View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DSpace@MIT

NIH Public Access
Author Manuscript
Inf Process Med Imaging. Author manuscript; available in PMC 2010 March 11.

Published in final edited form as:
Inf Process Med Imaging. 2009 ; 21: 398–410.

Exploratory fMRI Analysis without Spatial Normalization

Danial Lashkari and **Polina Golland**
Computer Science and Artificial Intelligence Laboratory, MIT, USA

Abstract

We present an exploratory method for simultaneous parcellation of multisubject fMRI data into functionally coherent areas. The method is based on a solely functional representation of the fMRI data and a hierarchical probabilistic model that accounts for both inter-subject and intra-subject forms of variability in fMRI response. We employ a Variational Bayes approximation to fit the model to the data. The resulting algorithm finds a functional parcellation of the individual brains along with a set of population-level clusters, establishing correspondence between these two levels. The model eliminates the need for spatial normalization while still enabling us to fuse data from several subjects. We demonstrate the application of our method on a visual fMRI study.

1 Introduction

Analyzing data from multisubject neuroimaging studies is one of the challenging aspects of functional brain imaging. Presence of multiple sources of variability significantly complicates inference from the evidence provided by a group of individual fMRI scans [1]. First, the fMRI response in each subject varies from experiment to experiment giving rise to intra-subject variability. Moreover, two distinct sources contribute to inter-subject variability of fMRI signals. Since the brain structure is highly variable across subjects, establishing accurate correspondence among anatomical images of different subjects is intrinsically difficult. In addition to this anatomical variability, functional properties of the same anatomical structures are likely to vary somewhat across subjects.

The conventional localization approach to fMRI analysis constructs statistical parametric maps (SPM) [2] that indicate the significance of activation based on the hypotheses specified a priori. In order to perform group analysis, the method assumes all subjects are normalized into a common anatomical space where we can average the response across subjects. The performance of this approach is thus constrained by the limitations of spatial normalization techniques and the unknown relationship between function and anatomy. These issues are extensively discussed and studied in the literature, and different ways have been suggested to tackle them in the traditional localization framework [3,4,5,6].

The alternative approach employs unsupervised learning techniques to analyze fMRI data in an exploratory fashion. Most methods consider raw fMRI time courses and use clustering [7] or Independent Component Analysis (ICA) [8,9] to estimate a decomposition of the data into a set of distinct time courses and their corresponding spatial maps. Some variants use information from the experimental setup to define a measure of similarity between voxels, effectively projecting the original high-dimensional time courses to a low dimensional feature space, and then perform clustering in the new space [10,11]. Application of these exploratory methods to multisubject data has mainly relied on the same spatial mapping paradigm which requires spatial normalization [7]. More recently, a technique for exploratory group analysis was proposed that represents a voxel time course by a normalized profile of response to different experimental conditions [12]. The method clusters voxels from all the subjects together in the common functional space defined by these vectors. Although the results

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

demonstrate the success of this approach in a multisubject experiment, such a model does not take any form of inter-subject variability into account.

Here, we present an exploratory method for simultaneous functional brain parcellation based on fMRI response in a cohort of subjects. The method also constructs cross-subject parcel correspondence through common functional labels. The parcellation procedure follows directly from a hierarchical probabilistic model that explicitly accounts for both intra-subject and inter-subject forms of variability in fMRI response. Similar to [12], our method takes advantage of a common functional space, rather than spatial normalization, to fuse data from different subjects. In effect, our results yield a parcel-level functional normalization of the data by associating groups of voxels from different subjects with the same population-level functional pattern.

Our method operates on vectors that represent the response of voxels to the experimental conditions. At the subject-level, we model the set of these response vectors as a mixture of several, distinct, functionally-defined parcels, each with a representative response vector. We assume that these subject-level representative vectors can be further clustered into a smaller set of population-level groups. The representative vector for each population-level cluster defines a certain pattern of functionality. In the framework of hypothesis-based localization methods, a similar hierarchical structure is employed in random effects analysis (RFX) to model the two distinct types of variability [13]. In addition to offering the advantage of generating hypotheses from data, our method also eliminates the need for spatial normalization due to its fully functional representation of fMRI signals.

2 Method

We represent the group data by a set of response vectors $y_v^s \in IR^N$ where $s \in \{1, \cdots, S\}$ indexes the subject and $v \in \{1, \cdots, V\}$ denotes different voxels of the corresponding subject. These response vectors can be constructed from raw fMRI time courses based on the details of the experimental protocol [10]. In our work, we use the estimated GLM regression coefficients [2] for different experimental conditions as the vector components.

2.1 Model

Our hierarchical model is comprised of two clustering layers, each accounting for one type of variability in the data. We call the subject-level groupings of voxels in each subject *parcels* and refer to the population-level groupings of such parcels as *clusters*. In what follows, we use lower case Latin letters for random variables and Greek letters for distribution parameters. To facilitate notation, we sometimes define meta-variables composed of a group of vector variables, in which case we denote them with bold font.

Subject-Level Model—We assume that each response vector belongs to one of K parcels. We let the hidden variable $c_v^s = \left[c_{v,1}^s, \cdots, c_{v,K}^s \right]$ be a binary indicator vector such that $c_{v,k}^s = 1$ if y_v^s belongs to parcel k, and $c_{v,k}^s = 0$ otherwise. In our generative model, we assume a likelihood model for the response vectors

$$f(y; m, \lambda) = e^{-\lambda \quad D(y,m)+Z(\lambda)} \tag{1}$$

parameterized by the mean parameter m and the variance-related parameter λ. Here, $D(\cdot, \cdot)$ defines a distance between y and m, and $Z(\cdot)$ is the log-partition function. For the subject s, we assume a common parameter λ^s for all parcels and allow K possible mean parameters

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$m=\left[m_k^s\right]$ defining the functional centers (representative response vectors) of the parcels. Given parcel centers $\boldsymbol{m}$ and memberships $c=[c_v^s]$, different response vectors are independent:

$$p_{\boldsymbol{y}|\boldsymbol{cm}}(\boldsymbol{y}|\boldsymbol{c},\boldsymbol{m};\lambda) = \prod_{s,v}\left[\prod_k \left(f\left(y_v^s;m_k^s,\lambda^s\right)\right)^{c_{v,k}^s}\right], \tag{2}$$

where $y=[y_v^s]$ is the set of combined response vectors. We do not put any prior on the assignment variables $\boldsymbol{c}$, which is equivalent to assuming that all voxels within a subject are equally likely a priori to be from any parcel.

To eliminate irrelevant effects of the response magnitude and normalize response vectors across subjects, we scale the response vectors to have unit length, i.e., $\|y\|=\sqrt{\langle y,y\rangle}=1$ where $\langle\cdot,\cdot\rangle$ is the inner product in IR^N. Since these normalized vectors lie on the unit hyper-sphere S^{N-1}, an appropriate likelihood model is the directional *von Mises-Fisher distribution* [14]

$$f_{\mathrm{v}}(y;m,\lambda) = \left(\frac{\lambda}{2\pi}\right)^{N/2}\frac{1}{\lambda\quad I_{N/2-1}(\lambda)}e^{\lambda\langle y,m\rangle}, \tag{3}$$

where $I_\gamma(\cdot)$ is the γ-th order modified Bessel function of the first kind. The parameter λ controls the concentration of the distribution around the mean direction m in a way similar to the precision (reciprocal of variance) parameter of a Gaussian distribution. The model could be expressed based on (1) as:

$$D_{\mathrm{v}}(y,m)=1-\langle y,m\rangle, \qquad Z_{\mathrm{v}}(\lambda)=\frac{N}{2}\log\frac{\lambda}{2\pi}-\log\lambda-\lambda-\log I_{N/2-1}(\lambda). \tag{4}$$

In Appendix B, we provide another instantiation of our derivations for a Gaussian likelihood.

Population-Level Model—Conventional mixture-model approach to clustering fMRI data uses model (2), where $\boldsymbol{m}$ is a set of parameters estimated separately for different subjects. In our hierarchical model, we treat the parcel centers as random variables generated from a higher-level distribution that characterizes the functional patterns shared among all subjects. More specifically, we assume that the vectors m_k^s in each subject are generated from a mixture of L components. The mixture components are centered around representative vectors $\mu=[\mu_l]_{l=1}^L$ and have subject-specific mixture weights $w^s=\left[w_l^s\right]_{l=1}^L$. We let z_k^s be the L-dimensional binary vector that indicates the cluster membership of m_k^s (similar to c_v^s in the subject-level model). Weights w^s serve as the parameters of the multinomial distribution that generates cluster membership vectors z_k^s:

$$p_{\boldsymbol{z}|\boldsymbol{w}}(\boldsymbol{z}|\boldsymbol{w})=\prod_{s,k}\left[\prod_l\left(w_l^s\right)^{z_{k,l}^s}\right]. \tag{5}$$

We can now form the likelihood model for the parcel centers

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$p_{\boldsymbol{m}|\boldsymbol{z}}(\boldsymbol{m}|\boldsymbol{z};\mu,\sigma) = \prod_{s,k}\left[\prod_{l}\left(f\left(m_k^s;\mu_l,\sigma_l\right)\right)^{z_{k,l}^s}\right], \tag{6}$$

where f is consistent with the likelihood model for response vector $\boldsymbol{y}$. Finally, we set the common prior over the set of weights w^s to be a Dirichlet distribution over an $(L-1)$-simplex of multinomial probability weights, parameterized by positive-valued vector $\alpha \in IR^L$:

$$p_{\boldsymbol{w}}(\boldsymbol{w};\alpha) = \prod_{s}\mathrm{Dir}(w^s;\alpha) = \prod_{s}\left[\frac{\Gamma(\alpha_o)}{\prod_{l'=l}^{L}\Gamma(\alpha_{l'})}\prod_{l=1}^{L}\left(w_l^s\right)^{\alpha_l-1}\right], \tag{7}$$

where $\alpha_o = \Sigma_{l'}\alpha_{l'}$ and $\Gamma(\cdot)$ is the Gamma function. Concentration of the distribution around the expected value $Ew_l^s = \frac{\alpha_l}{\alpha_o}$ is controlled by α_o.

Joint Model—Fig. 1(a) illustrates our model. We denote the set of all hidden variables by $\boldsymbol{h} = \{\boldsymbol{c}, \boldsymbol{m}, \boldsymbol{z}, \boldsymbol{w}\}$ and the set of all model parameters by $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \sigma, \alpha, \lambda\}$. We can now form the joint model:

$$\begin{aligned}\log p(\boldsymbol{y},\boldsymbol{h};\theta) = &-\sum_{s,v,k} c_{v,k}^s\left[\lambda^s D\left(y_v^s, m_k^s\right) - Z(\lambda^s)\right] + \sum_{s,k,l} z_{k,l}^s \log w_l^s \\ &-\sum_{s,k,l} z_{k,l}^s\left[\sigma_l D\left(m_k^s,\mu_l\right) - Z(\sigma_l)\right] + \sum_{s}\left[\sum_{l}\alpha_l \log w_l^s + B(\alpha)\right],\end{aligned} \tag{8}$$

where $B(\alpha) = \log\Gamma(\alpha_o) - \sum_{l'}\log\Gamma(\alpha_{l'})$ and the constant prior over $\boldsymbol{c}$ is dropped.

Fitting this model to the data could be cast as the maximum likelihood (ML) parameter estimation for the observed data:

$$\theta^* = \underset{\theta}{\mathrm{argmax}} \quad \log p_{\boldsymbol{y}}(\boldsymbol{y};\theta) = \underset{\theta}{\mathrm{argmax}} \quad \log\int_{\boldsymbol{h}} p(\boldsymbol{y},\boldsymbol{h};\theta)\,. \tag{9}$$

Solving this problem requires integrating the joint distribution over all possible states of the hidden variables $\boldsymbol{h}$. Because of the first three terms in (8) that involve interaction between hidden variables, the integration in (9) is hard.

Variational Bayes—We employ a Variational Bayes approach to solve the problem [15].

Accordingly, we define a parameterized distribution $q\left(\boldsymbol{h};\tilde{\theta}\right)$ on the hidden variables to approximate the posterior $p(\boldsymbol{h}|\boldsymbol{y};\boldsymbol{\theta})$. As illustrated in Fig. 1(b), we assume a fully factorable model q whose components are compatible with the original model, that is, multinomial for parcel/cluster memberships $\boldsymbol{c}$ and $\boldsymbol{z}$, f for parcel centers $\boldsymbol{m}$, and Dirichlet for the weights $\boldsymbol{w}$:

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$$\begin{aligned} q\left(\boldsymbol{h};\tilde{\theta}\right) &= q_{\boldsymbol{c}}\left(\boldsymbol{c};\tilde{\gamma}\right)\quad q_{\boldsymbol{m}}\left(\boldsymbol{m};\tilde{\mu},\tilde{\sigma}\right)\quad q_{\boldsymbol{z}}\left(\boldsymbol{z};\tilde{\beta}\right)\quad q_{\boldsymbol{w}}\left(\boldsymbol{w};\tilde{\alpha}\right) \\ &= \prod_{s,v}\left(\prod_{k}\left(\tilde{\gamma}_{v,k}^{s}\right)^{c_{v,k}^{s}}\right)\cdot\prod_{s,k}f\left(m_{k}^{s};\tilde{\mu}_{k}^{s},\tilde{\sigma}_{k}^{s}\right)\cdot\prod_{s,k}\left(\prod_{l}\left(\tilde{\beta}_{k,l}^{s}\right)^{z_{k,l}^{s}}\right)\cdot\prod_{s}\mathrm{Dir}\left(w^{s};\tilde{\alpha}^{s}\right), \end{aligned} \tag{10}$$

where $\tilde{\theta}=\left\{\tilde{\gamma},\tilde{\mu},\tilde{\sigma},\tilde{\beta},\tilde{\alpha}\right\}$ is the combined set of parameters. With our choice of von Mises model f_{V}, we find the expression for the Variational Bayes cost function, also called the free energy:

$$\begin{aligned} \mathcal{F}\left(\tilde{\theta};\theta\right) \triangleq\ & E_{q}\log q\left(\boldsymbol{h}\right)-E_{q}\log p\left(\boldsymbol{y},\boldsymbol{h};\theta\right) \\ =\ & \sum_{s,k}Z\left(\tilde{\sigma}_{k}^{s}\right)+\sum_{s,v,k}\tilde{\gamma}_{v,k}^{s}\left[\log\tilde{\gamma}_{v,k}^{s}+\lambda^{s}D_{\mathrm{V}}\left(y_{v}^{s},\tilde{\mu}_{k}^{s}\right)-Z\left(\lambda^{s}\right)\right] \\ & +\sum_{s,k,l}\tilde{\beta}_{k,l}^{s}\left[\log\tilde{\beta}_{k,l}^{s}-\left(\Psi\left(\tilde{\alpha}_{l}^{s}\right)-\Psi\left(\tilde{\alpha}_{o}^{s}\right)\right)+\sigma_{l}D_{\mathrm{V}}\left(\tilde{\mu}_{k}^{s},\mu_{l}\right)-Z\left(\sigma_{l}\right)\right] \\ & +\sum_{s}\left[\sum_{l}\left(\tilde{\alpha}_{l}^{s}-1\right)\left(\Psi\left(\tilde{\alpha}_{l}^{s}\right)-\Psi\left(\tilde{\alpha}_{o}^{s}\right)\right)+B\left(\tilde{\alpha}^{s}\right)\right] \\ & -S\left[\sum_{l}\left(\alpha_{l}-1\right)\left(\frac{1}{S}\sum_{s}\left(\Psi\left(\tilde{\alpha}_{l}^{s}\right)-\Psi\left(\tilde{\alpha}_{o}^{s}\right)\right)\right)+B\left(\alpha\right)\right], \end{aligned} \tag{11}$$

where $\Psi\left(x\right)=\frac{d}{dx}\log\Gamma(x)$ and $\tilde{\alpha}_{o}^{s}=\sum_{l}\tilde{\alpha}_{l}^{s}$. Appendix A provides the details of the derivation.

2.2 Algorithm

We alternate between minimizing the free energy over posterior and model parameters $\tilde{\theta}$ and $\boldsymbol{\theta}$ similar to the basic Expectation-Maximization algorithm. For a fixed $\boldsymbol{\theta}$, minimizing $\mathcal{F}$ over $\tilde{\theta}$ corresponds to finding the approximate posterior distribution, i.e., making *inference* using the given set of model parameters. Minimization of the free energy with respect to $\boldsymbol{\theta}$ corresponds to *learning* the model parameters for a particular posterior distribution.

Inference—We can split the posterior parameters in (11) into two sets of non-interacting parameters, $\left(\tilde{\gamma},\tilde{\beta}\right)$ and $\left(\tilde{\mu},\tilde{\sigma},\tilde{\alpha}\right)$. Fixing the values of either set, the optimization problem leads to a closed form solution for the parameters in the other set. Hence, we choose to use a coordinate descent approach which simplifies the update rules and significantly improves the overall speed of the algorithm. The efficiency of this approximation allows us to repeat the algorithm with numerous different initializations to better search the space of solutions.

For fixed $\left(\tilde{\mu},\tilde{\sigma},\tilde{\alpha}\right)$, the cost function for $\tilde{\gamma}$ involves only the second term of (11); the third term is only relevant for $\tilde{\beta}$. We find

$$\tilde{\gamma}_{v,k}^{s}\propto e^{-\lambda^{s}D_{\mathrm{V}}\left(y_{v}^{s},\tilde{\mu}_{k}^{s}\right)} \qquad , \qquad \tilde{\beta}_{k,l}^{s}\propto e^{\Psi\left(\tilde{\alpha}_{l}^{s}\right)}e^{-\sigma_{l}D_{\mathrm{V}}\left(\tilde{\mu}_{k}^{s},\mu_{l}\right)+Z_{\mathrm{V}}\left(\sigma_{l}\right)}. \tag{12}$$

The update rules for the parcel/cluster assignments are quite similar to the standard cluster assignment update rules. The term $e^{\Psi\left(\tilde{\alpha}_{l}^{s}\right)}$ acts as a prior weight for cluster *l* in subject *s*.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

For fixed parameters $\left(\tilde{\gamma}, \tilde{\beta}\right)$, the parameter $\tilde{\alpha}$ becomes fully decoupled from $\left(\tilde{\mu}, \tilde{\sigma}\right)$. The solution takes the form:

$$\tilde{\alpha}_l^s = \alpha_l + \sum_k \tilde{\beta}_{k,l}^s \quad , \quad \tilde{\mu}_k^s \propto \sum_l \tilde{\beta}_{k,l}^s \sigma_l \mu_l + \lambda^s \sum_v \tilde{\gamma}_{v,k}^s y_v^s. \tag{13}$$

The update for $\tilde{\alpha}_l^s$ combines current values of cluster weight α_l with the sum of multinomial weights $\tilde{\beta}$ assigned to cluster l. The update for $\tilde{\mu}_k^s$, which is further normalized to unit length, linearly combines the cluster centers and the response vectors, each with corresponding weights $\tilde{\beta}_{k,l}^s$ and $\tilde{\gamma}_{v,k}^s$. We note that $\tilde{\sigma}$ only appears in the first term of (11); the minimum is achieved when $\tilde{\sigma}_k^s \to \infty$. However, since $\tilde{\sigma}$ does not appear in the learning stage, this does not affect the rest of the derivations.

Learning—The von Mises parameters are found as

$$\mu_l \propto \sum_{s,k} \tilde{\beta}_{k,l}^s \tilde{\mu}_k^s, \qquad \sigma_l = A_N^{-1}\left(\frac{\left\|\Sigma_{s,k} \tilde{\beta}_{k,l}^s \tilde{\mu}_k^s\right\|}{\Sigma_{s,k} \tilde{\beta}_{k,l}^s}\right),$$
$$\lambda^s = A_N^{-1}\left(\frac{1}{V_s} \sum_{v,k} \tilde{\gamma}_{v,k}^s \left\langle \tilde{\mu}_k^s, y_v^s \right\rangle\right), \tag{14}$$

where $A_N(\lambda) = I_{N/2}(\lambda) I_{N/2-1}(\lambda)$, and μ_l vectors are normalized to unit length. The update rules are similar to the parameter estimation steps of the ordinary single level clustering models. The last term of (11) is simply the Dirichlet log-likelihood function with the observed statistics $\frac{1}{S}\sum_s \left(\Psi\left(\tilde{\alpha}_l^s\right) - \Psi\left(\tilde{\alpha}_o^s\right)\right)$. Estimation of the Dirichlet parameters α and computing the inverse of $A_N(\cdot)$ involve solving nonlinear equations. We employ standard zero-finding algorithms to solve both problems [16,17].

Implementation—It is not hard to show that our algorithm always converges to a local minimum of the free energy function. This implies that, similar to most clustering algorithms, the solution depends on the initialization and selecting reasonable initializations can substantially improve the results. To initialize the values of parameters $\left(\tilde{\theta}, \theta\right)$, we first run a simple mixture model clustering into L components on a group data pooled from all subjects. We use the resulting cluster centers and weights as initial values for cluster centers $\boldsymbol{\mu}$ and weights α. Then we cluster individual data sets separately to find K clusters for each subject. We use the corresponding centers and assignment probabilities as the initialization of posterior parcel center and weight parameters $\tilde{\mu}$ and $\tilde{\gamma}$. The rest of the parameters are initialized to one or appropriate random numbers. For the results presented in the next section, we used between 40 to 70 initializations depending on the number of cluster/parcels.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We empirically observed that the following order of update rules, in each iteration, yields better results: $\left(\tilde{\gamma},\tilde{\beta}\right)$, $(\boldsymbol{\mu},\sigma)$, $\left(\tilde{\mu},\tilde{\alpha}\right)$, and $(\boldsymbol{\mu},\sigma,\lambda,\alpha)$, respectively. We terminate the algorithm for each initialization when the relative change in the free energy between successive iterations falls bellow 10^{-9}.

Similar to all other exploratory methods, we face the challenge of determining the right number of components. In our method, the question becomes more important due to the hierarchical structure of the model where we need to determine both L and K. We use a heuristic method to select these numbers but, as we will see, the results show robustness to changing cluster and parcel numbers. When we incrementally increase the number of clusters L, the value of the cost function monotonically decreases. If we interpret the value of the drop $\Delta\mathcal{F}=\mathcal{F}_{L-1}-\mathcal{F}_{L}$ in the free energy as the gain achieved by reaching L from $L-1$, a *good* number of clusters is the one for which this gain is locally optimal. We use the same heuristic in selecting the parcel number K as well.

3 Experimental Results

We demonstrate application of our method in an fMRI study of category selectivity in the visual cortex. Prior hypothesis-based studies have localized selective regions in the ventral visual pathway for a few categories of visual stimuli such as bodies, faces, and places [18]. Here, we aim to employ our method to *discover* modules shared across all subjects that show distinct profiles of response to the variety of images presented during an fMRI experiment [12]. In our analysis, a cluster center μ_l represents the selectivity profile of a shared module and the corresponding cluster assignments $\tilde{\beta}^{s}_{k,l}$ establish functional correspondence among the parcels $\tilde{\mu}^{s}_{k}$ found with a similar type of selectivity in different subjects. Agreement of the discovered selectivity profiles with the prior results in the literature and their consistency across representations of different images from the same category serves as a validations of our method.

In our experiment, six subjects viewed stimuli from eight categories (faces, places, bodies, trees, cars, animals, shoes, tools, vases) in a block-design fMRI experiment. Each block lasted 16 seconds and contained 20 images from one category. Sets of eight blocks (one for each category) were separated by an interval of fixation. Each subject viewed between 16 and 29 blocks for each category. The images were acquired using a Siemens 3T scanner and a custom 32-channel coil (EPI, flip angle = 90°, TR = 2s, 28 axial slices, voxel size = 1.5 × 1.5 × 2mm^3). We performed motion correction, spike detection, intensity normalization, and Gaussian smoothing with a 3 mm kernel using FsFast [19].

As another way to validate our results, we split blocks of images from each original category into four groups and labeled them as four *new* categories, creating a total of 32 categories. Since the resulting cluster response vectors represent profiles of selectivity, we expect the four components that correspond to the same category to be close to each other. We considered the voxels that passed a stimulus-versus-fixation contrast at a significance level of 10^{-2} in order to constrain the analysis to the visually active cortex. To form the space of response vectors, we estimated the regression coefficients for each category based on the General Linear Model [2]. We then normalized the resulting vectors of the regression coefficients to unit length to construct vectors $y^{s}_{v} \in S^{N-1}$ for $N = 32$.

First, we ran the analysis for fixed $K = 15$ and varied L. The plot of the change in the free energy between two successive values of L, Fig 2(a), suggests the choice of 7 or 9 population-level clusters. Fig. 2(c) shows the resulting cluster centers for $K = 15$ and $L = 9$ sorted based

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

on their cluster weights α. Parcel centers $\tilde{\mu}_k^s$ assigned to each μ_l, such that $l=\text{argmax}_{l'} \tilde{\beta}_{k,l'}$, are also presented along with corresponding cluster centers. The first four large clusters do not show specific category selectivity. This is in agreement with the fact that a large part of the visual cortex, including the early V1 and V2 areas, is exclusively devoted to low level processing of the visual field. Next, we observe clusters with body selectivity (clusters 5 and 7), face selectivity (cluster 6), and scene selectivity (clusters 8 and 9), which correspond to the well-known body, face, and place selective areas in the visual cortex [18]. As we expect, most discovered response vectors show consistency across different blocks corresponding to the same category.

It can be seen that there are very few outlier parcels, such as the animal-selective parcel in cluster 7 denoted by a dashed black line. Since there is not enough evidence for the presence of such a pattern across subjects, the algorithm assigns the parcel to the closest cluster. All other subject parcels closely follow the selectivity profile of their population clusters. We can interpret the deviations of parcel centers $\tilde{\mu}_k^s$ around each cluster center μ_l as the inter-subject variability. For instance, as we expect, parcels assigned to the non-selective clusters show highly variable patterns while the variability is lower for the highly selective clusters. Fig. 2 (e) shows the estimated clusters for $K = 15$ and $L = 7$. These results are similar and confirm our observations above. We conclude that the algorithm successfully matched functionally similar parcels across subjects.

To investigate the effect of changing the number of subject parcels, we let $L = 8$ and change K. Fig 2(b) shows the incremental difference $\Delta\mathcal{F}$ for K which behaves much more smoothly compared to the one of L. Selecting $K = 16$, a value which is slightly higher than its local trend, the results are shown in Fig. 2(d). Most clusters are almost the same as those of $L = 7$ and $K = 15$, except cluster 8 that is comprised solely of parcels from a single subject and does not represent a trend shared across all subjects. This explains why the improvement achieved by changing the number of clusters from 7 to 8 is much smaller than that of 6 to 7, or 8 to 9 in Fig. 2(a). Overall, these results demonstrate a relative robustness in the estimation of cluster centers to changes in the number of population and subject-level clusters and parcels.

We also examine the spatial maps of voxels assigned to the identified parcels. Based on our model, the posterior probability that voxel v in subject s belongs to cluster l equals $\sum_k \tilde{\gamma}_{v,k}^s \tilde{\beta}_{k,l}^s$. In our results, most of these probabilities are very close to 0 or 1 yielding almost binary spatial maps for the clusters. Fig. 3 shows the spatial map for the scene-selective cluster 7 in Fig. 2(d) in two different subjects. The maps clearly exhibit similarities in several regions in spite of their sparsity. We emphasize that the algorithm uses no spatial information other than the Gaussian smoothing performed in the pre-processing stage.

4 Discussion and Conclusions

In this paper, we presented a probabilistic model for unsupervised parcellation of the brain based on multisubject fMRI data. Applying this method to data from a vision experiment on a cohort of 6 subjects, we were able to discover several types of category selectivity in the visual cortex that have been previously reported using hypothesis-driven methods. The discovered profiles show consistency across different image representations from the same category.

Our experimental results show relative robustness to changes in the number of parcels and clusters. However, it is hard to analyze all possible settings in order to find the best combination of component numbers. It is therefore interesting to set for discovery of the natural number of

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

parcels as suggested by the data. We plan to study possible generalizations of the current model to include the number of components as an unknown using nonparametric approaches.

Another direction for extending the current model is the choice of a prior for parcel memberships $\boldsymbol{c}$. The current model assumes uniform prior on the assignment variables. Since we expect the subject-level parcels to be also spatially clustered, we can assume a prior for $\boldsymbol{c}$ with spatial smoothness properties as a rigorous modeling replacement for the Gaussian smoothing in the pre-processing. Furthermore, we can employ spatial normalization to parameterize and learn a specific inter-subject spatial prior for $\boldsymbol{c}$, merging our method with the conventional approach to the population analysis. We aim to investigate ways for designing such models.

Acknowledgments

We thank Ed Vul and Nancy Kanwisher for providing us with the fMRI data. This research was supported in part by the MIT McGovern Institute Neurotechnology Program, the NSF CAREER grant 0642971, and the NIH grants NIBIB NAMIC U54-EB005149, and NCRR NAC P41-RR13218.

A Variational Inference

For a distribution q on the hidden variables $\boldsymbol{h}$, the free energy functional

$$\mathcal{F}[q, p(\theta)] \triangleq E_q \log q(\boldsymbol{h}) - E_q \log p(\boldsymbol{y}, \boldsymbol{h}; \theta) = E_q \log \frac{q(\boldsymbol{h})}{p(\boldsymbol{h}|\boldsymbol{y}; \theta)} - \log p(\boldsymbol{y}; \theta)$$

serves as a lower bound for $-\log p(\boldsymbol{y}; \boldsymbol{\theta})$ since the first term is $\mathcal{D}_{KL}\left(q \| p_{\boldsymbol{h}|\boldsymbol{y}}\right) \geq 0$.

With the distribution $q\left(\boldsymbol{h}; \tilde{\theta}\right)$ defined in (10), the free energy $\mathcal{F}\left(\tilde{\theta}; \theta\right)$ is now a function of both sets of posterior and model parameters. Using the expansion of $\log p(\boldsymbol{y}, \boldsymbol{h}; \boldsymbol{\theta})$ in (8), we obtain

$$\begin{aligned}\mathcal{F} = & \sum_{s,v,k} E_q\left[c_{v,k}^s\right] \log \tilde{\gamma}_{v,k}^s - \sum_{s,k}\left[\tilde{\sigma}_k^s E_q\left[D\left(m_k^s, \tilde{\mu}_k^s\right)\right] - Z\left(\tilde{\sigma}_k^s\right)\right] \\ & + \sum_{s,k,l} E_q\left[z_{k,l}^s\right] \log \tilde{\beta}_{k,l}^s + \sum_s\left[\sum_l\left(\tilde{\alpha}_l^s - 1\right) E_q\left[\log w_l^s\right] + B\left(\tilde{\alpha}^s\right)\right] \\ & + \sum_{s,v,k} E_q\left[c_{v,k}^s\right]\left[\lambda^s E_q\left[D\left(y_v^s, m_k^s\right)\right] - Z\left(\lambda^s\right)\right] - \sum_{s,k,l} E_q\left[z_{k,l}^s\right] E_q\left[\log w_l^s\right] \\ & + \sum_{s,k,l} E_q\left[z_{k,l}^s\right]\left[\sigma_l E_q\left[D\left(m_k^s, \mu_l\right)\right] - Z\left(\sigma_l\right)\right] - \sum_{s,l}\left(\alpha_l - 1\right) E_q\left[\log w_l^s\right] - S \quad B(\alpha).\end{aligned} \tag{15}$$

The factored structure of q and properties of the Dirichlet distribution imply

$$\begin{aligned} E_q\left[c_{v,k}^s\right] & = \tilde{\gamma}_{v,k}^s, \qquad E_q\left[z_{k,l}^s\right] \\ & = \tilde{\beta}_{k,l}^s, \qquad E_{q_w}\left[\log w_l^s\right] \\ & = \Psi\left(\tilde{\alpha}_l^s\right) - \Psi\left(\tilde{\alpha}_o^s\right), \end{aligned} \tag{16}$$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

where $\Psi(x) = \frac{d}{dx} \log \Gamma(x)$ and $\tilde{\alpha}_o^s = \sum_l \tilde{\alpha}_l^s$. From (4), the distance D_V is linear in the case of von Mises distribution. Thus, the value of a term $E_q[D_V(m, \eta)]$ under the distribution $q(m) = f_V\left(m; \tilde{\mu}_k^s, \tilde{\sigma}_k^s\right)$ is simply equal to $D_V\left(\eta, \tilde{\mu}_k^s\right)$. Substituting these expressions in (15) completes the computation of the free energy cost function in (11).

B Gaussian Model

We show the derivation for the case when the likelihood model f is the commonly used Gaussian distribution. Based on the general expression (1), $f_G(y; m, \lambda)$ is defined by $D_G(y, m) = \frac{1}{2}\|y - m\|^2$ and $Z_G(\lambda) = \frac{N}{2} \log \frac{\lambda}{2\pi}$. The interesting technical distinction between Gaussian and von Mises distributions appears in the computation of the free energy (15). The quadratic expression $E_q[D_G(m, \eta)]$ under the distribution $q(m) = f_G\left(m; \tilde{\mu}_k^s, \tilde{\sigma}_k^s\right)$ is equal to $D_G\left(\eta, \tilde{\mu}_k^s\right) + \frac{N}{2\tilde{\sigma}_k^s}$.

In the inference stage, we find the cluster and parcel assignment update rules:

$$\tilde{\gamma}_{v,k}^s \propto e^{-\lambda^s D_G\left(y_v^s, \tilde{\mu}_k^s\right) - \frac{N\lambda^s}{2\tilde{\sigma}_k^s}}, \quad \tilde{\beta}_{k,l}^s \propto e^{\Psi\left(\tilde{\alpha}_l^s\right)} \quad e^{-\sigma_l D_G\left(\tilde{\mu}_k^s, \mu_l\right) + Z_G(\sigma_l) - \frac{N\sigma_l}{2\tilde{\sigma}_k^s}}. \tag{17}$$

Comparing these rules with (12), we interpret the last terms in the exponents of (17) as a correction that accounts for the uncertainty in cluster center estimates $\tilde{\mu}$, reflected by the values of $\tilde{\sigma}$. While theupdatefor $\tilde{\alpha}$ remains the same, the mean and precision of parcel centers are updated as

$$\tilde{\mu}_k^s = \frac{\sum_l \tilde{\beta}_{k,l}^s \sigma_l \mu_l + \lambda^s \sum_v \tilde{\gamma}_{v,k}^s y_v^s}{\sum_l \tilde{\beta}_{k,l}^s \sigma_l + \lambda^s \sum_v \tilde{\gamma}_{v,k}^s}, \quad \tilde{\sigma}_k^s = V_s \lambda^s + \sum_l \tilde{\beta}_{k,l}^s \sigma_l. \tag{18}$$

Here, in contrast to the von Mises case, the finite values of $\tilde{\alpha}_k^s$ propagate the effects of our posterior uncertainty about $\boldsymbol{m}$ to the estimates of the other variables. In the learning stage, the Gaussian parameters are computed according to the update rules

$$\mu_l = \frac{1}{\sum_{s,k} \tilde{\beta}_{k,l}^s} \sum_{s,k} \tilde{\beta}_{k,l}^s \tilde{\mu}_k^s, \quad (\sigma_l)^{-1} = \frac{1}{\sum_{s,k} \tilde{\beta}_{k,l}^s} \sum_{s,k} \tilde{\beta}_{k,l}^s \left[\left(\tilde{\sigma}_k^s\right)^{-1} + \frac{1}{N}\|\tilde{\mu}_k^s - \mu_l\|^2\right],$$
$$(\lambda^s)^{-1} = \frac{1}{V_s} \sum_{v,k} \tilde{\gamma}_{v,k}^s \left[\left(\tilde{\sigma}_k^s\right)^{-1} + \frac{1}{N}\|y_v^s - \tilde{\mu}_k^s\|^2\right], \tag{19}$$

that again look quite similar to the conventional clustering update rules accompanied with the correction terms involving $\tilde{\sigma}$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

References

1. Brett M, et al. The problem of functional localization in the human brain. Nat. Rev. Neurosci 2002;3:243–249. [PubMed: 11994756]
2. Friston, KJ., et al., editors. Statistical Parametric Mapping. Academic Press; London: 2007.
3. Gee, JC., et al. Proc. SPIE Med. Imaging. Vol. 3034. 1997. Effect of spatial normalization on analysis of functional data.; p. 550-560.
4. Thirion B, et al. Dealing with the shortcomings of spatial normalization: multisubject parcellation of fMRI datasets. Hum. Brain Mapp 2006;27:678–693. [PubMed: 16281292]
5. Thirion B, et al. Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. NeuroImage 2007;35:105–120. [PubMed: 17239619]
6. Thirion B, et al. Structural analysis of fMRI data revisited: improving the sensitivity and reliability of fMRI group studies. TMI 2007;26:1256–1269.
7. Golland, P., et al. Detection of spatial activation patterns as unsupervised segmentation of fMRI data.. In: Ayache, N.; Ourselin, S.; Maeder, A., editors. MICCAI 2007, Part I. LNCS. Vol. 4791. Springer; Heidelberg: 2007. p. 110-118.
8. McKeown JM, et al. Analysis of fMRI data by blind separation into independent spatial components. Hum. Brain Mapp 1998;10:160–178. [PubMed: 9673671]
9. Beckmann CF, Smith SM. Tensorial extensions of independent component analysis for group fMRI data analysis. NeuroImage 2005;25(1):294–311. [PubMed: 15734364]
10. Goutte C, et al. On clustering fMRI time series. NeuroImage 1999;9:298–310. [PubMed: 10075900]
11. Thirion, B.; Faugeras, O. Feature detection in fMRI data: the information bottleneck approach.. In: Ellis, RE.; Peters, TM., editors. MICCAI 2003. LNCS. Vol. 2879. Springer; Heidelberg: 2003. p. 83-91.
12. Lashkari, D., et al. Discovering structure in the space of activation profiles in fMRI.. In: Metaxas, D.; Axel, L.; Fichtinger, G.; Székely, G., editors. MICCAI 2008, Part I. LNCS. Vol. 5241. Springer; Heidelberg: 2008. p. 1015-1024.
13. Penny, WD.; Holmes, A. Random effects analysis.. In: Frackowiak, RSJ.; Friston, KJ.; Frith, CD., editors. Human brain function II. Elsevier; Oxford: 2003.
14. Mardia KV. Statistics of directional data. J. R. Statist. Soc. Series B 1975;37:349–393.
15. Jordan MI, et al. An introduction to variational methods for graphical models. Mach. Learn 1999;37:183–233.
16. Blei DM, et al. Latent Dirichlet allocation. J. Mach. Learn. Res 2003;3:993–1022.
17. Banerjee A, et al. Clustering on the unit hypersphere using von Mises-Fisher distribution. J. Mach. Learn. Res 2005;6:1345–1382.
18. Kanwisher, NG. The ventral visual object pathway in humans: evidence from fMRI.. In: Chalupa, L.; Werner, J., editors. The Visual Neurosciences. MIT Press; Cambridge: 2003.
19. http://surfer.nmr.mgh.harvard.edu/fswiki/FsFast

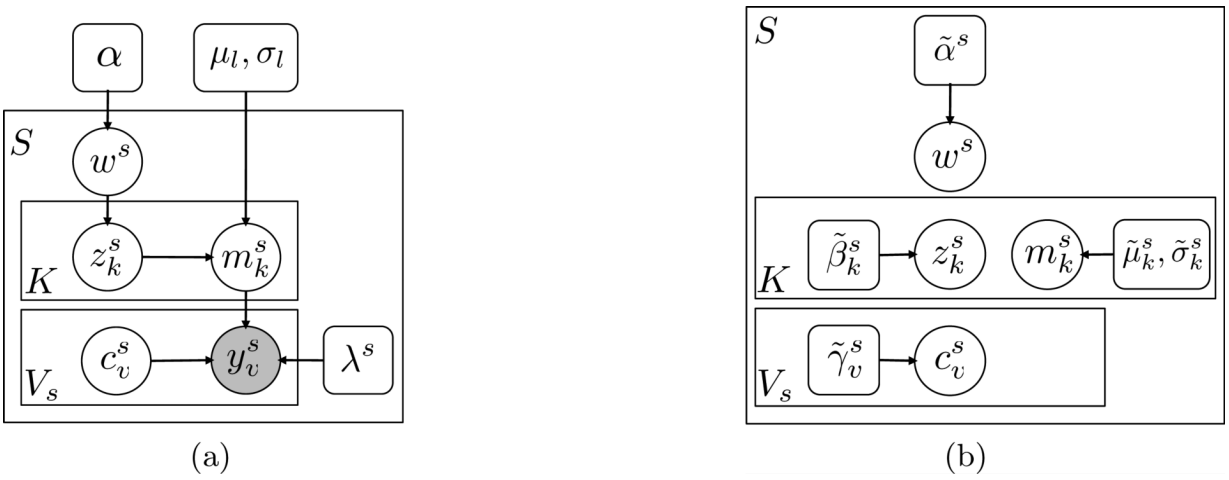

Fig. 1.

Graphical model showing the structure of (a) the model $p(\boldsymbol{y}, \boldsymbol{h}; \boldsymbol{\theta})$, and (b) distribution $q\left(\boldsymbol{h}; \tilde{\theta}\right)$ that approximates the posterior $p(\boldsymbol{h}|\boldsymbol{y}; \boldsymbol{\theta})$ over the hidden variables. Parameters and random variables are illustrated by squares and circles, respectively.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

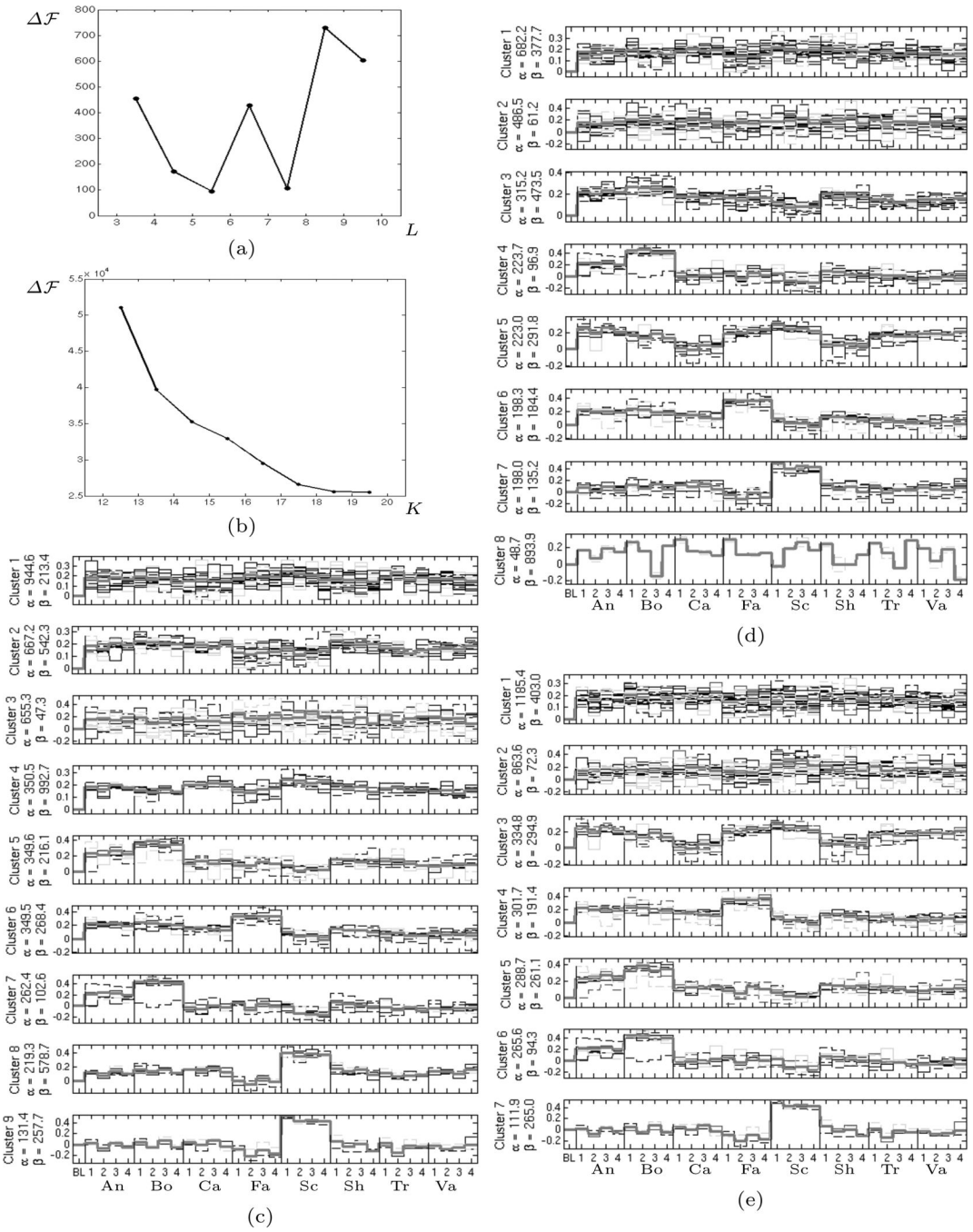

Fig. 2.

Improvement in the final value of free energy achieved by incrementally increasing (a) L for $K = 15$, and (b) K for $L = 8$. Each data point in these plots shows the change in $\mathcal{F}$ between the two neighboring numbers of clusters or parcels. (c)-(e) Components of the individual and population-level cluster centers for (K, L)=(9, 15), (8, 16), and (7, 15), respectively. For each cluster center μ_l (thick red line), the parcel centers $\tilde{\mu}_k^s$ assigned to this cluster based on weights $\tilde{\beta}_{k,l}^s$ are shown for all the 6 subjects (thin lines). The labels denote the baseline (BL) and the

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

four coefficients for each category of images: Animals, Bodies, Cars, Faces, Scenes, Shoes, Trees, Vases.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

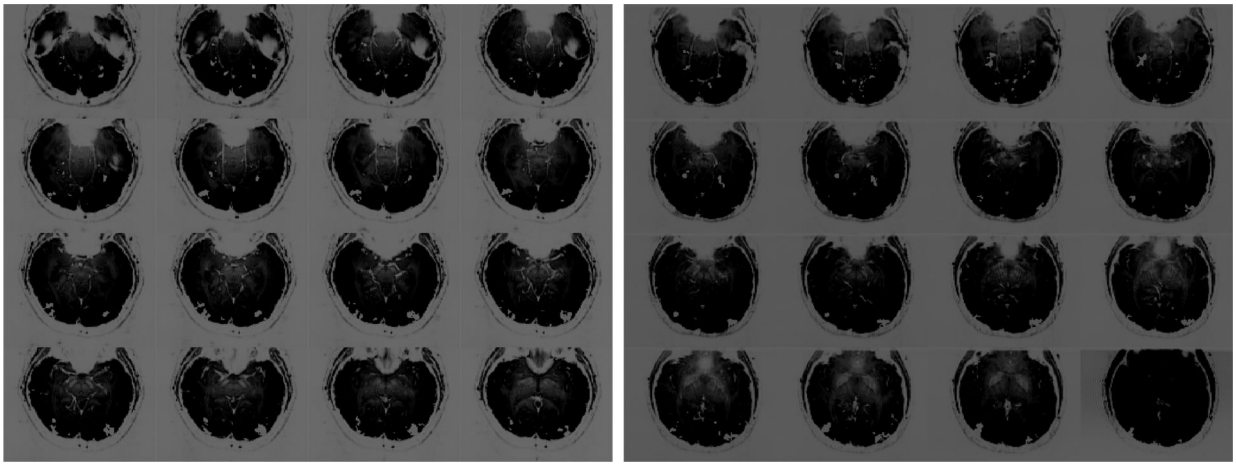

Fig. 3.
Spatial maps of the parcels assigned to the scene selective cluster 7 in Fig. 2(d) in two different subjects

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript