex*, and *anterior cingulate cortex*, regions previously found to be important in the

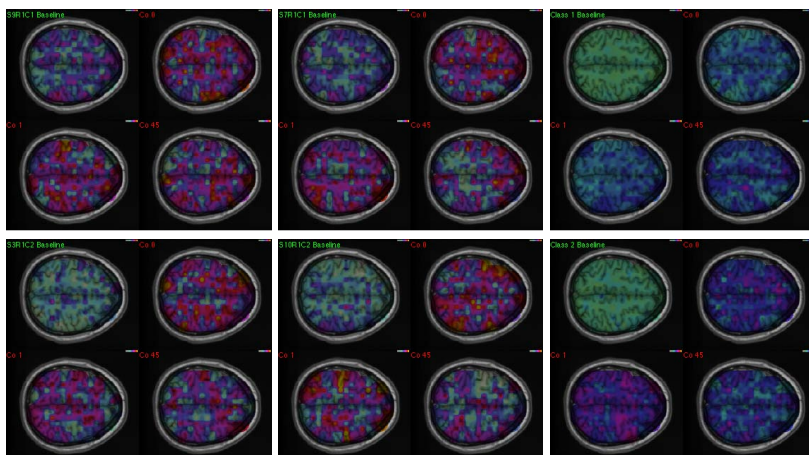


Fig. 2. Functional integration within the brain: the model map density values were projected back to the spatial coordinate system. Top row: two series for drug abusers, and median values, bottom row: two series for controls, and median values. For each subject baseline, and the three conditions are depicted. The differences between baseline (top left in each panel) and reward conditions are less pronounced for the drug abusers.

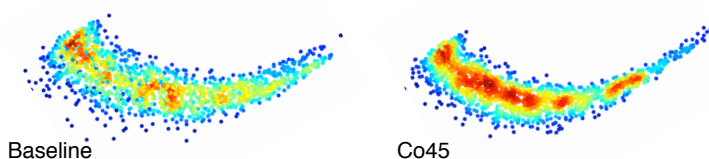


Fig. 3. Two model maps for one subject: one corresponding to baseline, one to Co45

context of reward processing [3]. Scatter indicates the level of integration of the activity in this region. We also report qualitative results for the model maps, to demonstrate the visual information they provide for data exploration, and visualize the model map density - the connectivity - in the spatial brain space.

Results: In Tab. 1 the increase in model map density relative to BL is reported, and plots for the absolute values for both drug abusers, and control subjects are depicted. For all subjects, the density values for Co0/1/45 are significantly higher than for BL ($p=0.05$). The relative increase is more pronounced for control subjects. An interesting observation is that the scatter of the points in the model map that correspond to the cortical regions is significantly lower for group A, compared to group B during BL ($p=0.034$). This indicates that these regions behave in a more coherent manner during baseline for drug abusers. In Fig. 2 the model map density is visualized on corresponding structural (T1) data. The top row shows two examples, and the median values for group A, while the bottom row shows the same for group B. The more pronounced increase of connectivity

for group B is visible in this example. The median images show a similar trend for the entire population. However, since these results exhibited a certain level of heterogeneity more research is needed to understand their implications. In Fig. 3 the model maps are shown for conditions BL, and Co45. The color encodes the local density in the map. The rather sparse structure with scattered local density maxima of the BL map changes to a dense map for Co45. This corresponds to increased interactions between brain regions.

4 Conclusion

We propose a method to capture, and to represent the functional geometry of brains in model maps. The representation is built by quantifying the joint modeling behavior of individual BOLD signals over time for all positions in the brain. The model map represents the functional relations between brain positions. In addition to mere signal strength, the distance in the map can indicate subtle co-dependencies between regions, while the density represents the functional connectivity to the rest of the brain. Experimental results demonstrate that the method extracts clinically relevant information, and can be used for data exploration, since in contrast to seed-point based approaches, it is built in an unsupervised manner, and encodes the structure of the entire brain in a single map. We expect this method to be relevant when searching for differentiating regions in the brain for the detection of certain changes over time. The functional connectivity can be used to establish functional trajectories. Promising directions both from the practical as well as theoretical point of view are to introduce dependencies that go beyond pair-wise interactions between the nodes and to relax the multi-variate Gaussian assumption. The proposed method could have an impact on understanding the relationships of measurements from heterogeneous domains (functional, anatomical), which is one of the challenges in the field of biological and biomedical analysis.

References

1. Hasson, U., Furman, O., Clark, D., Dudai, Y., Davachi, L.: Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* 57(3), 452–462 (2008)
2. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R.: Intersubject synchronization of cortical activity during natural vision. *Science* 303(5664), 1634–1640 (2004)
3. Goldstein, R.Z., Alia-Klein, N., Tomasi, D., Zhang, L., Cottone, L.A., Maloney, T., Telang, F., Caparelli, E.C., Chang, L., Ernst, T., Samaras, D., Squires, N.K., Volkow, N.D.: Is decreased prefrontal cortical sensitivity to monetary reward associated with impaired motivation and self-control in cocaine addiction? *Am. J. Psychiatry* 164(1), 43–51 (2007)
4. Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E.: Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* 56(1), 171–184 (2007)

5. Fox, M.D., Raichle, M.E.: Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8(9), 700–711 (2007)
6. Brett, M., Johnsrude, I.S., Owen, A.M.: The problem of functional localization in the human brain. *Nat. Rev. Neurosci.* 3(3), 243–249 (2002)
7. Rissanen, J.: Modeling by shortest data description. *Automatica* 14, 465–471 (1978)
8. Langs, G., Paragios, N.: Modeling the structure of multivariate manifolds: Shape maps. In: *Proc. of CVPR 2008* (2008)
9. Coifman, R.R., Lafon, S.: Diffusion maps. *Appl. Comp. Harm. An.* 21, 5–30 (2006)
10. Yan, J., Pollefeys, M.: A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3954, pp. 94–106. Springer, Heidelberg (2006)
11. Meyer, F.: Learning and predicting brain dynamics from fMRI: a spectral approach. In: *Wavelet XII conference, Proceedings of SPIE*, vol. 6701(2007)
12. Shen, X., Meyer, F.: Analysis of Event-Related fMRI Data Using Diffusion Maps. *Inf. Process Med. Imaging* 19, 652–663 (2005)
13. Wang, Y., Xia, J.: Functional interactivity in fMRI using multiple seeds' correlation analyses – novel methods and comparisons. In: Karssemeijer, N., Lelieveldt, B. (eds.) *IPMI 2007*. LNCS, vol. 4584, pp. 147–159. Springer, Heidelberg (2007)
14. Golland, P., Golland, Y., Malach, R.: Detection of spatial activation patterns as unsupervised segmentation of fMRI data. *Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Assist. Interv.* 10(Pt 1), 110–118 (2007)
15. Shannon, C.: A mathematical theory of communication. *Bell Systems Technical Journal* 27, 379–423 (1948)
16. Davies, R.H., Twining, C., Cootes, T.F., Waterton, J.C., Taylor, C.J.: A minimum description length approach to statistical shape modeling. *IEEE TMI* 21(5), 525–537 (2002)
17. Chung, F.R.: *Spectral Graph Theory*. American Mathematical Society (1997)