NeuroImage 123 (2015) 185–199

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Anatomically-adapted graph wavelets for improved group-level fMRI activation mapping

Hamid Behjat [a,*], Nora Leonardi [b], Leif Sörnmo [a], Dimitri Van De Ville [b,c]

[a] Biomedical Signal Processing Group, Department of Biomedical Engineering, Lund University, Lund, Sweden
[b] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[c] Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland

ARTICLE INFO

Article history:
Received 30 January 2015
Accepted 2 June 2015
Available online 7 June 2015

Keywords:
Statistical parametric mapping (SPM)
Functional MRI
Spectral graph theory
Graph wavelets
Wavelet thresholding

ABSTRACT

A graph based framework for fMRI brain activation mapping is presented. The approach exploits the spectral graph wavelet transform (SGWT) for the purpose of defining an advanced multi-resolutional spatial transformation for fMRI data. The framework extends wavelet based SPM (WSPM), which is an alternative to the conventional approach of statistical parametric mapping (SPM), and is developed specifically for group-level analysis. We present a novel procedure for constructing brain graphs, with subgraphs that separately encode the structural connectivity of the cerebral and cerebellar gray matter (GM), and address the inter-subject GM variability by the use of template GM representations. Graph wavelets tailored to the convoluted boundaries of GM are then constructed as a means to implement a GM-based spatial transformation on fMRI data. The proposed approach is evaluated using real as well as semi-synthetic multi-subject data. Compared to SPM and WSPM using classical wavelets, the proposed approach shows superior type-I error control. The results on real data suggest a higher detection sensitivity as well as the capability to capture subtle, connected patterns of brain activity.

Introduction

Functional magnetic resonance imaging (fMRI) is a key modality to localize brain activity based on the blood-oxygen-level-dependent (BOLD) signal (Ogawa et al., 1993). The most widely used approach in fMRI activation mapping is a mass univariate hypothesis-driven method that is implemented in many software packages such as statistical parametric mapping (SPM) (Frackowiak et al., 1997; Friston et al., 1994). Using regressors defined by the experimental paradigm, a general linear model (GLM) is fitted to the time course of every voxel of the brain, followed by a statistical test of a linear combination of the fitted parameters, leading to a statistical map indicating evidence for stimulus-related brain activity. Since using a Bonferroni correction is too conservative, SPM deals with the multiple comparison problem based on Gaussian random field theory (GRFT) (Poline et al., 1997). A key characteristic of GRFT is that it requires initial smoothing of the functional data by a fixed Gaussian filter. This pre-filtering not only is required to control the spatial smoothness of the data to comply with GRFT, but it also serves as a means to improving the signal-to-noise ratio (SNR) by virtue of the matched filter argument. However, such linear isotropic filtering comes at the expense of a loss in fine spatial details of the underlying activity.

* Corresponding author at: Department of Biomedical Engineering, Lund University, Box 118, 221 00 Lund, Sweden.
E-mail address: hamid.behjat@bme.lth.se (H. Behjat).

http://dx.doi.org/10.1016/j.neuroimage.2015.06.010
1053-8119/

As an alternative to GRFT, spatial wavelet transforms have been proposed as a means to non-linearly denoise functional data within frameworks of both classical inference (e.g., Aston et al., 2005; Ruttimann et al., 1998; Soleymani et al., 2009; Van De Ville et al., 2004, 2007; Wink and Roerdink, 2004) and Bayesian inference (e.g., Flandin and Penny, 2007; Sanyal and Ferreira, 2012). Since brain activity is highly localized in space (Bullmore et al., 2004), the property of sparse signal representation in the wavelet domain makes it possible to encode a cluster of active voxels with only a few coefficients. Such representation enhances the SNR as the background noise remains equally distributed among the wavelet coefficients, and thus, coefficient-wise statistical testing provides a higher sensitivity than voxel-wise testing. Wavelet-based SPM (WSPM) (Van De Ville et al., 2007) has the unique feature of treating thresholding within the wavelet domain as a denoising step only, and the statistical testing is deferred to a second thresholding on the reconstructed map within the spatial domain.

Accounting for intra-subject gray matter structure

Gaussian filters as well as standard wavelets such as those deployed by WSPM share several basic properties: they are (i) isotropic in structure, (ii) defined within regular Euclidean spaces (either a square in 2-D space or a cube in 3-D space) and (iii) stationary and quasi shift-invariant, meaning that their structure does not vary as applied to different regions within a volume. To various extents, these properties are opposed to the expected geometrical properties of the activation pattern. Since the gray matter (GM), within which the BOLD response is expected, has a convoluted structure, isotropically shaped activation patterns that cross boundaries of GM are unlikely. Moreover, due to the differences in the structure of the sulci and gyri across the brain, *intra-subject* variability of GM geometry is widely observed (Fischl et al., 2002; Riviere et al., 2002). Thus, it is essential to construct filters that adapt to the intricately convoluted GM domain rather than assuming the spatial characteristics of the underlying signal independent of its location. As a step in this direction, surface-based approaches have been proposed that restrict the analysis to the cortex by using reconstructions of the cortical surface. One such approach is the anatomically-informed basis function (AIBF) method proposed in Kiebel et al. (2000) and Kiebel and Friston (2002), where a forward model is determined for solving the inverse problem of explaining the distribution of the functional data using circular Gaussian basis functions defined on the cortical surface. In other approaches, here collectively referred to as cortical surface mapping (CSM), an interpolation scheme is used to map the functional data to the extracted cortical surface, followed by iteratively smoothing the data on the surface using different procedures such as diffusion smoothing (Andrade et al., 2001), heat kernel smoothing (Chung et al., 2005; Hagler et al., 2006) and spline smoothing (Qiu et al., 2006). Nevertheless, the problem of loss in spatial accuracy remains in CSM due to the irreversible smoothing. Aside from that, the mapping of volumetric data to a surface is challenging due to the variability in cortical thickness.

In the present paper, we introduce an alternative approach where we define a *volumetric* GM domain with the help of graph theory, where the graph vertices correspond to irregularly sampled points of the 3-D Euclidean space. Numerous neuroimaging applications have benefited from brain data being modeled as graphs and graph signals (Bullmore and Sporns, 2009; Richiardi et al., 2013). Here, we propose constructing brain graphs that encode local structural connectivity of GM *geometry* (irregular domain in 3-D), as opposed to the surface-based approaches which mainly incorporate cortical *topology* (2-D surface that is folded). Functional data can then be modeled as a scalar function (signal) defined on the vertices, and graph filters that diffuse only within the GM volume can be constructed. As such, the performance in fMRI brain activation mapping can be improved by attenuating the effect of non-signal components that originate from outside the GM.

With the increased interest in graph approaches to data analysis, a great amount of research has been devoted to generalizing signal processing operations to the graph setting (Shuman et al., 2013). This includes wavelet transforms, with the spectral graph wavelet transform (SGWT) proposed in Hammond et al. (2011) being an example. To prevent linear irreversible smoothing and to perform analysis at multiple scales, we propose the tight-frame SGWT (Leonardi and Van De Ville, 2013) to construct GM-adapted wavelets that are utilized to implement an advanced spatial transformation on fMRI data, integrated within the statistical analysis of the WSPM framework.

Accounting for inter-subject GM variability

Group-level fMRI activation mapping is further complicated by the *inter-subject* GM variability that is important to address. This variability renders the need for normalization of functional data to a template space, which, in turn, leads to better domain matching across subjects and improved statistical power as activations better overlap. Due to the observed difference in the extent of geometrical GM variety in the cerebrum and the cerebellum across subjects, it is advantageous to define cerebral and cerebellar template spaces separately.

The geometry of the cerebral cortex is not consistent across subjects. Although there are similarities in terms of the main fissures, the GM foldings are very inconsistent across individuals even in standard populations (Mangin et al., 2004; Riviere et al., 2002), see Fig. 1. The most commonly used cortical templates are based on either the anatomy of a single subject (Tzourio-Mazoyer et al., 2002) or the ensemble average over many subjects, such as the ICBM-152 (Evans et al., 1993) that defines the Montreal Neurological Institute (MNI) space. Such templates can be viewed as two extremes in GM representation: single subject templates take no account for inter-subject variability, and the group averaged templates, such as the ICBM-152, lack fine anatomical detail of the cerebral GM, which makes both categories unsuitable for our purpose. To address this problem, study specific template construction methods such as DARTEL (Ashburner, 2007) are of great benefit. The fast diffeomorphic image registration scheme proposed by Ashburner is among the best performing (Klein et al., 2009) and can produce a detailed group-averaged template GM through iterative, nonlinear warping of the segmented GM of a set of subjects.

The structural variability within the cerebellum is lower than in the cerebral cortex, since the cerebellar structure is relatively consistent across individuals in terms of the number and shape of its fissures (see Fig. 1). This observation has made it possible to create atlas templates of the cerebellum that prevent a loss in spatial accuracy of the anatomical detail. The spatially unbiased infra-tentorial (SUIT) cerebellum template (Diedrichsen, 2006) is the most accurate cerebellar template available to date. Compared to the ICBM 152 template (Evans et al., 1993) that is designed through averaging of T1 scans from 152 different subjects, SUIT is constructed from scans of 20 subjects, and at the same time, has the unique feature of being spatially unbiased; that is, the location of each of the structures is equal to its expected location in the MNI space across subjects (Diedrichsen, 2006).

Therefore, we propose the use of the SUIT atlas as the basis for defining a canonical cerebellar subgraph and the DARTEL for constructing study-dependent template cerebral subgraphs. A full GM-adapted brain graph is then defined by merging the two subgraphs.

Overview

The paper is organized as follows. In the Transform-based SPM (tSPM) section, WSPM is reviewed by generalizing the framework such that it 1) incorporates any linear spatial transform and 2) is set out for group-level analysis. In the Spectral graph wavelet transform (SGWT) section, we review the necessary concepts from graph theory and wavelet design. In the Spectral graph wavelet based SPM (tSPM^{sgwt}) section, we introduce the construction of GM-adapted graphs and wavelets, the required preprocessing steps and contrast mappings. In the Datasets section, we introduce a real dataset as well as the design

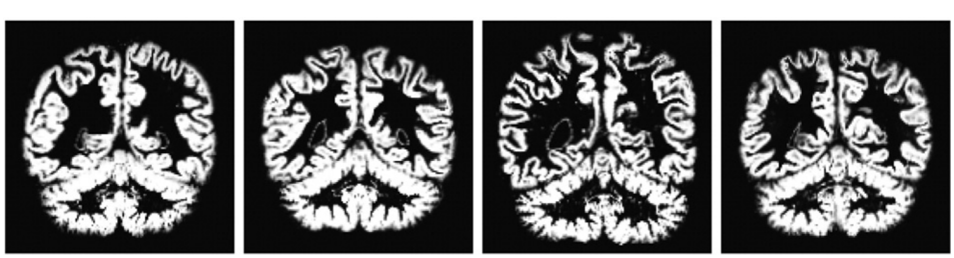


Fig. 1. Segmented gray matter of four individuals from an experimental dataset (see Experimental dataset section) illustrating the inter-subject variability. The variability is less significant in the cerebellum, as opposed to the cerebrum where the pattern of folding varies greatly from one individual to the other.

of a semi-synthetic multi-subject dataset. In the Results section, we demonstrate the utility of our approach by applying it to the semi-synthetic and experimental datasets, followed by a discussion in the Discussion section.

Methods

For the data structure we assume J subjects, where each subject has a structural scan $\{\mathcal{S}_j\}_{j=1}^{J}$ and a set of N_t functional volumes (across time), each containing N_v voxels. N_{r1} and N_{r2} regressors are assumed for the first level (subject-level) and second level (group-level) analysis, respectively. $\{\boldsymbol{X}_{1,j}\}_{j=1}^{J}$ and $\boldsymbol{X}_2$ denote the corresponding first level and second level design matrices which are of size $N_t \times N_{r1}$ and $J \times N_{r2}$, respectively.

Transform-based SPM (tSPM)

We present the WSPM in a more general way as transform-based SPM (tSPM) such that it: 1) incorporates any invertible linear spatial transform and 2) addresses fMRI activation mapping at the group-level.

The invertible linear spatial transform is a mapping from the voxel space $\mathbb{R}^{N_m}$ to $\mathbb{R}^{N_T}$, where $N_m \leq N_v$ and N_T denotes the dimension of the transform domain. $\boldsymbol{T} = [\boldsymbol{\zeta}_1|\boldsymbol{\zeta}_2|\ldots|\boldsymbol{\zeta}_{N_T}]^T$ denotes the corresponding orthogonal transformation matrix, with $\{\boldsymbol{\zeta}_k\}_{k=1}^{N_T}$ being the basis vectors spanning the transform domain. For the classical wavelets such as the discrete wavelet transform (DWT) used in Van De Ville et al. (2004, 2007), $N_m = N_v = N_T$, where the second equality follows from the DWT being non-redundant. For the graph setting proposed in this paper, N_m is the number of graph vertices, and $N_T = (S+1) \times N_m$, where S denotes the number of wavelet scales of the spectral graph wavelet transform (SGWT) that is a redundant graph transform.

Temporal modeling at the first level

Temporal modeling of the data is implemented in the native spatial domain, leading to a parameter map (effect size) for each subject. The approach is mass univariate where a general linear model (GLM) is fitted to the time course of each voxel $i \in \{1, \ldots, N_m\}$ for all subjects $j \in \{1, \ldots, J\}$ as

$$\boldsymbol{v}_{i,j} = \boldsymbol{X}_{1,j}\boldsymbol{\beta}_{i,j} + \boldsymbol{\epsilon}_{i,j}, \tag{1}$$

where $\boldsymbol{v}_{i,j}$ is an $N_t \times 1$ vector representing the time course of the ith voxel of the jth subject, $\boldsymbol{\beta}_{i,j}$ is an $N_{r1} \times 1$ vector of regression parameters and $\boldsymbol{\epsilon}_{i,j}$ is the vector of residual errors. Using the estimated $\boldsymbol{\beta}_{i,j}$, the effect sizes are constructed as

$$p_{i,j} = \boldsymbol{c}_1^T\boldsymbol{\beta}_{i,j} \tag{2}$$

where $\boldsymbol{c}_1$ is an $N_{r1} \times 1$ first level contrast vector that defines how the estimated regression parameters should be combined based on the hypothesis at test. By vectorizing $p_{i,j}$, the first level parameter maps (contrast maps) of the subjects are obtained as

$$\boldsymbol{p}_j = \left[p_{1,j}, p_{2,j}, \cdots, p_{N_m,j}\right]^T. \tag{3}$$

Transform-domain spatial modeling

The first level parameter maps are then taken into the spatial transform domain by applying the transform matrix $\boldsymbol{T}$ as

$$\boldsymbol{P}' = \boldsymbol{T}\left[\boldsymbol{p}_1|\boldsymbol{p}_2|\cdots|\boldsymbol{p}_J\right], \tag{4}$$

where $\boldsymbol{P}'$ is an $N_T \times J$ matrix, each column representing the transform domain coefficients of one subject.

In order to implement a standard group-level random-effects inference across subjects, a second-level GLM is fitted to the change in value of each coefficient across subjects, i.e., the rows of matrix $\boldsymbol{P}'$ denoted with $\boldsymbol{P}'_{k,:}$, as

$$\boldsymbol{P}'_{k,:} = \mathbf{X}_2\boldsymbol{\beta}_k + \boldsymbol{\epsilon}_k, \tag{5}$$

where $\boldsymbol{\beta}_k$ is an $N_{r2} \times 1$ vector of regression parameters and $\boldsymbol{\epsilon}_k$ is the vector of residual errors. The effect size μ_k and its uncertainty s_k for each coefficient k can then be obtained as

$$\mu_k = \boldsymbol{c}_2^T\boldsymbol{\beta}_k, \tag{6}$$

$$s_k = {\boldsymbol{\epsilon}_k}^T\boldsymbol{\epsilon}_k\boldsymbol{c}_2^T\left(\boldsymbol{X}_2^T\boldsymbol{X}_2\right)^{-1}\boldsymbol{c}_2, \tag{7}$$

respectively, where $\mathbf{c}_2$ is an $N_{r2} \times 1$ second-level contrast vector.[1]

Transform-domain denoising

At this stage, the main idea is to threshold the estimated effect sizes μ_k as a means to denoise the parameter maps. This is done by constructing transform-domain t-values corresponding to the coefficients as

$$t_k = \frac{\mu_k}{\sqrt{\frac{s_k}{J - \mathrm{rank}(\boldsymbol{X}_2)}}}, \tag{8}$$

where $\{t_k\}_{k=1}^{N_T}$ follow a Student's t-distribution with J degrees of freedom under the null hypothesis that stimulus-related activity is driven by noise only.

By subjecting t_k to a transform domain threshold $\tau_{_T}$, the reconstructed, denoised second level parameter map after bias correction is obtained as

$$\hat{u}_i = \min\left(\sum_{k=1}^{N_T} H(|t_k| - \tau_{_T})\mu_k\boldsymbol{\zeta}_k[i], \sum_{k=1}^{N_T}\mu_k\boldsymbol{\zeta}_k[i]\right), \tag{9}$$

where $i = 1, \ldots, N_m$, $H(\cdot)$ is the Heaviside step function and $\sum_{k=1}^{N_T}\mu_k\boldsymbol{\zeta}_k[i]$ denotes the elements of the unprocessed map (i.e., the linear estimate). The optimal value of $\tau_{_T}$ is obtained in combination with a spatial domain threshold $\tau_{_S}$ (see below Eq. (10)) such that the null hypothesis rejection probability in the spatial domain is properly controlled; see Van De Ville et al. (2004) for further details and derivations.

Spatial-domain statistical inference

The final detected parameter map is created by constructing spatial domain t-values and performing statistical significance testing by subjecting them to a spatial domain threshold $\tau_{_S}$ as (Van De Ville et al., 2004)

$$u_i = H\left(\frac{\hat{u}_i}{\sum_{k=1}^{N_T}\mu_k|\boldsymbol{\zeta}_k[i]|} - \tau_{_S}\right)\hat{u}_i. \tag{10}$$

To address the multiple comparison problem, the desired significance level is adjusted using Bonferroni correction when computing the optimal combination of $\tau_{_S}$ and $\tau_{_T}$.

Spectral graph wavelet transform (SGWT)

We now give a brief review of the SGWT on which our framework is based, and refer to Hammond et al. (2011) for further details.

[1] In particular, for the one sample t-test that is used for analyzing the datasets in this study, $\boldsymbol{X}_2 = [1, \cdots, 1]^T$, $\boldsymbol{c}_2 = 1$. As such, $\boldsymbol{\beta}_k$ will be a scalar value, and μ_k and s_k will correspond to unbiased estimates of the coefficients' sample mean and variance, respectively.

Classical wavelets

Classical wavelets defined in the Euclidean domain are constructed by shifting and scaling a mother wavelet $\boldsymbol{\psi}$, as $\boldsymbol{\psi}_{s,l}(x) = \frac{1}{s}\boldsymbol{\psi}\left(\frac{x-l}{s}\right)$, where $\boldsymbol{\psi}_{s,l}(x)$ denotes the wavelet at scale s and location l. To generalize wavelets to graphs, the analogue of the Fourier domain is required for graphs. In the Fourier domain, the classical continuous wavelet is given by

$$\boldsymbol{\psi}_{s,l}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{\boldsymbol{\psi}}(s\omega)e^{-j\omega l}e^{j\omega x}\mathrm{d}\omega, \tag{11}$$

where $\hat{\boldsymbol{\psi}}$ denotes the Fourier transform of $\boldsymbol{\psi}$. From Eq. (11), it is clear that the scaling parameter s solely affects the argument of $\hat{\boldsymbol{\psi}}$, and the shifting is represented by multiplication with the complex exponential e^{-jwl}. These two observations are essential when generalizing the wavelet transform to graphs.

Graphs and their spectra

Turning to spectral graph wavelets, a brief description of some basic notions of graphs and their spectra is in place. An undirected binary graph is described by its set of vertices $\mathcal{V}$ and edges $\mathcal{E}$ as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each edge is defined by a pair of unordered indices (m, n). For a graph with $|\mathcal{V}| = N_g$ vertices and no self-loops, the symmetric adjacency matrix $\boldsymbol{A}$ is given by the off-diagonal elements

$$a_{m,n} = \begin{cases} 1 & \text{if } (m,n)\in\mathcal{E}, \\ 0 & \text{otherwise}. \end{cases} \tag{12}$$

The degree matrix $\boldsymbol{D}$ is diagonal with elements $d_m = \sum_n a_{m,n}$. The normalized graph Laplacian matrix $\boldsymbol{L}$ is given by

$$\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}, \tag{13}$$

where $\boldsymbol{I}$ denotes the identity matrix. Assuming a single connected graph, and noting that $\boldsymbol{L}$ is symmetric and positive semi-definite, the eigendecomposition of $\boldsymbol{L}$ leads to a set of N_g real, non-negative eigenvalues, i.e., $0 = \lambda_1 \le \lambda_2 \cdots \le \lambda_{N_g} := \lambda_{max}$. The multiplicity of eigenvalues equal to zero reflects the number of connected components in the graph. The corresponding eigenvectors $\{\boldsymbol{\chi}_n\}_{n=1}^{N_g}$ form a complete set of orthonormal vectors, which defines the graph spectral domain (Chung, 1997).

Spectral graph wavelets

The wavelet shifting and scaling operations are not straightforward to generalize for signals lying on a non-Euclidean domain such as graph signals. Hammond et al. (2011) have tackled this issue by taking the design to the graph spectral domain and deriving the equivalence with the Fourier domain of conventional wavelets. Since the complex exponentials $\{e^{j\omega x}\}$ are both the basis functions of the Fourier transform, cf. Eq. (11), and the eigenfunctions of the 1-D Laplacian operator, an analogue spectral design for graphs can be defined based on the graph eigenvectors $\{\boldsymbol{\chi}_n\}_{n=1}^{N_g}$. Thus, the spectral graph wavelet functions $\{\boldsymbol{\psi}_{s,l}\}_{s=1}^{S-1}$ and scaling function $\boldsymbol{\phi}_l$ localized at each graph vertex l can be defined as

$$\boldsymbol{\psi}_{s,l}(x) = \sum_{n=1}^{N_g} g(\rho_s\lambda_n)\boldsymbol{\chi}_n^*(l)\boldsymbol{\chi}_n(x), \tag{14}$$

$$\boldsymbol{\phi}_l(x) = \sum_{n=1}^{N_g} h(\lambda_n)\boldsymbol{\chi}_n^*(l)\boldsymbol{\chi}_n(x), \tag{15}$$

where $\{\rho_s\}_{s=1}^{S-1}$ denote the scaling parameters, and $g(\cdot)$ and $h(\cdot)$ are wavelet and scaling generating kernels defined as weighted windows on the graph spectrum, respectively. An example of such a frame can be constructed by defining Meyer-like wavelet/scaling generating kernels in the spectral graph domain as shown in Appendix A.

The spectral graph wavelet and scaling coefficients of a graph signal $\boldsymbol{f}\in\mathbb{R}^{N_g}$ at scale s and location l are computed as

$$\boldsymbol{w}_{\psi}(s,l) = \langle \boldsymbol{f}, \boldsymbol{\psi}_{s,l}\rangle \overset{(14)}{=} \sum_{n=1}^{N_g} g(\rho_s\lambda_n)\hat{\boldsymbol{f}}(n)\boldsymbol{\chi}_n(l), \tag{16}$$

$$\boldsymbol{w}_{\phi}(l) = \langle \boldsymbol{f}, \boldsymbol{\phi}_l\rangle \overset{(15)}{=} \sum_{n=1}^{N_g} h(\lambda_n)\hat{\boldsymbol{f}}(n)\boldsymbol{\chi}_n(l), \tag{17}$$

respectively, where $\langle\cdot,\cdot\rangle$ denotes the inner product, and $\hat{\boldsymbol{f}}$ is the spectral representation of the graph signal $\boldsymbol{f}$ given by $\left\{\hat{\boldsymbol{f}}(n) = \langle \boldsymbol{f}, \boldsymbol{\chi}_n\rangle\right\}_{n=1}^{N_g}$.

Spectral graph wavelet based SPM (tSPM^{sgwt})

In what follows, the SGWT is adapted to a GM brain graph and integrated in the tSPM framework for group-level activation mapping, leading to a new graph based fMRI activation mapping referred to as tSPM^{sgwt}. First, the construction of GM-adapted graphs is introduced, followed by the design of the corresponding graph wavelets and a description of the required preprocessing steps and contrast mappings.

GM-adapted graphs

Defining a graph based on the GM geometry of the brain $\mathcal{G}_{gm}$ is an essential step in our approach. $\mathcal{G}_{gm}$ is constructed as a graph with two subgraph components: a cerebral subgraph $\mathcal{G}_{cbr}$ constructed using the GM information acquired from structural scans of multiple subjects, and a cerebellar subgraph $\mathcal{G}_{cbl}$ constructed based on the SUIT template atlas.

Cerebellar graph construction ($\mathcal{G}_{cbl}$). The cerebellar graph is designed as follows. First, the SUIT cerebellum template $\mathcal{S}_{suit}$ (Diedrichsen, 2006), defined in MNI coordinates at 1 mm resolution, is segmented using the unified segmentation algorithm (Ashburner and Friston, 2005) to extract its GM probability map, denoted $\mathcal{M}_{suit}$, with voxel probability values $p_i \in [0, 1]$. Next, a cerebellar GM mask is defined by intersecting the thresholded $\mathcal{M}_{suit}$ (threshold value 0.5) with the SUIT probabilistic atlas of cerebellar lobules $\mathcal{A}_{suit}$ (Diedrichsen et al., 2009). Intersection of the mask with $\mathcal{A}_{suit}$ ensures that only those voxels which define the cerebellar structure are kept, and that the brainstem is excluded. The resulting map is then morphologically filtered to remove isolated voxels; i.e., a voxel is defined as isolated if it is not adjacent to any other voxel within its 6-connected neighborhood in 3-D (see Fig. 2). We denote the resulting mask with $\mathcal{M}_{cbl}$ and the binary values of this mask with $v_i \in \{0, 1\}$.

The resolution of $\mathcal{M}_{cbl}$ is 1 mm isotropic and needs to be downsampled to the functional resolution. Therefore, we propose to

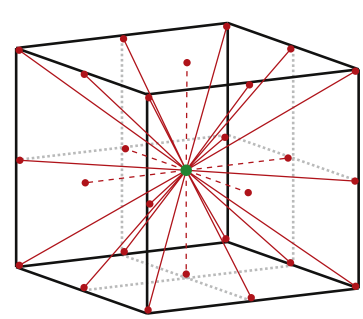


Fig. 2. Definition of neighborhood connectivities. The red dashed lines indicate the directions defining the 6-connectivity neighborhood with respect to the central green point, whereas the red dashed lines together with the red solid lines define the 26-connectivity neighborhood.

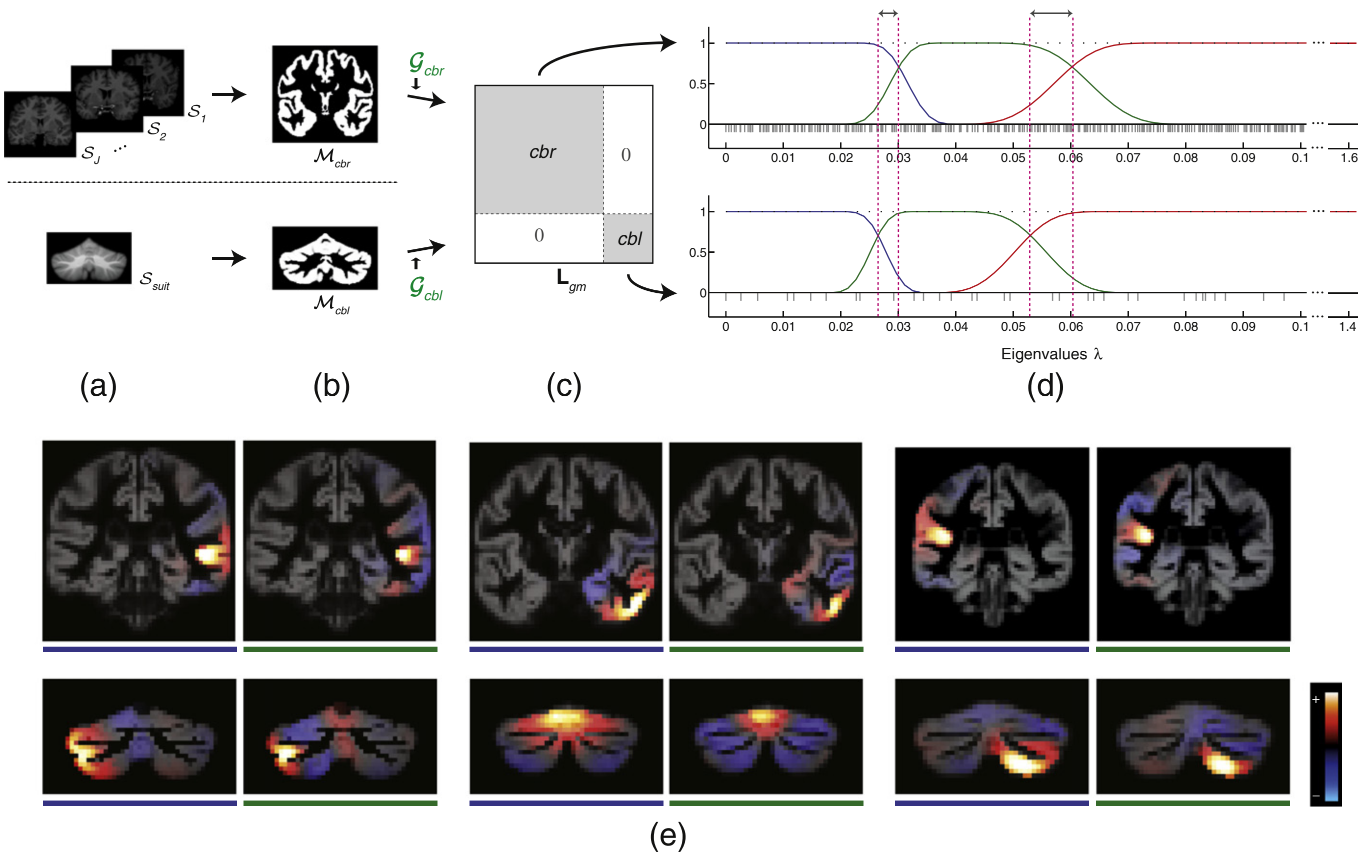


Fig. 3. GM-adapted graph and wavelet design. (a) Structural scans of the subjects (top) and the SUIT template atlas (bottom) are used to construct (b) the cerebral (top) and cerebellar (bottom) template GM masks, which are the basis for designing local structural connectivity subgraphs $\mathcal{G}_{cbr}$ and $\mathcal{G}_{cbl}$, respectively. (c) Block diagonal Laplacian matrix defining the spectrum of $\mathcal{G}_{gm}$. (d) Meyer-like windowing functions $h(\rho_S\lambda)$ and $\{g(\rho_s\lambda)\}_{s=1,2}$ at the lower end of the spectra of $\mathcal{G}_{cbr}$ and $\mathcal{G}_{cbl}$. A comparison of the two frames reveals the difference in the support of the corresponding scaling (blue) and wavelet (green and red) kernels of the two subgraphs (indicated by the arrows and dashed lines). The gray colored vertical parallel bars along the horizontal axis indicate the position of the eigenvalues within the spectral range. (e) Realizations of the GM-adapted wavelets localized at several different regions of the cerebrum (top row) and the cerebellum (bottom row) overlaid on GM. The figures are color coded (blue and green) corresponding to the spectral designs shown in (d); i.e. blue: ϕj and green: $\psi_{1,j}$, for six different j denoting the indices of the graph vertices where the wavelets are localized. Note that although the wavelets diffuse within 3-D space, only images of a coronal slice are illustrated.

first filter the mask weighted with the probability map, using a moving average 3-D box filter with window sizes approximating the functional voxel size in each dimension, followed by thresholding (threshold value 0.5). In particular, the downsampled voxels are obtained as

$$v_i' = H\left(\frac{1}{M}\sum_{m=1}^{M} v_m p_m - 0.5\right), \tag{18}$$

where $H(\cdot)$ is the Heaviside step function and M denotes the number of nearest neighbor voxels in the 1 mm mask required to interpolate the downsampled voxel. The resulting mask is denoted with $\mathcal{M}_{cbl}^{d}$. A binary cerebellar graph, denoted $\mathcal{G}_{cbl}$, is then constructed by considering the non-zero voxels in $\mathcal{M}_{cbl}^{d}$ as vertices and assigning edges by computing connections between adjacent voxels assuming 26-connectivity in 3-D, see Fig. 2.

Cerebral graph design ($\mathcal{G}_{cbr}$). The structural scans $\{\mathcal{S}_j\}_{j=1}^{J}$ are segmented, resulting in a set of GM probability maps, denoted with $\{\mathcal{M}_j\}_{j=1}^{J}$. The DARTEL iterative scheme of averaging and diffeomorphic registration of the probability maps is incorporated to construct a GM template for the group of subjects within the dataset. The algorithm converges after several iterations, resulting in a detailed GM tissue probability template at 1 mm resolution $\mathcal{M}_{tmp}$, and a set of flow fields $\{\mathcal{T}_j\}_{j=1}^{J}$ describing the deformation required for mapping $\{\mathcal{M}_j\}_{j=1}^{J}$ to this template.

$\mathcal{M}_{tmp}$ is not necessarily aligned to the MNI coordinates, and, therefore, a second level of transformation to the MNI space is required. Thus, $\mathcal{M}_{tmp}$ is updated by subjecting it to an affine transformation $\mathcal{T}_{tmp}$ that registers it to SPM's GM tissue probability map. For reference, we denote the thresholded (threshold value 0.5) and binarized version of $\mathcal{M}_{tmp}$ with $\mathcal{M}_{cbr}$. After downsampling $\mathcal{M}_{tmp}$ to the functional resolution, a binary cerebral GM mask is created by first thresholding the mask (threshold value 0.5), followed by excluding those voxels that lie within its intersection with $\mathcal{M}_{cbl}^{d}$. The mask is then morphologically filtered to remove isolated voxels (6-connectivity in 3-D). The nonzero voxels in the resulting mask are treated as graph vertices, and the edges are defined with the same approach as described for the cerebellar case, leading to a binary cerebral graph that we denote with $\mathcal{G}_{cbr}$.

Gray matter adapted wavelet design

The constructed subgraphs $\mathcal{G}_{cbr}$ and $\mathcal{G}_{cbl}$ are disconnected within $\mathcal{G}_{gm}$. Thus, the corresponding full brain Laplacian matrix $\boldsymbol{L}_{gm}$ is block diagonal (Fig. 3(c)). Let N_{cbr} and N_{cbl} denote the number of vertices in each subgraph. As such, the eigenvectors of $\boldsymbol{L}_{gm}$ comprise two sets of distinct basis $\{\chi_n^{cbr}\}_{n=1}^{N_{cbr}}$ and $\{\chi_n^{cbl}\}_{n=1}^{N_{cbl}}$, each separately spanning the cerebral and cerebellar GM domains, respectively. A tight frame design is then used to partition the spectral range of each subgraph and to allocate the support of each wavelet scale. We choose $g(\cdot)$ and $h(\cdot)$ (cf. Spectral graph wavelet transform (SGWT) section) such that the set of functions $\{\boldsymbol{\psi}_{s,l}, \boldsymbol{\phi}_l\}_{s\in S, l\in\mathcal{V}}$ forms a tight frame in $\ell_2(\mathcal{V})$ (Leonardi and Van De Ville,

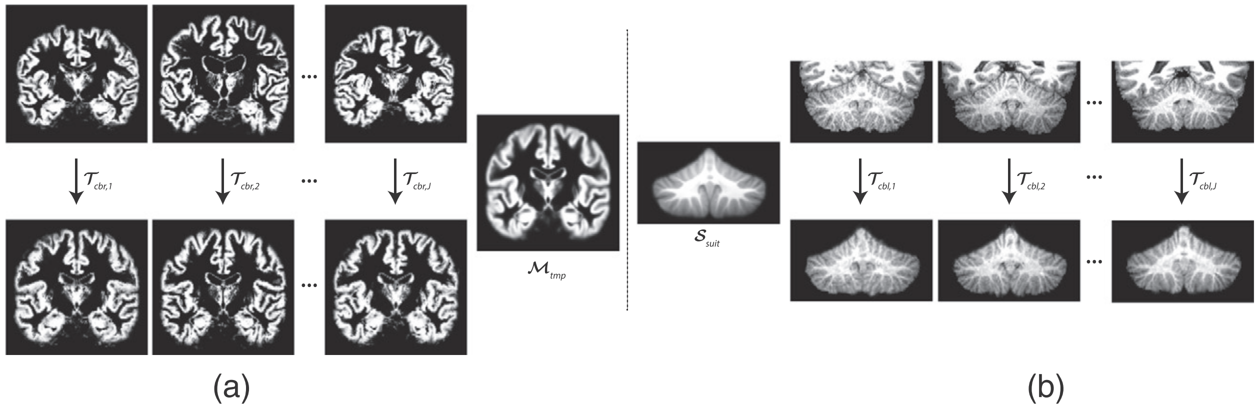


Fig. 4. Addressing inter-subject GM variability. The required transformations for normalizing the first level contrasts to the spaces of $\mathcal{G}_{cbr}$ and $\mathcal{G}_{cbl}$, denoted $\{\mathcal{T}_{cbr,j}\}_{j=1}^{J}$ and $\{\mathcal{T}_{cbl,j}\}_{j=1}^{J}$, respectively, are estimated based on transforming (a) the GM segmentations to $\mathcal{M}_{tmp}$ and (b) the cerebellum structural scans to $\mathcal{S}_{suit}$, respectively. Note that the normalized structural data (bottom row) are not directly used by tSMP^{sgwt}, but rather the estimated transformations.

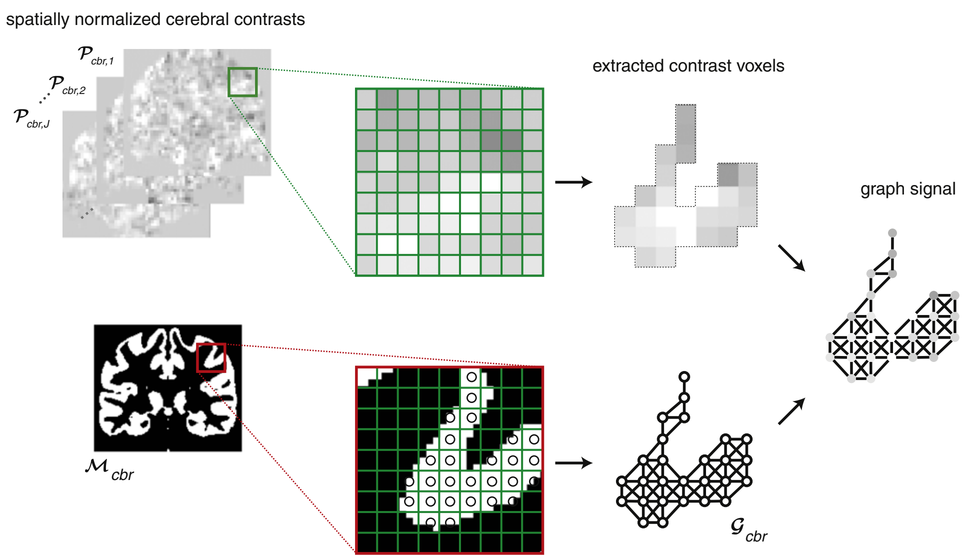


Fig. 5. Illustration of mapping an fMRI contrast map to a graph signal. Note that the graph design and mapping procedure is implemented in 3-D space, but for ease of illustration a 2-D example is presented.

2013). The tight frame construction is of importance as it leads to energy conservation between the original and transformed domain and enables an efficient inverse transform using the adjoint operator (Benedetto and Fickus, 2003). Also, we prefer the wavelet kernels at different scales to function as scaled bandpass filters and the scaling function as a lowpass filter. Thus, we use the Meyer-like spectral wavelet frame as proposed in Leonardi and Van De Ville (2013) (Appendix A).

Due to the large size of $\mathcal{G}_{gm}$,[2] diagonalising $\boldsymbol{L}_{gm}$ is computationally cumbersome. Therefore, we find it advantageous to use a polynomial approximation scheme for estimating the scaled generating kernels $g(\cdot)$ (see Eq. (14)) by low-order polynomials as proposed in Hammond et al. (2011). In doing so, the wavelet coefficients at each scale are obtained by applying a polynomial of $\boldsymbol{L}_{gm}$ to the data, and only an estimate of the range of the eigenvalues for each block of $\boldsymbol{L}_{gm}$ is required. $\boldsymbol{L}_{gm}$ has two zero eigenvalues as both $\mathcal{G}_{cbr}$ and $\mathcal{G}_{cbl}$ are constructed such that their single connectivity is ensured and are, at the same time, mutually disconnected. Thus, the lower spectral bound for both $\mathcal{G}_{cbr}$ and $\mathcal{G}_{cbl}$, i.e., λ_1, is 0. However, their upper spectral bounds depend on the complexity present in the GM structure of the respective region, with a higher value expected for $\mathcal{G}_{cbr}$ due to its more intricate geometrical detail. For $\mathcal{G}_{cbl}$ that is constructed to form a template that is canonical in nature, the upper bound is fixed, i.e., $\lambda_{max} = 1.4$, whereas it varies for $\mathcal{G}_{cbr}$ as it is constructed for each dataset separately. For the dataset tested in this study (cf. Experimental dataset section), $\lambda_{max} = 1.6$. This difference in upper bound, in turn, affects the support of the cerebellar and cerebral frames in terms of the range and width of the kernels, see Fig. 3(d). Fig. 3(e) shows several realizations of cerebellar and cerebral GM-adapted wavelets constructed using the spectral kernels shown in Fig. 3(d).

Functional data pre-processing & contrast mapping

The functional volumes are corrected for slice-timing and realigned with the first acquired image. The head movement parameters are then estimated and used as covariates in the design matrices $\{\boldsymbol{X}_{1,j}\}$. For each subject, the structural volumes are co-registered with the mean functional volume.

In order to reduce interpolation effects, all computations of the first level analysis are done in the native subject space, thus, no normalization is done on the functional data. Instead, the resulting first level

[2] The construction of $\mathcal{G}_{cbl}$ is study-independent, and the resulting $\mathcal{G}_{cbl}$ has 4465 vertices. The size of $\mathcal{G}_{cbr}$ depends on the dataset used. Using the dataset used in this study (see Experimental dataset section) leads to a $\mathcal{G}_{cbr}$ with 29,439 vertices, and thus, a $\mathcal{G}_{gm}$ with 33,904 vertices.

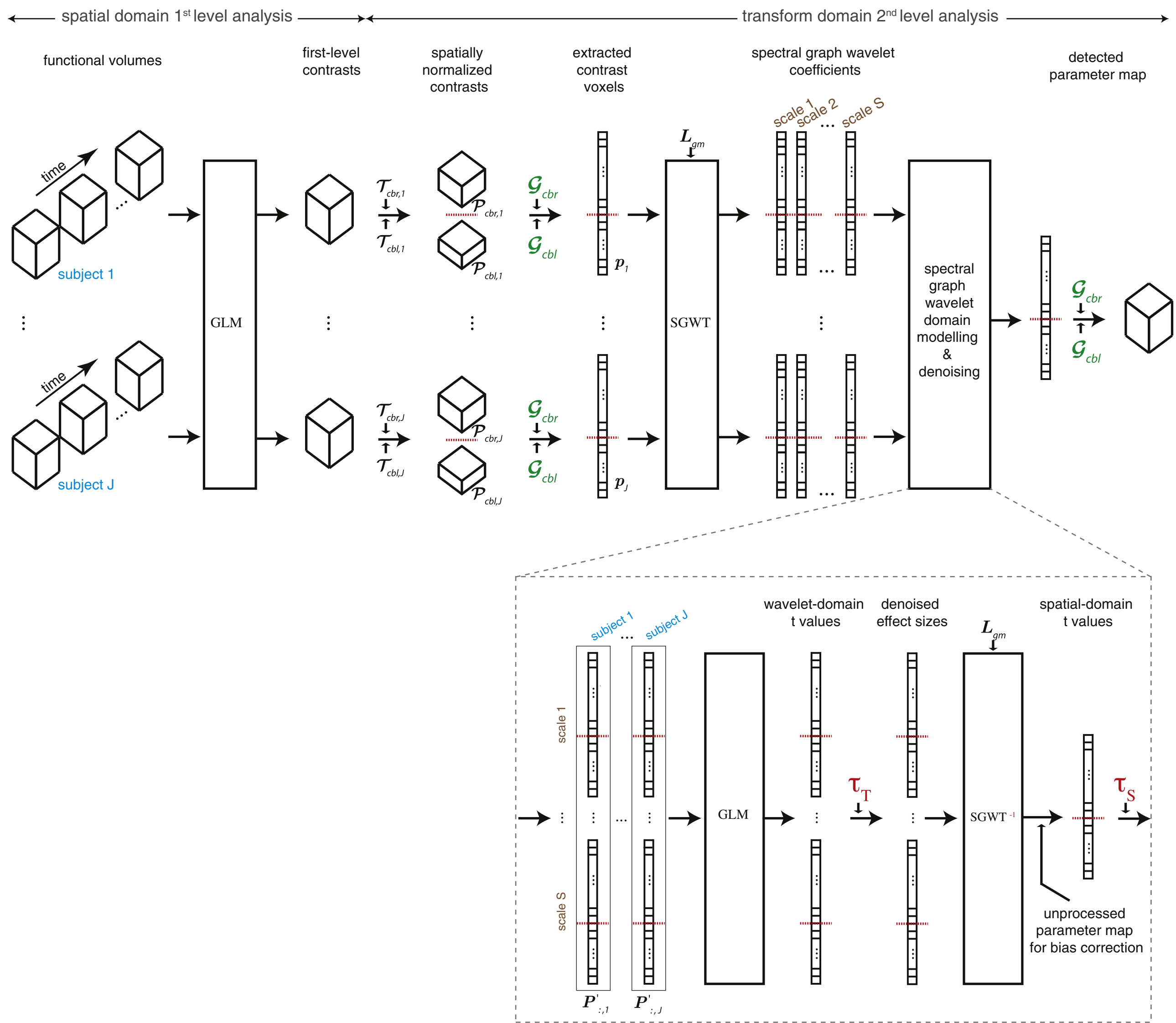


Fig. 6. Transform-based SPM using GM-adapted spectral graph wavelets (tSPM^{sgwt}).

parameter maps $\{\mathcal{P}_j\}_{j=1}^{J}$ [3] are normalized, which leads to normalized cerebral and cerebellar contrast maps denoted with $\{\mathcal{P}_{cbr,j}\}_{j=1}^{J}$ and $\{\mathcal{P}_{cbl,j}\}_{j=1}^{J}$, respectively. The required transformations for this normalization, denoted $\{\mathcal{T}_{cbr,j}\}_{j=1}^{J}$ and $\{\mathcal{T}_{cbl,j}\}_{j=1}^{J}$, respectively, are estimated based on transforming the subjects' structural data to the respective template domains.

Having designed $\mathcal{G}_{cbr}$, the GM of each subject can be constructed through warping $\mathcal{M}_{tmp}$, using the inverse of the estimated flow fields $\{\mathcal{T}_j\}_{j=1}^{J}$. Therefore, as the functional and structural volumes are co-registered, applying the same transformation to the corresponding contrast maps results in their within-subject registration. As such, a better overlap of effects can be gained for second level analysis, which can consequently lead to increased statistical power. The same reasoning holds for the cerebellum. $\{\mathcal{T}_{cbr,j}\}_{j=1}^{J}$ are constructed by combining $\{\mathcal{T}_j\}_{j=1}^{J}$ with $\mathcal{T}_{tmp}$ (cf. GM-adapted graphs section). The cerebellar structure of each subject is first annotated using a semi-automatic procedure, and the transformation is determined by mapping it to the SUIT template (Diedrichsen, 2006). The SUIT template is already defined in MNI space and thus no extra transformations are required. Fig. 4 illustrates examples of the resulting normalized structural data after applying the normalization transformations, both on the cerebrum and the cerebellum.

The co-registration and normalization procedures are crucial steps in the proposed framework due to the requirement of a one-to-one correspondence between the voxels of the functional and structural data as the constructed graphs are based on anatomical data. Fig. 5 illustrates the mapping of an fMRI contrast map to a graph signal. Note that only those voxels of the contrast volumes with a graph vertex counterpart are extracted, and their values are considered as elements of a signal lying on the graph. The extracted contrast voxels from both the cerebrum and cerebellum are vectorized, their SGWT is computed and the resulting set of wavelet coefficients are fed to transform-domain modeling. Fig. 6 illustrates an overview of tSPM^{sgwt}.

Datasets

Experimental dataset

Data from 26 healthy adults performing a slow event-related Eriksen flanker task were studied (Kelly et al., 2008). In each trial,

[3] $\mathcal{P}_j$ denotes the same first level parameter maps as $\boldsymbol{p}_j$ given in Eq. (3) but in non-vectorized format.

the subjects used one of two buttons to distinguish between congruent and incongruent trials (inter-trial interval varied between 8 s and 14 s with a mean of 12 s). In congruent trials, the flanking arrows pointed in the same direction as the central arrow (e.g., <<<<<), whereas in more demanding incongruent trials the flanking arrows pointed in opposite direction (e.g., <<><<).

Functional data were acquired using a 3 T scanner, where contiguous whole-brain functional volumes were obtained using echo planar imaging during each of the two flanker task blocks (TE = 30 ms, TR = 2000 ms, flip angle = 8°, matrix = 64 × 64, 40 slices, FOV = 192 mm, voxel size = 3 × 3 × 4 mm^3, 146 volumes). T1-weighted anatomical scans were collected using an MPRAGE sequence (TE = 3.93 ms, TR = 2500 ms, flip angle = 8°, 176 slices, FOV = 256 mm, voxel size = 30 × 20 × 10 mm^3). Using this dataset, a cerebrum graph $\mathcal{G}_{crb}$ with 29,439 vertices was created.

Semi-synthetic dataset

To evaluate the proposed algorithm, a semi-synthetic group fMRI dataset with known ground truth underlying simulated brain activity was created, using the anatomical scans of the 26 subjects of the flanker task dataset. The dataset was created to account for inter-subject variability in strength and location of the activity as well as to simulate a realistic activity pattern. Functional contrasts with known subject-specific ground truth activation patterns that diffuse according to the GM of each subject were created as follows.

First, for each subject j, the estimated flow field $\mathcal{T}_{cbr,j}$ (cf. Functional data pre-processing & contrast mapping section) was applied to its GM probability map, $\mathcal{M}_j$, resulting in a deformed-warped GM in MNI space. The resulting probability maps were then smoothed (FWHM 2 mm), binarized by thresholding at 50%, leading to a set of masks used to construct an unweighted graph for each subject with adjacency matrices $\{\boldsymbol{A}_j\}_{j=1}^{26}$. Three 5 × 5 × 5 mm cubic regions (125 voxels), two in the cerebrum and one in the cerebellum, were then chosen as candidates for three activation centers. The centers were spatially jittered by randomly picking three voxels (one from each region), creating a set of indicator vectors $\{\boldsymbol{x}_j\}_{j=1}^{26}$ defining the voxel location of the three centers for each subject. An activation pattern $\boldsymbol{y}_j$ that diffuses from the three center points along the individual's GM domain was constructed by consecutive application of the corresponding adjacency matrix $\boldsymbol{A}_j$ to $\boldsymbol{x}_j$ and confining the elements of $\boldsymbol{y}_j$ to the range [0, 1] as

$$\boldsymbol{y}_j(i) = \begin{cases} 1 & \text{if } \boldsymbol{x}'_j(i) > 1, \\ \sqrt[n]{\boldsymbol{x}'_j(i)} & \text{otherwise}, \end{cases} \tag{19}$$

where $\boldsymbol{x}_j{}' = \boldsymbol{A}_j^n \boldsymbol{x}_j$ and n defines the extent of the diffusion (Fig. 7). With this construction, not only do the patterns vary between subjects, but also the distribution of their values. The resulting patterns were downsampled to 3-mm resolution.

We treat these patterns as the ground truth first level contrasts for the individual subjects (i.e., one pattern per subject). Using the average of all 26 patterns, those voxels whose activity exceeded 0.5 and whose location corresponded to a voxel in $\mathcal{M}_{tmp}$ with at least 50% probability were defined as ground truth (Fig. 8(a)). The individual activity patterns were then corrupted with additive white Gaussian noise with variance $\sigma^2 = 1$.

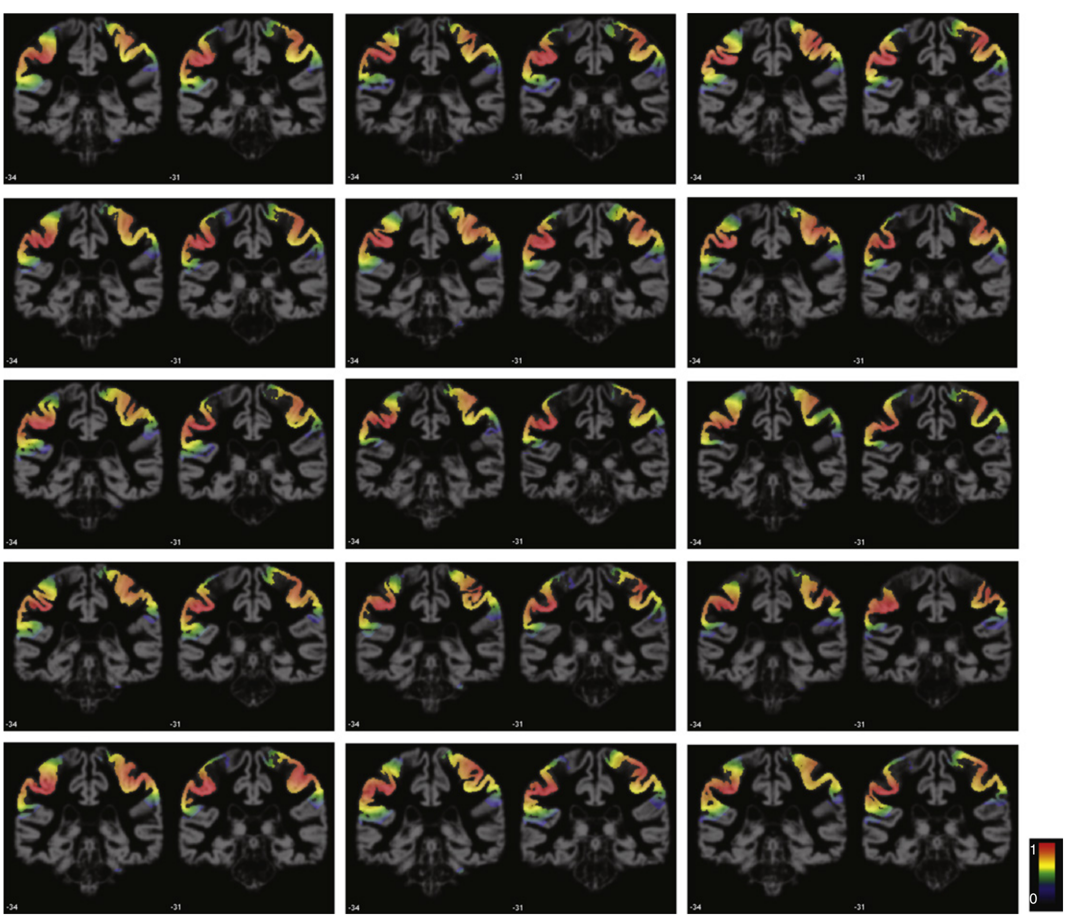
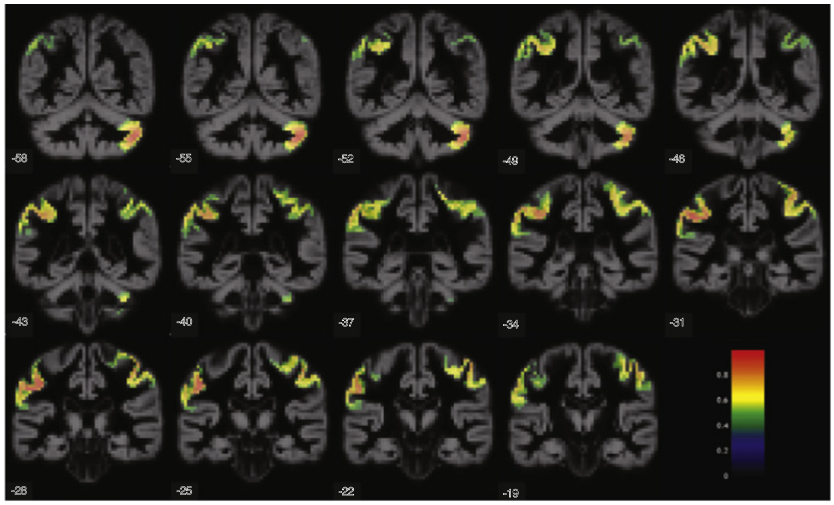
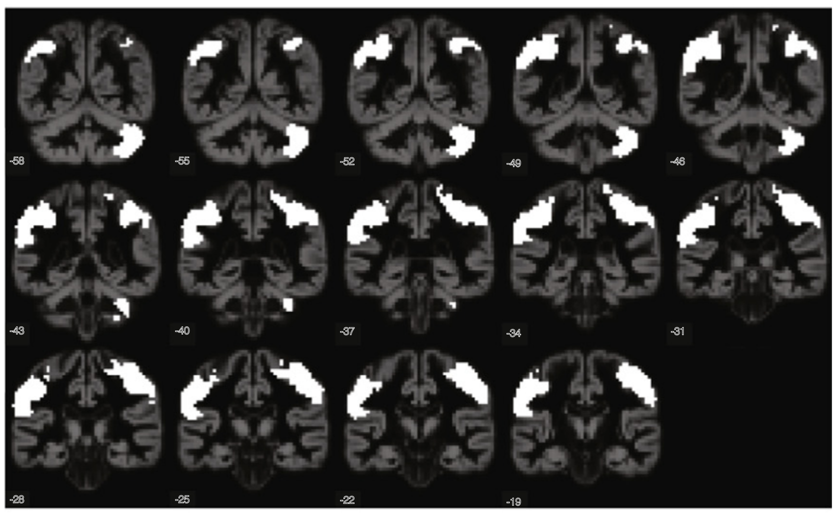


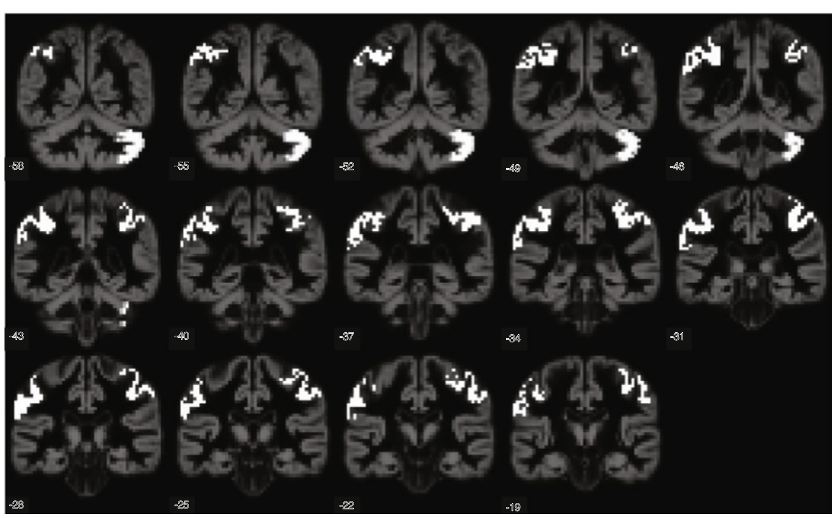
Fig. 7. Realizations of synthetic activity patterns $\boldsymbol{y}_j$ for several subjects, overlaid on each subject's GM.



(a)



(b)



(c)

Fig. 8. Semi-synthetic data. (a) Defined ground truth activation map, and activation maps detected by (b) SPM^{6mm} and (c) tSPM^{sgwt}, all overlaid on $\mathcal{M}_{tmp}$.

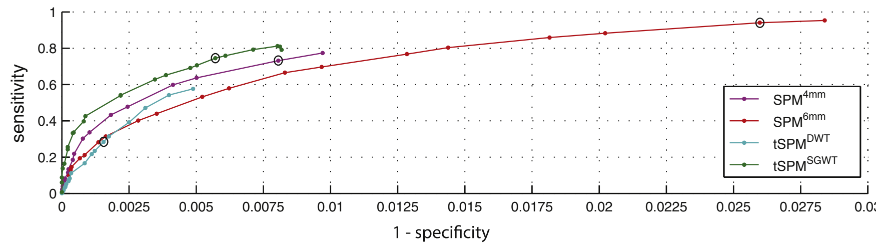


Fig. 9. Performance ROC curves for the four approaches. The circled points display the results achieved when testing at a significance level of $p < 0.05$ FWE corrected (i.e., corresponding to Table 1).

Results

Setup and performance measures

For comparative purposes, we implemented SPM (isotropic Gaussian smoothing with FWHM of 4 and 6 mm, denoted SPM^{4mm} and SPM^{6mm}, respectively), and transform-based SPM using standard orthogonal wavelets (2-D + Z, redundant), denoted tSPM^{dwt}, and GM-adapted wavelets, denoted tSPM^{sgwt}. SPM^{4mm} was performed to illustrate the effect of smoothing on spatial accuracy and detail of detections.

For both datasets, we present the results in terms of the number of detections. For the semi-synthetic data, we also present the number of true positives (TP) and false positives (FP), and the receiver operating characteristic (ROC) curve by varying *p*-values. Moreover, to indicate the detections that lie outside the expected GM region where a BOLD signal is expected, the detections are also categorized based on their affinity to the underlying tissue; GM50 denotes a mask covering the regions with a GM probability of at least 50%, i.e., greater than that of the probability of being white matter (WM) or cerebrospinal fluid (CSF). This metric provides an intuitive quality measure of activation mapping, especially for real data where the ground truth is not known. As an initial but reasonable approximation, detections that do not intersect with GM50 can be assumed to be FPs.

Semi-synthetic data

Table 1 presents the results in applying the different methods on the semi-synthetic data, when testing at a significance level of $p < 0.05$ familywise error (FWE) corrected. Both wavelet approaches have a significantly better control over FPs (type-I error control) compared to SPM^{6mm} (95 and 183 vs. 1590 FPs). tSPM^{sgwt} shows a significantly higher sensitivity compared to tSPM^{dwt} (1398 vs. 538 TPs). Although SPM^{6mm} also exhibits a higher sensitivity than tSPM^{sgwt} (1785 vs. 1398 TPs), it lacks spatial accuracy (1590 vs. 183 FPs). For example, the lack of spatial accuracy can be observed when comparing SPM^{6mm} and tSPM^{sgwt} detections at coronal slices −34 to −25 in Figs. 8(b)–(c), respectively, to the ground truth in Fig. 8(a).

Many SPM^{6mm} detections are outside GM50 (approximately 30%). The classification of FPs with respect to GM50 shows that the specificity of tSPM^{sgwt} is higher than SPM^{6mm} not only in total, but also when only taking detections within GM50 into account (183 vs. 644 FPs). In other words, the fine details detected by tSPM^{sgwt} cannot be obtained by intersecting SPM^{6mm} detections with the GM50 mask. Note that SPM^{6mm} has 387 extra TPs than tSPM^{sgwt} (i.e., 1785 − 1398), as compared to 461 extra FPs within GM50 (i.e., 644 − 183), which indicates less than random TP detection. This lack of detail in SPM^{6mm} detections is mainly due to the smoothing phase, though leading to higher sensitivity. By reducing the amount of smoothing, we observe that the sensitivities of SPM^{4mm} and tSPM^{sgwt} are similar (1389 vs. 1398 TPs) while tSPM^{sgwt} preserves its better specificity (493 vs. 183 FPs).

Fig. 9 shows ROC curves illustrating the specificity–sensitivity trade-off of the approaches as a function of the significance level. At the same level of specificity, tSPM^{sgwt} consistently shows superior sensitivity when compared to the other three approaches. Note that specificity of both tSPM^{dwt} and tSPM^{sgwt} is confined to high values due to the inverse inter-relation of the spatial and transform domain thresholds (i.e., τ_S and τ_T) that are inversely proportional (Van De Ville et al., 2004); as the significance level input to the algorithm drops below $\alpha = 0.01$, a reasonable low enough limit, τ_T significantly decreases, leading to an excess increase in τ_S, which in turn restricts detections and the sensitivity.

Table 1
Detection performance on the semi-synthetic dataset.

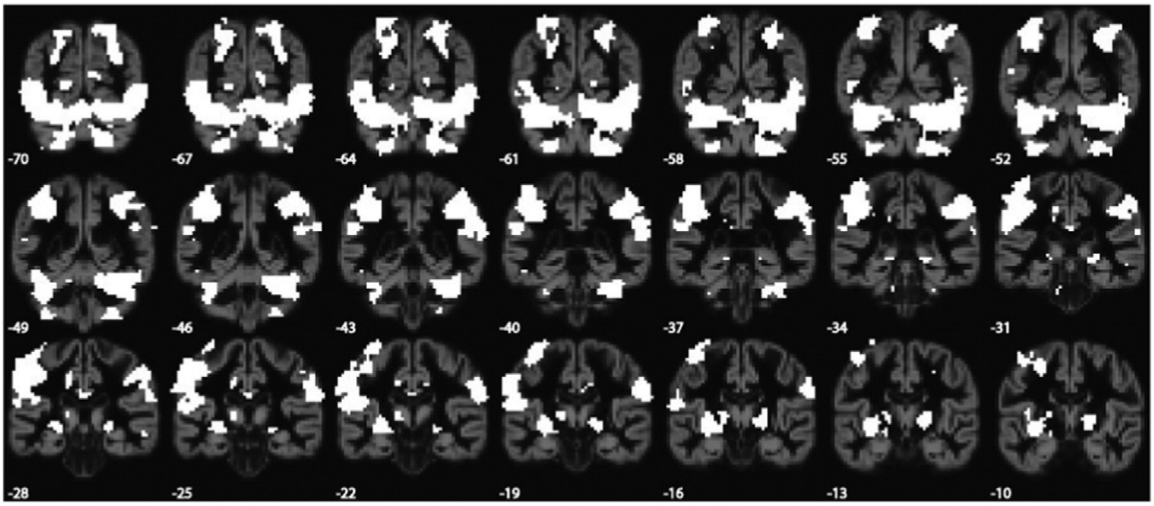
Method	SPM^{4mm}	SPM^{6mm}	tSPM^{dwt}	tSPM^{sgwt}
No. of detections − total	1882	3375	633	1581
No. of detections − in GM50	1587	2429	566	1581
No. of TPs/(sensitivity)	1389/(73%)	1785/(94%)	538/(28%)	1398/(75%)
No. of FPs − total/(specificity)	493/99.2%	1590/(97.4%)	95/(99.9%)	183/(99.4%)
No. of FPs − in GM50	198	644	28	183

Experimental data

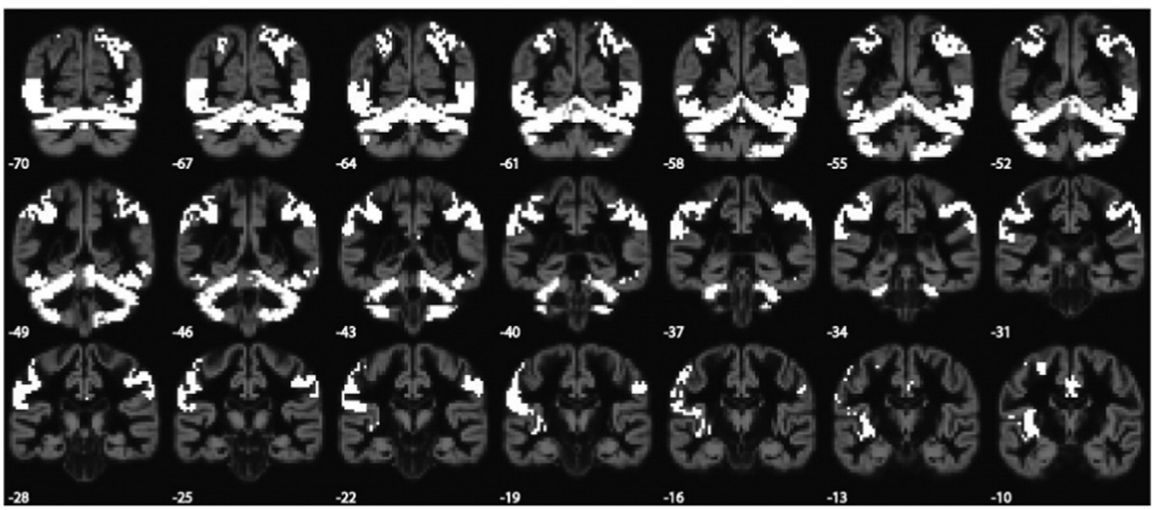
Table 2 presents the results for the flanker task dataset, where the incongruent active contrast was studied; i.e., detecting activations during incongruent trials.[4] All tests were performed at a significance level of $p < 0.05$ FWE corrected. A first observation is that both SPM^{6mm} and tSPM^{dwt} result in more detections than tSPM^{sgwt} (9678 and 7707 vs. 7274 detections, respectively). However, restricting the detections to GM50 within which tSPM^{sgwt} functions, we see that tSPM^{sgwt} outperforms both SPM^{6mm} and tSPM^{dwt} by 20% and 50% more detections, respectively (7274 vs. 6165 and 4973 detections, respectively).

Figs. 10(a)–(b) show the detection maps using SPM^{6mm} and tSPM^{sgwt}, respectively. tSPM^{sgwt} detections exhibit more spatial details, as evidenced for instance by the subtle patterns observed in the left and right upper cerebrocortical regions in slices −49 to −31. Moreover, tSPM^{sgwt} has better ability to detect connected patterns — a result which is elucidated by re-plotting SPM^{6mm} detections while keeping only those detections that fall within GM50, see Fig. 10(c). Examples of this phenomenon can be observed by comparing the detections in the cerebral cortex in slices −61 to −31 in Figs. 10(b)–(c).

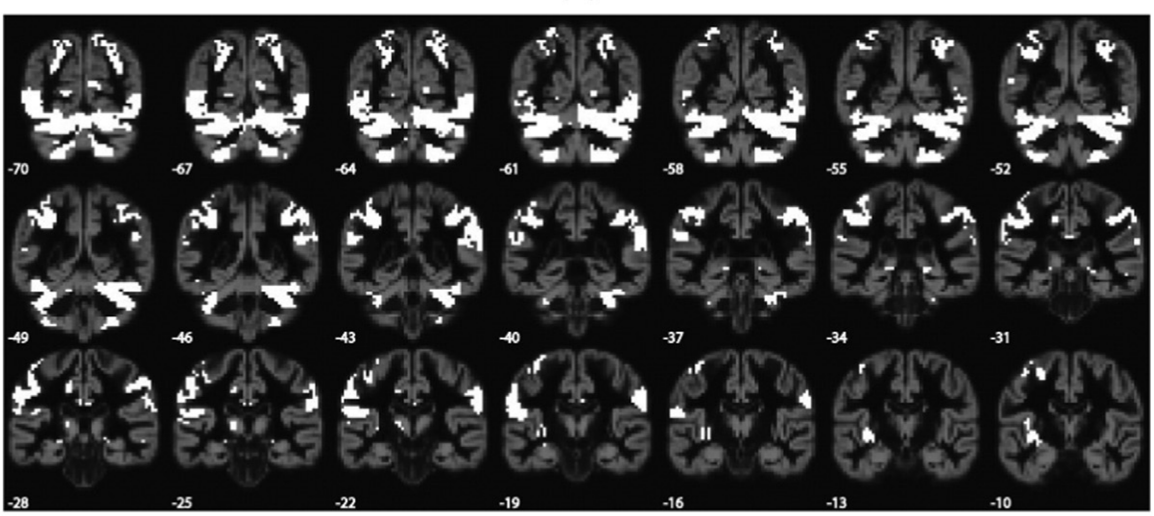
[4] For each subject j, $\boldsymbol{X}_{1,j} = [\boldsymbol{x}_{1,j}|\boldsymbol{x}_{2,j}|\boldsymbol{r}_{1,j}|\boldsymbol{r}_{2,j}| \cdots |\boldsymbol{r}_{6,j}, \mathbf{1}]$, where $\boldsymbol{x}_{1,j}$ and $\boldsymbol{x}_{2,j}$ are binary vectors indicating the onset of congruent and incongruent trials, respectively, $\{\boldsymbol{x}_{k,j}\}_{K=1,\cdots,6}$ are the estimated head movement parameters used as additional regressors, and the last column $\mathbf{1}$ is a constant vector which models the average activity. $\boldsymbol{c}_1$ is set to $[0, 1, 0, \cdots, 0]^T$ for detecting voxels whose activation increase in response to incongruent trials.



(a)



(b)



(c)

Fig. 10. Experimental data. Activation maps detected by (a) SPM^{6mm} and (b) tSPM^{sgwt}. (c) Activations detected by SPM^{6mm} that lie within GM50. All maps are overlaid on $\mathcal{M}_{tmp}$.

With respect to detection sensitivity within GM50, tSPM^{sgwt} in general performs better with additional activations compared to SPM^{6mm} in several regions, see cerebellar region in slices −55 to −34. It should also be noted that the detections of SPM^{6mm} include deep nuclei regions such as basal ganglia and thalamus, see medial detections in slices −19 to −10 in Fig. 10(a), which are not

Table 2
Detection performance on the experimental dataset.

Method	SPM^{4mm}	SPM^{6mm}	$tSPM^{dwt}$	$tSPM^{sgwt}$
No. of detections — total	5963	9678	7707	7274
No. of detections — in GM50	4182	6165	4973	7274

part of the graph definition of $tSPM^{sgwt}$ and thus, remain undetected by $tSPM^{sgwt}$.

Discussion

Improved spatial localization and empirical sensitivity

The enhanced spatial localization for $tSPM^{sgwt}$ compared to that of $tSPM^{dwt}$ and SPM is due to several reasons. First and foremost, prior knowledge about how and where the BOLD signal emerges is taken into account advantageously. WM is an example of a tissue with very low neuro-vascular coupling due to its sparse vascularization, which implies a weak or absent BOLD signal within WM (Logothetis and Wandell, 2004). As such, exploiting this prior knowledge about tissue–function relationship is essential. The prevention of detections in WM or CSF, which would be mainly driven by noise, also prevents these contributions from penetrating and getting mixed with the functional signal in GM. Although even within SPM, statistical analysis can be restricted to a subset of the voxels corresponding to GM, data still needs to be smoothed in order to be sufficiently in line with the Gaussian random field theory (Poline et al., 1997). Secondly, compared to SPM, there is no irreversible spatial smoothing of the functional data in wavelet approaches, which in turn results in enhanced spatial accuracy in detecting subtle activity patterns. The conventional approach of spatial smoothing with a non-adaptive Gaussian filter trades increased sensitivity for loss of information on the spatial extent and shape of the activation areas (Jo et al., 2008; Mikl et al., 2008). It has also been shown that isotropic Gaussian smoothing across GM and WM (i.e., without any GM constraint) can lead to displacement of activation peaks in *t*-value maps towards WM due to the difference in noise variance of the two tissue types (Reimold et al., 2006). Thirdly, the strong control of FPs using tSPM is linked to the underlying theory in selecting the threshold values through the bound over the null hypothesis rejection probability (Van De Ville et al., 2004). On the other hand, SPM yields FWE rates that are higher than the expected one, as recently reported in Eklund et al. (2012). This effect is decreased by reducing the amount of smoothing, as was observed when using SPM^{4mm}.

Fixing the desired significance, we observe that $tSPM^{sgwt}$ outperforms SPM and $tSPM^{dwt}$ in terms of sensitivity, see Fig. 9. Although the use of classical wavelets has shown good sensitivity in single-subject studies (Van De Ville et al., 2004, 2007), it performs worse in group studies due to inter-subject variability that is not countered by smoothing as in SPM, in particular, when there is only partial overlap between activations. Moreover, although the approach is multi-resolution, the wavelets dilate along the axes of the Cartesian coordinate system at a scale progression that is too fast. Instead, GM-adapted wavelets lead to much higher sensitivity compared to classical wavelets (i.e., 1398 vs. 483 detections, respectively) thanks to the adaptive scaling of wavelets that respect the GM domain. Moreover, empirical assessment of sensitivity on real data, by considering detections within GM50, reflects the higher sensitivity of $tSPM^{sgwt}$ compared to both $tSPM^{dwt}$ and SPM (see Table 2 and Fig. 10). Designing basis functions that dilate only within GM, not only prevents high noise data from WM or CSF to be “mixed” with GM data, but also leads to a more significant overlap of pure activations that follow GM foldings, which in turn, results in higher sensitivity.

Graph design for the brain

A binary graph design is adopted in the proposed framework. Our preliminary tests showed that weighting the edges based on the Euclidean distance does not improve the results, as the Euclidean distances between connected vertices in an isotropic 3-D grid are in the same range ($1, \sqrt{2}$ or $\sqrt{3}$ voxels, see Fig. 2). Also, defining weights based on GM probabilities did not improve the results (Behjat et al., 2014). Thus, the binary design was favored over an edge weighting scheme due to its simplicity. This construction resembles an $\mathcal{E}$-neighborhood graph commonly used in spectral clustering (Von Luxburg, 2007), with $\mathcal{E}$ equal to a distance of $\sqrt{3}$ voxels.

There are two main reasons why the design of a graph with region-specific subgraphs was preferred over the design of a single connected brain graph. First, the importance of separating the cerebrum from the cerebellum is to prevent activations from the ventral occipital lobe “bleeding” into the cerebellum, and vice versa. Examples of this phenomenon can be observed in SPM detections, see slices −70: −49 in Fig. 10(a). Second, the geometrical properties of the GM are different for the cerebrum and the cerebellum. At the resolution of currently available structural MRI scans and segmentation algorithms, there are more fine details observed within the whole range of the cerebral structure than within the more coarse geometry of the cerebellum. As such, constructing subgraphs specific to the cerebrum and cerebellum leads to wavelets specifically designed for each region, as the wavelet frame is based on their corresponding spectra.

By comparing the cerebellar mask $\mathcal{M}_{suit}$ to the cerebellar region of the average GM map $\mathcal{T}_{tmp}$, it can be observed that more spatial detail is preserved within the former, see Fig. 11. Moreover, the use of such a template mask leads to a study-independent canonical GM graph, with advantages similar to those of the SUIT atlas itself. The GM template constructed from SUIT was therefore preferred over the cerebellar region of $\mathcal{M}_{tmp}$.

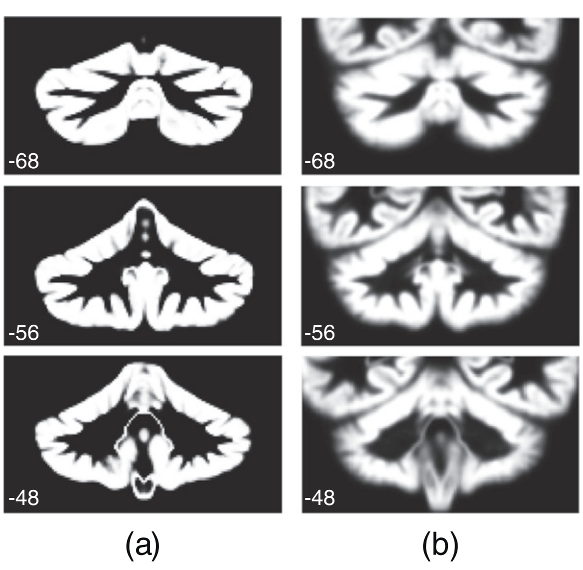


Fig. 11. Template cerebellar GM. Three coronal slices of (a) $\mathcal{M}_{suit}$ constructed using SUIT template and (b) the cerebellar region of $\mathcal{M}_{tmp}$ constructed using DARTEL. Finer spatial detail is observed in $\mathcal{M}_{suit}$ than in $\mathcal{M}_{tmp}$.

Comparison with other anatomically constrained methods

Previous studies following CSM approaches (Andrade et al., 2001; Chung et al., 2005; Hagler et al., 2006; Qiu et al., 2006), which anatomically constrain the activation mapping, suffer from a few shortcomings. First, the interpolation effects related to the projection of 3-D volumes onto the cortical surface is a major challenge (Grova et al., 2006; Operto et al., 2008). Second, smoothing leads to spatial inaccuracy and artificial shift of activations, be it using conventional volume-based approaches (e.g., SPM) or CSM (Jo et al., 2007, 2008; Mikl et al., 2008). Although the cortical 2-D diffusion smoothing of CSM approaches, rather than 3-D isotropic Gaussian smoothing of conventional volume-based approaches, prevents overlap of activation centers that are geodesically distant but close in a Euclidean sense (e.g., two points, at the opposite sides of a sulcus) (Andrade et al., 2001), the resulting smoothed signal can still be influenced by sources that are geodesically adjacent. tSPM^{sgwt} overcomes these two limitations by not smoothing the data in the first place and keeping the analysis within the native voxel-space. By modeling the GM as a graph based on the local neighborhood structure, we implicitly define a manifold (with non-zero thickness) on which wavelets dilate with respect to geodesic distances on this structure. In this sense, the design allows upscaling of local properties to a global scale. At the coarse scale, although tSPM^{sgwt} also implicitly inherits smoothing in applying the lowpass scaling function $h(\cdot)$ that can be interpreted as equivalent to Gaussian pre-filtering (Van De Ville et al., 2003), this is fundamentally different from pre-smoothing the data. First, the scaling coefficients are thresholded leading to *non-linear* denoising as opposed to the *linear* filtering done in pre-smoothing where all fine details of the data, be it noise or signal, are removed,[5] which in turn, leads to a loss in spatial detail of the underlying activity patterns. Second, the wavelet coefficients that survive thresholding are then used, together with the scaling coefficients, to reconstruct the de-noised signal, unlike the *irreversible* pre-smoothing as in CSM or SPM.

It is interesting to compare tSPM^{sgwt} with the AIBF approach (Kiebel et al., 2000; Kiebel and Friston, 2002). First, both approaches keep modeling in the spatial and temporal domain separated. However, temporal modeling precedes spatial modeling in tSPM^{sgwt}, i.e., spatial modeling is deferred to the second level analysis (cf. Transform-based SPM (tSPM) section), whereas the opposite is performed in AIBF. If single subject analysis similar to the current graph-based proposal is desired; i.e., only first level analysis, the original joint spatio-temporal model of WSPM as proposed in Van De Ville et al. (2004) can be used (Behjat et al., 2013). Second, tSPM^{sgwt} shares similarities with AIBF in the sense that smoothing of the data is prevented. They both project the data into a transform domain: AIBF to the space spanned by circular, user-specified wide Gaussian basis functions with local support defined on the flattened cortical surface, and tSPM^{sgwt} to a multi-scale GM domain spanned by GM-adapted wavelets defined within the voxel-space. The basis set is then used to represent, by a linear combination, the functional observations (BOLD volumes in AIBF and first level contrast maps in tSPM^{sgwt}). In tSPM^{sgwt}, the modeling is done within the wavelet domain, whereas in AIBF it is done in the voxel-space by back-projecting the constructed basis to the voxel-space. As such, AIBF is also prone to interpolation effects in back-projecting the basis from the cortical surface to the voxel-space, similar to that explained earlier with respect to mapping functional data from the voxel-space to the surface. Third, in tSPM^{sgwt} denoising is performed by subjecting the wavelet coefficients to a threshold (cf. Transform-based SPM (tSPM) section) and applying the inverse transform to project the data back to the native space, whereas in AIBF noise reduction is inherent in the anatomically constraint least-squares modeling. Fourth, in multi-subject AIBF (Kiebel and Friston, 2002), a single subject's cortical surface is used as a canonical surface to construct the basis set and to normalize functional data from all subjects. Here, we made use of recent population-level template atlas constructions as well as deformation algorithms to construct a group-level template of GM, in an attempt to address the observed inter-subject variability of GM structure more elegantly (see Fig. 4).

Another limitation of surface-based approaches is their inability to analyze non-cortical regions, such as the cerebellum and the deep nuclei (e.g., thalamus and basal ganglia), as they require a surface reconstruction, which would necessitate additional and tailored representation steps. As such, volumetric techniques, such as the present proposal, allow more easily to include non-cortical structures.

We also mention an alternative design of anatomically adaptive wavelets proposed in Ozkaya and Van De Ville (2011) and Ozkaya (2012). The design is based on the lifting scheme (Sweldens, 1996) where an irregular domain can be iteratively sectioned into a nested family of partitions at different spatial scales. As such, the basis construction in Ozkaya (2012) becomes shift variant (due to subsampling), while the wavelet basis in SGWT spans the space of eigenfunctions of $\boldsymbol{L}$ that are defined on the GM domain. Moreover, our approach has been directly devised for multi-subject analysis and as such addresses inter-subject GM variability. As both proposals use the segmented GM as the starting point for their construction, the lifting-based anatomically adapted wavelets can also benefit from the GM template as proposed in the present study.

Limitations

Although we have shown the applicability of tSPM^{sgwt} in enhancing fMRI activation mapping, there are still aspects that can be improved. The present framework lacks a systematic approach in determining the spectral coverage for each wavelet scale; this is a limitation for the SGWT design in general, as also reported in other applications than fMRI (e.g., Kim et al., 2014; Li and Hamza, 2013). The adopted spectral partitioning in the design has been found empirically by visual assessment of the wavelets and their characteristic scale. The current setting should nevertheless reasonably generalize to other datasets due to the general similarity in the extent and nature of the resulting group-level template GM domains.

Another limitation is that sub-cortical regions, such as the basal ganglia and thalamus, are currently not part of the graph. As such, several detections by SPM in the sub-cortical area were missed by tSPM^{sgwt} (cf. Experimental data section in Results). Although there is a precise atlas of deep cerebellar nuclei available (Diedrichsen et al., 2011), there does not seem to exist a detailed population-level atlas (mask) for the sub-cortical nuclei such as the thalamus and pallidum. Available atlases such as the AAL atlas (Tzourio-Mazoyer et al., 2002) are based on a single subject anatomy, and, when incorporated within our GM template, regions for the sub-cortical nuclei were too coarse and collapsed with the nearby delicate GM structure as defined by G_{cbr}. Therefore, these regions were decided not to be included in the current graph design. Provided a suitable atlas becomes available, it should be possible to include these regions as extra subgraphs to $\mathcal{G}_b$. Another option is to incorporate tractography information provided by diffusion tensor imaging data (Le Bihan et al., 2001). A similar idea in constructing such hybrid connectivity graphs has been recently proposed to improve electroencephalography-based source estimation (Hammond et al., 2013).

The computational burden of the proposed approach is another potentially limiting factor. The most costly part of the implementation is the absolute-value wavelet reconstruction needed for the spatial thresholding (cf. denominator of Eq. (10)). Since the graph wavelets are unique, they need to be computed explicitly at each and every vertex and scale. Although this computational burden can be reduced by pre-computing and storing the absolute-value-wavelets $\left\{\{,|\boldsymbol{\psi}_{s,l}|,|\boldsymbol{\phi}_l|\}_{s=1}^{S}\right\}_{l=1}^{N_{cbl}}$ for $\mathcal{G}_{cbl}$, this should still be done for each study-dependent $\mathcal{G}_{cbr}$. Another possibility would be to perform the reconstructions in parallel on the computer's graphic card, as such applicability has been shown for other fMRI analysis procedures (Eklund et al., 2013).

[5] Note that the Fourier transform of a Gaussian is also a Gaussian, and thus, applying such a Gaussian filter to the data corresponds to multiplying the spectrum of the data with a Gaussian.

Extension to structural studies

Another research avenue that can directly benefit from the proposed GM-adapted wavelet construction are structural studies. Euclidean wavelet approaches have been used to classify structural brain data (Canales-Rodriguez et al., 2013; Lao et al., 2004) as a means to assess structural morphometric differences between different populations of subjects. They have also been used to discriminate between healthy and pathological tissue by characterizing subtle changes in brain structure in a variety of diseases such as Alzheimer's disease, mild cognitive impairment and multiple sclerosis (Hackmack et al., 2012; Harrison et al., 2010). Interestingly, the recent proposal in Kim et al. (2014), also uses the SGWT to derive multi-scale shape descriptors that can be used to detect group-level effects. However, the approach uses cortical surface reconstructions, and as such, it comes with benefits and limitations of interpolation between the surface and volume as we discussed earlier. Nevertheless, the approach can be easily extended using the proposed volumetric GM graph and wavelet design.

Conclusion

We have extended fMRI activation mapping based on spatial multi-scale transforms to exploit the geometrical structure of the GM. We leveraged recent advances in graph-based wavelet design to incorporate this prior knowledge in the transformation. The procedure included the construction of a GM-adapted graph, including different subgraphs for cerebral and cerebellar regions. The wavelet-based SPM framework was presented to incorporate any linear spatial transform, including the spectral graph wavelet transform. Experimental and simulated results showed the potential of the proposed approach in terms of improved specificity and sensitivity for multi-subject studies, and to reveal fine-grained activity patterns.

Software

The source code of the proposed method will be made available to the community at miplab.epfl.ch/software/.

Acknowledgments

This work was supported by the Swedish Research Council under grant #2009-4584, the Swiss National Science Foundation under grant PP00P2-146318 and the Center for Biomedical Imaging (CIBM) of the Geneva and Lausanne Universities, the EPFL, and the Leenaards and Louis-Jeantet Foundations. Preliminary versions of the proposed graph wavelet-based fMRI activation mapping approach, but for single subject studies, were presented in Behjat et al. (2013, 2014).

Appendix A. Tight spectral graph Meyer-like wavelet frame

An example of tight graph wavelet frame can be constructed by defining Meyer-like wavelet and scaling kernels in the spectral graph domain as (Leonardi and Van De Ville, 2013)

$$g(\lambda) = \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{a}-1\right)\right)\right) & \forall\lambda\in]a, Ma] \\ \cos\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{Ma}-1\right)\right)\right) & \forall\lambda\in]Ma, M^2a] \\ 0 & \text{elsewhere} \end{cases} \tag{A.1}$$

$$h(\lambda) = \begin{cases} 1 & \forall\lambda\in[0, a] \\ \cos\left(\frac{\pi}{2}\nu\left(q\left(\frac{\lambda}{a}-1\right)\right)\right) & \forall\lambda\in]a, Ma] \\ 0 & \text{elsewhere} \end{cases} \tag{A.2}$$

respectively, where M denotes the dilation factor and is set by $q \in \mathbb{Z}$ as $M = \frac{q+1}{q}$, $a = (q-\epsilon)b, \epsilon\in]0, (1+M)^{-1}]$ and $a, b \in \mathbb{R}^+$. The classical dyadic dilation corresponds to $M = 2$. The J wavelet scales are defined as $\left\{t_j = \frac{a}{\lambda_{max}}M^j\right\}_{j=1,\dots,J}$. The resulting construction leads to a set of kernels where the support of each kernels is a strict subset of the eigenvectors of $\mathcal{L}$ (i.e., bandpass filters), and the union of the functions $|g(t_j\lambda)|^2$ and $|h(t_j\lambda)|^2$ forms a partition of unity, i.e., $\sum_{j=1}^{J}|g(t_j\lambda)|^2 + |h(t_j\lambda)|^2 = 1$.

References

Andrade, A., Kherif, F., Mangin, J., Worsley, K.J., Paradis, A., Simon, O., Dehaene, S., Bihan, D.L., Poline, J., 2001. Detection of fMRI activation using cortical surface mapping. Hum. Brain Mapp. 12 (2), 79–93.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38, 95–113.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26, 839–851.

Aston, J., Gunn, R., Hinz, R., Turkheimer, F., 2005. Wavelet variance components in image space for spatiotemporal neuroimaging data. NeuroImage 25 (1), 159–168.

Behjat, H., Leonardi, N., Van De Ville, D., 2013. Statistical parametric mapping of functional MRI data using wavelets adapted to the cerebral cortex. Proc. IEEE Int. Symp. Biomed. Imaging, pp. 1070–1073.

Behjat, H., Leonardi, N., Sörnmo, L., Van De Ville, D., 2014. Canonical cerebellar graph wavelets and their application to fMRI activation mapping. Proc. IEEE Int. Conf. Eng. Med. Biol. Soc., pp. 1039–1042

Benedetto, J., Fickus, M., 2003. Finite normalized tight frames. Adv. Comput. Math. 18 (2–4), 357–385.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186–198.

Bullmore, E., Fadili, J., Maxim, V., Sendur, L., Whitcher, B., Suckling, J., Brammer, M., Breakspear, M., 2004. Wavelets and functional magnetic resonance imaging of the human brain. NeuroImage 23 (Suppl. 1), S234–S249.

Canales-Rodriguez, E.J., Radua, J., Pomarol-Clotet, E., Sarro, S., Aleman-Gomez, Y., Iturria-Medina, Y., Salvador, R., 2013. Statistical analysis of brain tissue images in the wavelet domain: wavelet-based morphometry. NeuroImage 72, 214–226.

Chung, F., 1997. Spectral Graph Theory. AMS, Providence, RI.

Chung, M.K., Robbins, S.M., Dalton, K.M., Davidson, R.J., Alexander, A.L., Evans, A.C., 2005. Cortical thickness analysis in autism with heat kernel smoothing. NeuroImage 25 (4), 1256–1265.

Diedrichsen, J., 2006. A spatially unbiased atlas template of the human cerebellum. NeuroImage 33, 127–138.

Diedrichsen, J., Balsters, J.H., Flavell, J., Cussans, E., Ramnani, N., 2009. A probabilistic MR atlas of the human cerebellum. NeuroImage 46, 39–46.

Diedrichsen, J., Maderwald, S., Küper, M., Thürling, M., Rabe, K., Gizewski, E.R., Ladd, M.E., Timmann, D., 2011. Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure. NeuroImage 54 (3), 1786–1794.

Eklund, A., Andersson, M., Josephson, C., Johannesson, M., Knutsson, H., 2012. Does parametric fMRI analysis with SPM yield valid results? — an empirical study of 1484 rest datasets. NeuroImage 61 (3), 565–578.

Eklund, A., Dufort, P., Forsberg, D., LaConte, S.M., 2013. Medical image processing on the GPU — past, present and future. Med. Image Anal. 17 (8), 1073–1094.

Evans, A., Louis Collins, D., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M., 1993. 3D statistical neuroanatomical models from 305 MRI volumes. Nucl. Sci. Symp. Med. Imaging Conf., pp. 1813–1817

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33 (3), 341–355.

Flandin, G., Penny, W.D., 2007. Bayesian fMRI data analysis with sparse spatial basis function priors. NeuroImage 34 (3), 1108–1125.

Frackowiak, R.S., Friston, K., Frith, C., Dolan, R., Mazziotta, J., 1997. Human Brain Function. Academic Press.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J., 1994. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Grova, C., Makni, S., Flandin, G., Ciuciu, P., Gotman, J., Poline, J.B., 2006. Anatomically informed interpolation of fMRI data on the cortical surface. NeuroImage 31 (4), 1475–1486.

Hackmack, K., Paul, F., Weygandt, M., Allefeld, C., Haynes, J.D., 2012. Multi-scale classification of disease using structural MRI and wavelet transform. NeuroImage 62 (1), 48–58.

Hagler Jr., D.J., Saygin, A.P., Sereno, M.I., 2006. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. NeuroImage 33 (4), 1093–1103.

Hammond, D., Vandergheynst, P., Gribonval, R., 2011. Wavelets on graphs via spectral graph theory. Appl. Comput. Harmon. Anal. 30 (2), 129–150.

Hammond, D.K., Scherrer, B., Warfield, S.K., 2013. Cortical graph smoothing: a novel method for exploiting DWI-derived anatomical brain connectivity to improve EEG source estimation. IEEE Trans. Med. Imaging 32 (10), 1952–1963.

Harrison, L.C., Raunio, M., Holli, K.K., Luukkaala, T., Savio, S., Elovaara, I., Soimakallio, S., Eskola, H.J., Dastidar, P., 2010. MRI texture analysis in multiple sclerosis: toward a clinical analysis protocol. Acad. Radiol. 17 (6), 696–707.

Jo, H.J., Lee, J.M., Kim, J.H., Shin, Y.W., Kim, I.Y., Kwon, J.S., Kim, S.I., 2007. Spatial accuracy of fMRI activation influenced by volume- and surface-based spatial smoothing techniques. NeuroImage 34 (2), 550–564.

Jo, H.J., Lee, J.M., Kim, J.H., Choi, C.H., Gu, B.M., Kang, D.H., Ku, J., Kwon, J.S., Kim, S.I., 2008. Artificial shifting of fMRI activation localized by volume- and surface-based analyses. NeuroImage 40 (3), 1077–1089.

Kelly, A.M.C., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., Milham, M.P., 2008. Competition between functional brain networks mediates behavioral variability. NeuroImage 39, 527–537. (The dataset used is available online at https://openfmri.org/dataset/ds000102).

Kiebel, S., Friston, K.J., 2002. Anatomically informed basis functions in multisubject studies. Hum. Brain Mapp. 16 (1), 36–46.

Kiebel, S.J., Goebel, R., Friston, K.J., 2000. Anatomically informed basis functions. NeuroImage 11 (6), 656–667.

Kim, W.H., Singh, V., Chung, M.K., Hinrichs, C., Pachauri, D., Okonkwo, O.C., Johnson, S.C., 2014. Multi-resolutional shape features via non-Euclidean wavelets: applications to statistical analysis of cortical thickness. NeuroImage 93, 107–123.

Klein, A., Andersson, J., Ardekani, B., Ashburner, J., Avants, B., Chiang, M., Christensen, G., et al., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46 (3), 786–802.

Lao, Z., Shen, D., Xue, Z., Karacali, B., Resnick, S.M., Davatzikos, C., 2004. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. NeuroImage 21 (1), 46–57.

Le Bihan, D., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N., Chabriat, H., 2001. Diffusion tensor imaging: concepts and applications. J. Magn. Reson. Imaging 13 (4), 534–546.

Leonardi, N., Van De Ville, D., 2013. Tight wavelet frames on multislice graphs. IEEE Trans. Signal Process. 61, 3357–3367.

Li, C., Hamza, A.B., 2013. A multiresolution descriptor for deformable 3D shape retrieval. Vis. Comput. 29 (6–8), 513–524.

Logothetis, N., Wandell, B., 2004. Interpreting the BOLD signal. Annu. Rev. Physiol. 66, 735–769.

Mangin, J.F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Regis, J., 2004. A framework to study the cortical folding patterns. NeuroImage 23, S129–S138.

Mikl, M., Marecek, R., Hlustík, P., Pavlicová, M., Drastich, A., Chlebus, P., Brázdil, M., Krupa, P., 2008. Effects of spatial smoothing on fMRI group inferences. Magn. Reson. Imaging 26 (4), 490–503.

Ogawa, S., Menon, R.S., Tank, D.W., Kim, S.G., Merkle, H., Ellermann, J.M., Ugurbil, K., 1993. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. Biophys. J. 64 (3), 803–812.

Operto, G., Bulot, R., Anton, J.L., Coulon, O., 2008. Projection of fMRI data onto the cortical surface using anatomically-informed convolution kernels. NeuroImage 39 (1), 127–135.

Ozkaya, S., 2012. Randomized Wavelets on Arbitrary Domains and Applications to Functional MRI Analysis (Ph.D. thesis). Princeton University.

Ozkaya, S., Van De Ville, D., 2011. Anatomically adapted wavelets for integrated statistical analysis of fMRI data. Proc. IEEE Int. Symp. Biomed. Imaging, Chicago, IL, pp. 469–472.

Poline, J.B., Worsley, K.J., Evans, A.C., Friston, K.J., 1997. Combining spatial extent and peak intensity to test for activations in functional imaging. NeuroImage 5, 83–96.

Qiu, A., Bitouk, D., Miller, M.I., 2006. Smooth functional and structural maps on the neocortex via orthonormal bases of the Laplace–Beltrami operator. IEEE Trans. Med. Imaging 25 (10), 1296–1306.

Reimold, M., Slifstein, M., Heinz, A., Mueller-Schauenburg, W., Bares, R., 2006. Effect of spatial smoothing on t-maps: arguments for going back from t-maps to masked contrast images. J. Cereb. Blood Flow Metab. 26 (6), 751–759.

Richiardi, J., Achard, S., Bunke, H., Van De Ville, D., 2013. Machine learning with brain graphs: predictive modeling approaches for functional imaging in systems neuroscience. IEEE Signal Process. Mag. 30 (3), 58–70.

Riviere, D., Mangin, J.F., Papadopoulos-Orfanos, D., Martinez, J.M., Frouin, V., Regis, J., 2002. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. Med. Image Anal. 6 (2), 77–92.

Ruttimann, U.E., Unser, M., Rawlings, R.R., Rio, D., Ramsey, N.F., Mattay, V.S., Hommer, D., Frank, J., Weinberger, D.R., 1998. Statistical analysis of functional MRI data in the wavelet domain. IEEE Trans. Med. Imaging 17 (2), 142–154.

Sanyal, N., Ferreira, M.A., 2012. Bayesian hierarchical multi-subject multiscale analysis of functional MRI data. NeuroImage 63 (3), 1519–1531.

Shuman, D.I., Narang, S.K., Frossard, P., Ortega, A., Vandergheynst, P., 2013. The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process. Mag. 30 (3), 83–98.

Soleymani, M., Hossein-Zadeh, G., Soltanian-Zadeh, H., 2009. Fixed and random effect analysis of multi-subject fMRI data using wavelet transform. J. Neurosci. Methods 176 (2), 237–245.

Sweldens, W., 1996. The lifting scheme: a custom-design construction of biorthogonal wavelets. Appl. Comput. Harmon. Anal. 3 (2), 186–200.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289.

Van De Ville, D., Blu, T., Unser, M.A., 2003. Wavelets versus resels in the context of fMRI: establishing the link with SPM. SPIE's Symposium on Optical Science and Technology: Wavelets X vol. 5207, pp. 417–425.

Van De Ville, D., Blu, T., Unser, M., 2004. Integrated wavelet processing and spatial statistical testing of fMRI data. NeuroImage 23, 1472–1485.

Van De Ville, D., Seghier, M., Lazeyras, F., Blu, T., Unser, M., 2007. WSPM: wavelet-based statistical parametric mapping. NeuroImage 37, 1205–1217.

Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17 (4), 395–416.

Wink, A.M., Roerdink, J.B., 2004. Denoising functional MR images: a comparison of wavelet denoising and gaussian smoothing. IEEE Trans. Med. Imaging 23 (3), 374–387.