NeuroImage 50 (2010) 1085–1098

Contents lists available at ScienceDirect

NeuroImage

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

Discovering structure in the space of fMRI selectivity profiles

Danial Lashkari [a,*], Ed Vul [b], Nancy Kanwisher [b], Polina Golland [a]

[a] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] Brain and Cognitive Science Department, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 16 September 2009
Revised 21 December 2009
Accepted 23 December 2009
Available online 4 January 2010

Keywords:
fMRI
Clustering
High level vision
Category selectivity

ABSTRACT

We present a method for discovering patterns of selectivity in fMRI data for experiments with multiple stimuli/tasks. We introduce a representation of the data as profiles of selectivity using linear regression estimates, and employ mixture model density estimation to identify functional systems with distinct types of selectivity. The method characterizes these systems by their selectivity patterns and spatial maps, both estimated simultaneously via the EM algorithm. We demonstrate a corresponding method for group analysis that avoids the need for spatial correspondence among subjects. Consistency of the selectivity profiles across subjects provides a way to assess the validity of the discovered systems. We validate this model in the context of category selectivity in visual cortex, demonstrating good agreement with the findings based on prior hypothesis-driven methods.

Introduction

Standard fMRI experiments investigate the functional organization of the brain by contrasting the response to two or more sets of stimuli or tasks that are hypothesized to be treated differently by the brain. An *activation map* is generated by statistically comparing the fMRI response of each voxel to one set of tasks or stimuli versus another. The consistency of these activation maps across subjects is commonly evaluated by aligning the brain data across multiple subjects in a common anatomical space using spatial normalization. Using such voxel-wise correspondence across subjects, statistical analyses can test whether each voxel produces a higher response in one condition than another, consistently across subjects. In an alternative *region of interest* (ROI) method, discrete activation foci are functionally identified within each individual subject, and the responses of a given ROI to new conditions are then statistically compared across subjects.

These standard practices have generated a wealth of knowledge about the functional organization of the human brain, most of it unknown just 20 years ago. However the standard methods are also subject to two important limitations: (i) they can only test hypotheses generated by the experimenter; they cannot discover new structure in the fMRI response, and (ii) they assume some consistency across subjects in the spatial pattern of activation across the brain. Here, we introduce a new method that avoids both of these limitations, enabling us to discover patterns of functional response that are found robustly across subjects. These patterns do not have to be hypothesized a priori and do not have to correspond to voxels that exhibit spatial contiguity or spatial consistency across subjects.

Our exploratory approach introduces the concept of a *selectivity profile*, which is a simple characterization of the function of a voxel in terms of its response to each of the different experimental conditions. The experimental conditions in this approach can number in the tens or even hundreds, instead of the two to eight conditions used in most imaging studies. We aim to discover selectivity profiles that best explain the entire data set.

In the conventional univariate approach, the response of the entire population of voxels is not considered as a whole; tests are performed separately on single voxels to examine the significance of *a priori* hypothesized activations. In contrast, we devise a model that explains the selectivity profiles of all voxels by grouping them into a number of *systems* (clusters), each with a distinct, representative selectivity profile across the stimuli/tasks in the experiment. Once a small set of robust systems is discovered, we can map the location of the voxels that correspond to each system to find out where they are in the brain, if they are spatially contiguous, and if they are in similar locations across subjects. With our new method, the answers to these questions are now genuine discoveries, not assumptions built into the method.

Our method offers an additional advantage for group analysis. In the conventional hypothesis-driven analysis, in order to analyze a cohort of subjects, we need to first normalize different subjects into a common anatomical space. Since brain structure is highly variable across subjects, establishing accurate correspondences among anatomical images of different subjects is intrinsically challenging (Gee et al., 1997; Thirion et al., 2006; Thirion et al., 2007a; 2007b). In addition to this anatomical variability, functional properties of the same anatomical structures are likely to vary somewhat across

* Corresponding author. Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
E-mail addresses: danial@mit.edu (D. Lashkari), evul@mit.edu (E. Vul), ngk@mit.edu (N. Kanwisher), polina@csail.mit.edu (P. Golland).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.12.106

subjects (Brett et al., 2002). This variability presents a fundamental obstacle for anatomically-constrained characterization of functional systems.

Since we are primarily interested in the systems with selectivity profiles that are shared across subjects, the space of these profiles for a certain experimental setup can act as a common space for representing data from different subjects. Hence, our model does not require spatial information, and we can analyze group data from different runs and subjects without a need for spatial normalization. Furthermore, rather than relying on spatial consistency to establish the validity of selectivity patterns, we employ *functional* consistency defined as the robustness of the estimated profiles across subjects. Because our representation of the data relies solely on the functional response, and the basic analysis knows nothing about the location of each voxel, the resulting systems we discover are not constrained to be spatially clustered together or in similar locations across subjects.

Several exploratory, unsupervised learning methods have been previously applied to the analysis of fMRI data. Some methods consider the spatial patterns of response across many voxels; for instance, they may apply an algorithm such as agglomerative clustering to these spatial patterns to infer the hierarchical grouping of stimuli (Kriegeskorte et al., 2008). In contrast, we seek groups of voxels that respond similarly across a large set of images and to characterize the nature of their response.

Most exploratory methods that attempt to partition the set of voxels consider raw fMRI time courses and use clustering (Baumgartner et al., 1997; Baumgartner et al., 1998; Moser et al., 1997; Golay et al., 1998; Filzmoser et al., 1999; Fadili et al., 2000; Golland et al., 2007) or Independent Component Analysis (ICA) (McKeown et al., 1998; Beckmann and Smith, 2004, 2005; Calhoun et al., 2001a, 2001b) to estimate a decomposition of the data into a set of distinct time courses of interest and their localization maps. However, these methods do not readily express the relationship between the discovered time courses and the functional response of voxels to the experimental conditions. Some variants employ information from the experimental setup to define a measure of similarity between voxels, effectively projecting the original high-dimensional time courses onto a low dimensional feature space, followed by clustering in the new space (Goutte et al., 1999; Goutte et al., 2001; Thirion and Faugeras, 2003). Nevertheless, these methods mainly focus on the spatial maps of the clusters corresponding to the activation of interest as the main result of the analysis. Few exploratory or multivariate methods systematically address the issue of group analysis (Calhoun et al., 2001b).

In this paper, we employ the studies of category selectivity in the visual pathway as a concrete example for the applications of our method in fMRI studies with a rich space of experimental conditions. Functional MRI studies of vision have provided a great example for the success of fMRI in revealing structure in brain's functional organization (Grill-Spector and Malach, 2004). Using the conventional hypothesis-driven approach, fMRI data have identified a handful of regions with specific category selectivity in the visual cortex (Kanwisher, 2003). For instance, the fusiform face area and occipital face area (FFA and OFA) are associated with face selectivity (McCarthy et al., 1997; Kanwisher et al., 1997; Kanwisher and Yovel, 2006; Rossion et al., 2003) while the parahippocampal place area (PPA) (Epstein and Kanwisher, 1998; Burgess et al., 1999; Aguirre et al., 1998) and the extrastriate body area (EBA) (Downing et al., 2001; Schwarzlose et al., 2005; Peelen and Downing, 2007) exhibit high selectivity for places and body parts, respectively. In addition, an object-selective region in the lateral occipital complex (Malach et al., 1995) and a small area in the left fusiform gyrus with selectivity to letter strings and visually-presented words (Baker et al., 2007) have been identified and further characterized. However, the collection of all currently known category selective areas constitutes a small part of the visual pathway and accounts for a limited set of categories in the visual world. Further studies to find other selective areas have generally failed, leaving many questions unanswered (Downing et al., 2006).

We face several methodological challenges in proceeding with the present hypothesis-driven approach. With the increasing number of image categories included in the experiments, it becomes more challenging to explore the entire set of possible patterns of selectivity by only comparing voxel responses to two conditions at a time. In principle, for any candidate category, we have to search for a brain region that shows significant activation in pairwise contrasts with all the other categories in the experiment. Moreover, we should also consider possible meta-categories (categories comprised of multiple other categories) in the experiment that might form natural classes of selectivity for a brain system. For instance, if we have images of both human faces and human bodies in the experiment, we can also consider the set of all these images as one candidate category. Finally, the conventional approach uses spatial contiguity of the regions and their spatial consistency across subjects as the only method to evaluate whether a system is truly selective (Kanwisher et al., 1997; Spiridon et al., 2006). Hence, anatomical variability of functional systems in the brain fundamentally restricts the power of this method.

As an illustration of how our method can address the basic limitations of the conventional hypothesis-driven method, we apply it to the data from an experiment investigating category selectivity in the visual cortex. Our results replicate prior findings obtained via numerous hypothesis-driven fMRI studies. We robustly discover the known selectivity patterns, further opening up the possibility of studies on richer sets of stimuli.

Methods

In this section, we define the main elements of our approach in three steps. First, we introduce the space of selectivity profiles and explain how this representation enables discovery of the patterns of selectivity in the brain systems of interest. Second, we formulate our model for the analysis of the data, represented in the space of selectivity profiles, and derive an algorithm for estimating the model parameters from fMRI data. Third, we present our approach to validating the results based on our definition of cross-subject and within-subject consistency measures. In the last part, we discuss alternative validation procedures.

Space of selectivity profiles

The goal of our method is to discover the types of response specificity that appear across multiple voxels in our data. Hence, we choose to represent an fMRI time course by a profile that characterizes the selectivity of the voxel response to the set of all experimental conditions. The notion of selectivity focuses on the *relative* response; for instance, the rough definition of a selective system in high level vision states that a selective voxel's response to the preferred stimulus is at least twice as high as its response to other stimuli (Op de Beeck et al., 2008). This definition involves only a ratio of the responses and is independent of their overall magnitude. Our profile representation, therefore, aims to capture this relative response of the voxels based on their observed BOLD time course.

It is customary to use regression to estimate the contributions of different experimental conditions to the BOLD signal. In this set up, we can represent the BOLD response $x_v \in IR^T$ of voxel v at the T time points as:

$$x_v = H\alpha_v + G\beta_v + \varepsilon_v, \tag{1}$$

where the columns of matrices H and G are the temporal regressors corresponding to the protocol-independent nuisance factors and the

D experimental conditions, respectively. Assuming white temporal noise, $\varepsilon_v \sim \mathcal{N}(0, \sigma_v^2 I)$, the least square solution yields the estimates of the regression coefficients β_v and $\hat{\alpha}_v$:

$$[\hat{\alpha}_v \hat{\beta}_v] = \left(A^t A\right)^{-1} A^t x_v, \tag{2}$$

where $A = [H\ G]$. Component i of the estimated vector β_v is commonly interpreted as a measure of the response of voxel v to stimulus i. The above model is usually used for assessing the statistical significance of different hypotheses about the response in the framework of the General Linear Models (GLM) (Friston et al., 1994). More accurate models of fMRI signal also account for autocorrelations present in the covariance structure of the temporal noise ε_v (Aguirre et al., 1997; Woolrich et al., 2001), but the above simple model is adequate for the demonstration of our method.

We define the *voxel selectivity profile* to be a vector containing the estimated regression coefficients for the experimental conditions, normalized to unit magnitude, that is,

$$y_v = \frac{\hat{\beta}_v}{||\hat{\beta}_v||}, \tag{3}$$

where $||a|| = \sqrt{\langle a, a\rangle}$ with $\langle\ ,\ \rangle$ denoting the inner product. Selectivity profiles lie on a hyper-sphere S^{D-1} and imply a pattern of selectivity to the D experimental conditions defined by a direction in the corresponding D-dimensional space. Normalization removes the contribution of the overall magnitude of response and presents the estimated response as a ratio with respect to this overall response. Furthermore, it is well-known that the magnitude of overall BOLD response of the voxel is mainly a byproduct of irrelevant variables such as distance from major vessels or general response to the type of stimuli used in the experiment (Friston et al., 2007). This provides another justification for the normalization of response vectors, in addition to our interest in representing selectivity as a relative measure of response.

Fig. 1A illustrates the population of unnormalized estimated vectors of regression coefficients β_v for all the voxels identified by the conventional hypothesis-driven analysis as selective for one of three different conditions. The differences between voxels with different types of selectivity are not well expressed in this representation; there is no apparent separation between different groups of voxels. We also note that there is an evident overlap between the sets of voxels assigned to these different patterns of selectivity. The standard analysis in this case uses a contrast comparing each of the three conditions of interest with a fourth experimental condition; therefore, it is possible for a voxel to appear selective for all three of these contrasts. In order to explain the selectivity of such a voxel, we can define a novel type of selectivity towards a meta-category which is composed of the three categories represented by these contrasts. The same argument can be applied to any combinations of the categories presented in the experiment to form various, new candidates as possible types of selectivity.

Fig. 1B shows the selectivity profiles y_v formed for the same data set. We observe that the voxels associated with different types of activation become more separated, exhibiting an arrangement that is similar to a clustering structure. Furthermore, it is easy to see that the set of voxels shared among all three patterns of selectivity has a distinct structure of its own, mainly concentrated around a direction close to $[111]^t / \sqrt{3}$ on the sphere. We interpret the center of a cluster of selectivity profiles as a representative for the type of selectivity shared among the neighboring profiles on the sphere.

Although the clusters of the profiles are not well separated, the arrangement of concentrations of profiles on the sphere can carry important information about the types of selectivity more heavily represented in the data. This information becomes more interesting as the number of dimensions (experimental conditions) grows and the overall density of profiles on the sphere decreases. This motivates us to consider application of mixture model density estimation, the probabilistic modeling formulation of clustering (McLachlan and Peel, 2000), to the set of selectivity profiles. Each component in the mixture model represents a cluster of voxels, i.e., a functional system, concentrated around a central direction on the sphere. The corresponding cluster center, which we call *system selectivity profile*, specifies that system's type of selectivity.

Model

Let $\{y_v\}_{v=1}^{V}$ be a set of selectivity profiles of V brain voxels. We assume the vectors are generated *i.i.d.* by a mixture distribution

$$p\left(y; \{q_k, m_k\}_{k=1}^{K}, \lambda\right) = \sum_{k=1}^{K} q_k f(y; m_k, \lambda), \tag{4}$$

where $\{q_k\}_{k=1}^{K}$ are the weights of K components and $f(\cdot, m, \lambda)$ is the likelihood of the data parameterized by m and λ. We assume that the likelihood model describes simple directional distribution on the hyper-sphere and choose the von Mises–Fisher distribution (Mardia, 1975) for the mixture components:

$$f(y; m, \lambda) = C_D(\lambda) e^{\lambda \langle y, m\rangle}, \tag{5}$$

where inner product corresponds to the correlation of the two vectors on the sphere. Note that this model is in agreement with the notion that on a hyper-sphere, correlation is the natural measure of similarity between two vectors. The distribution is an exponential function of the correlation between the vector y (voxel selectivity profile) and the mean direction m (system selectivity profile). The normalizing constant $C_D(\lambda)$ is defined in terms of the γ-th order modified Bessel function of the first kind I_γ:

$$C_D(\lambda) = \frac{\lambda^{D/2-1}}{(2\pi)^{D/2} I_{D/2-1}(\lambda)}. \tag{6}$$

The concentration parameter λ controls the concentration of the distribution around the mean direction m similar to the reciprocal of variance for Gaussian models. In general, mixture components can have distinct concentration parameters but in this work, we use the same parameter for all the clusters to ensure a more robust estimation. This model has been previously employed in the context of clustering (Banerjee et al., 2006).

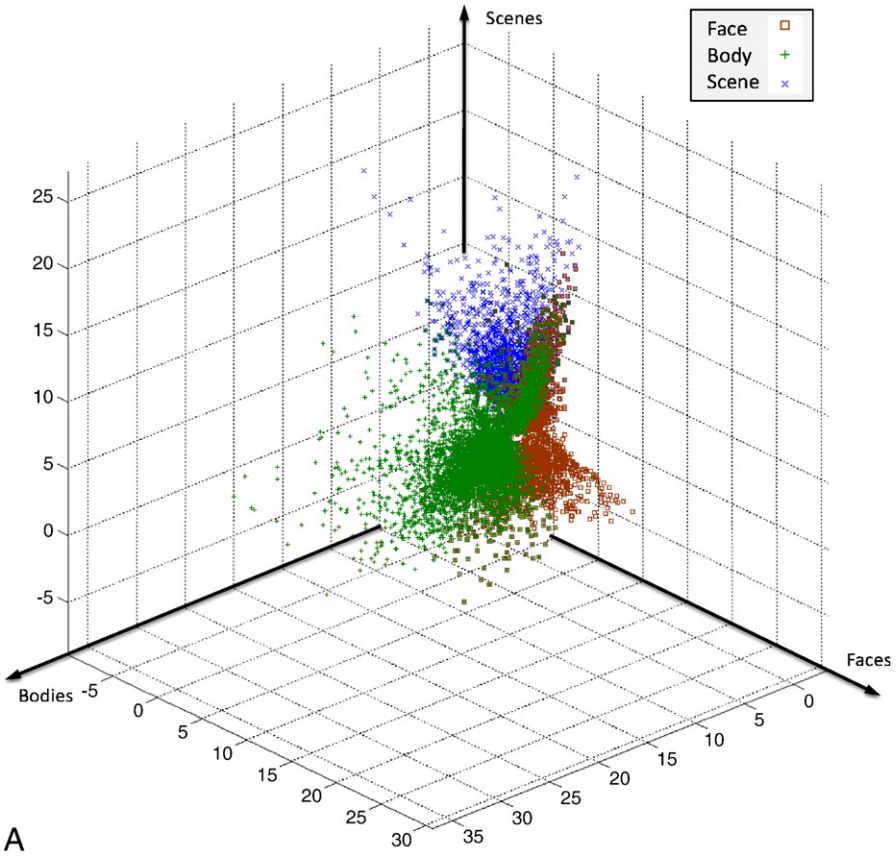
We formulate our problem as a maximum likelihood estimation:

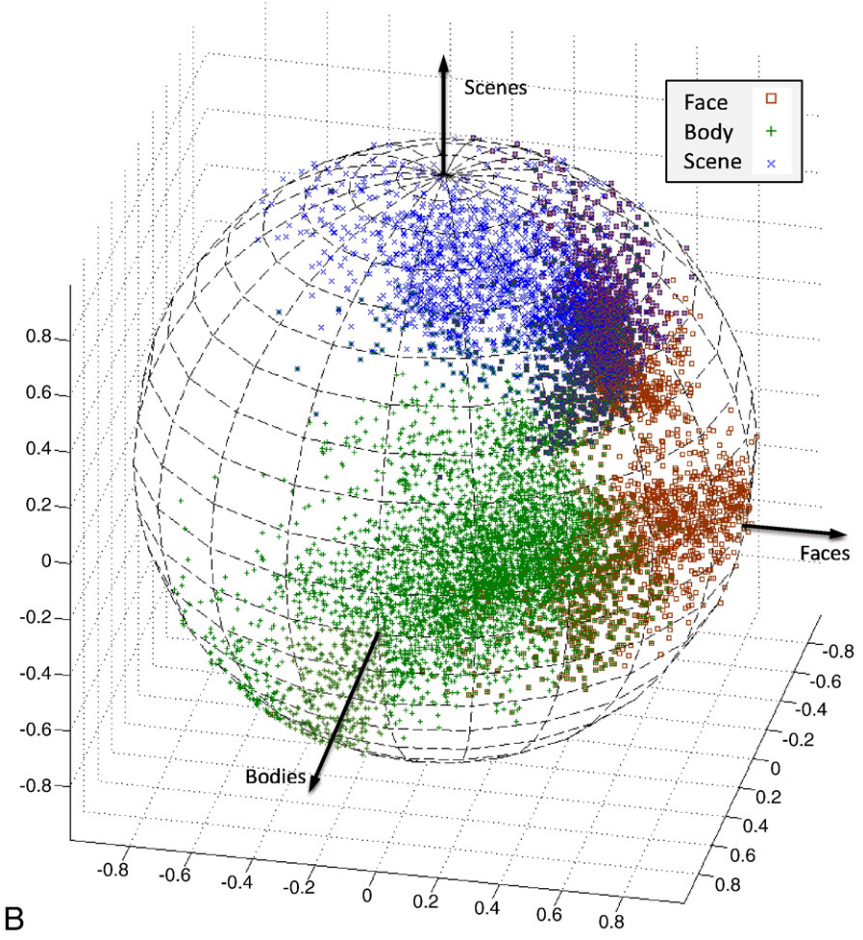
$$\left(\left\{q_k^*, m_k^*\right\}_{k=1}^{K}, \lambda^*\right) = \underset{\{q_k, m_k\}_{k=1}^{K}, \lambda}{argmax} \sum_{v=1}^{V} log\, p\left(y_v; \{q_k, m_k\}_{k=1}^{K}, \lambda\right). \tag{7}$$

Employing the Expectation–Maximization (EM) algorithm (Dempster et al., 1977) to solve the problem involves adding membership variables $p(k|y_v)$, for $k = 1, \ldots, K$, that describe the posterior probability that voxel v is associated with the mixture component k. The details of the EM derivation are presented in Appendix A.

Starting with initial values $\{q_k^{(0)}, m_k^{(0)}\}_{k=1}^{K}$ and $\lambda^{(0)}$ for the model parameters, we iteratively compute the posterior assignment probabilities $p(k|y_v)$ and then update the parameters $\{q_k, m_k\}_{k=1}^{K}$ and λ. In the E-step, we fix the model parameters and update the system memberships:

$$p^{(t)}(k|y_v) = \frac{q_k^{(t)} e^{\lambda^{(t)} \langle y_v, m_k^{(t)}\rangle}}{\sum_{k'=1}^{K} q_{k'}^{(t)} e^{\lambda^{(t)} \langle y_v, m_{k'}^{(t)}\rangle}}. \tag{8}$$


Scenes
Face
Body
Scene
Faces
Bodies
A


Scenes
Face
Body
Scene
Faces
Bodies
B

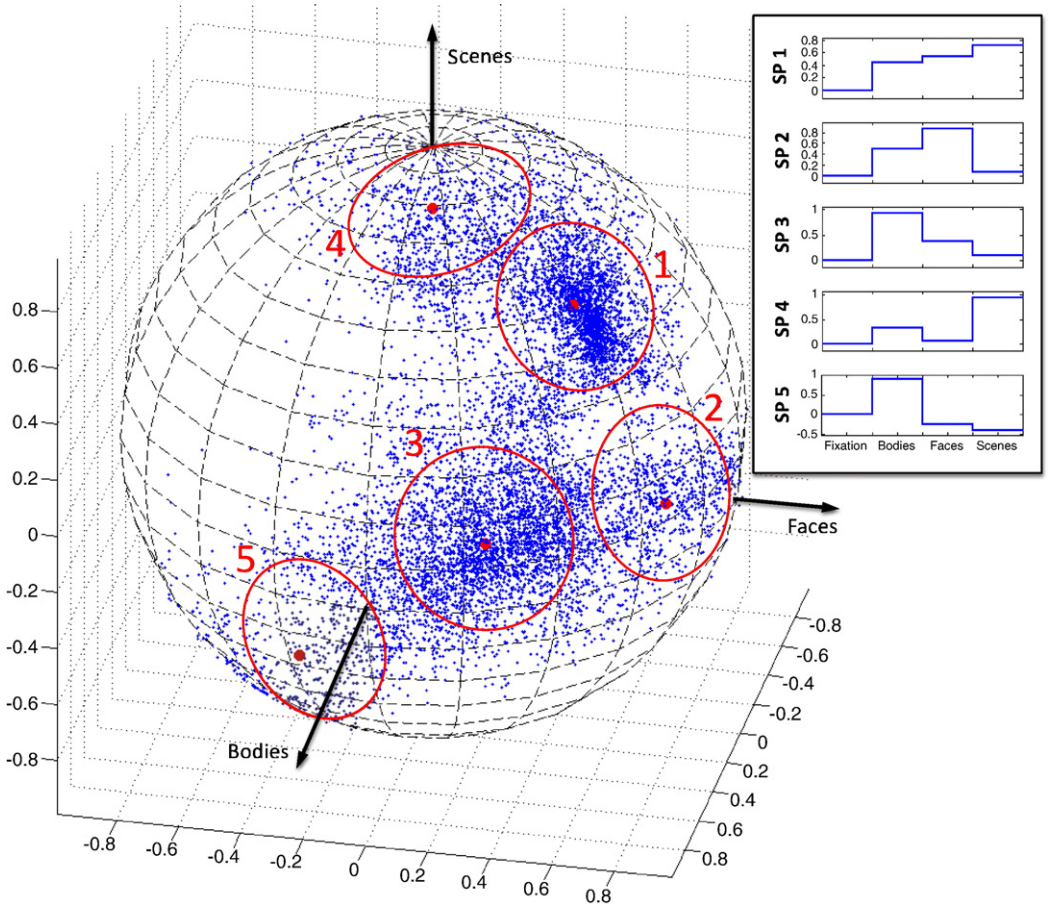


Fig. 2. The results of mixture model density estimation with 5 components for the set of selectivity profiles in Fig. 1B. The resulting system selectivity profiles (cluster centers) are denoted by the red dots; circles around them indicate the size of the corresponding clusters. The box shows an alternative presentation of the selectivity profiles where the values of their components are shown along with zero for fixation. Since this format allows presentation of the selectivity profiles in general cases with $D > 3$, we adopt this way of illustration throughout the paper. The first selectivity profile, whose cluster includes most of the voxels in the overlapping region, does not show a differential response to our three categories of interest. Selectivity profiles 2, 3, and 4 correspond to the three original types of activation preferring faces, bodies, and scenes, respectively. Selectivity profile 5 shows exclusive selectivity for bodies along with a slightly negative response to other categories.

In the M-step, we update the model parameters:

$$q_k^{(t+1)} = \frac{1}{V}\sum_{v=1}^{V} p^{(t)}(k|y_v), \tag{9}$$

$$m_k^{(t+1)} = \frac{\sum_{v=1}^{V} y_v p^{(t)}(k|y_v)}{\left\|\sum_{v=1}^{V} y_v p^{(t)}(k|y_v)\right\|}. \tag{10}$$

After computing the updated cluster centers $m_k^{(t+1)}$, the new concentration parameter $\lambda^{(t+1)}$ is found by solving the nonlinear equation

$$A_D\left(\lambda^{(t+1)}\right) = r^{(t+1)} \tag{11}$$

for positive values of $\lambda^{(t+1)}$, where

$$r^{(t+1)} = \frac{1}{V}\sum_{k=1}^{K}\sum_{v=1}^{V} p^{(t)}(k|y_v)\langle m_k^{(t+1)}, y_v\rangle \tag{12}$$

and the function $A_D(\cdot)$ is defined as

$$A_D(\lambda) = \frac{I_{D/2}(\lambda)}{I_{D/2-1}(\lambda)}. \tag{13}$$

Our algorithm for solving this equation is presented in Appendix B. Iterating the set of E-step and M-step updates until convergence, we find K system selectivity profiles m_k and a set of soft assignments $p(k|y_v)$ for $k = 1, \ldots, K$. The assignments $p(k|y_v)$, when projected to the anatomical locations of voxels, define the spatial maps of the discovered systems.

Fig. 2 illustrates 5 systems and the corresponding profiles of selectivity found by this algorithm for the population of voxels shown in Fig. 1B. As expected, the analysis identifies clusters of voxels exclusively selective for one of the three conditions, but also finds a cluster selective for all three conditions along with a group of body selective voxels that show inhibition towards other categories. More complex profiles of selectivity such as the two latter cases cannot be easily detected with the conventional method.

Cross-subject consistency analysis

Consider a group study where a group of subjects take part in an fMRI experiment. We denote a voxel in an experiment with S subjects by y_v^s, where $s \in \{1, \ldots, S\}$ is the subject index, and v is the voxel index as before. We aim to discover the brain systems with distinct profiles of selectivity that are shared among all subjects. Let us assume that

Fig. 1. An example of voxel selectivity profiles in the context of a study of visual category selectivity. The block design experiment included several categories of visual stimuli such as faces, bodies, scenes, and objects, defined as different experimental conditions. (A) Vectors of estimated regression coefficients $\beta = [\beta_{\text{Faces}}, \beta_{\text{Bodies}}, \beta_{\text{Scenes}}]^t$ for the voxels detected as selective to bodies, faces, and scenes in one subject. As is common in the field, the conventional method detects these voxels by performing significance tests comparing voxel's response to the category of interest and its response to objects. (B) The corresponding selectivity profiles y formed for the same group of voxels.

two selectivity profiles y_v^s and $y_{v'}^{s'}$, corresponding to voxel v of subject s and voxel v' of subject s', belong to the same selective system in the two brains. The overall magnitude of response of these two voxels can be different but the two profile vectors have to still reflect the corresponding type of selectivity. Therefore, they should resemble each other as well as the selectivity profile of the corresponding system. This suggests that we can fuse data from different subjects and cluster them all together in order to improve the estimates of system selectivity profiles. This approach can be thought of as a simple model that ignores possible small variability in subject-specific selectivity profiles of the same system, similar to the way that fixed effect analysis simplifies the more elaborate hierarchical model of random effect analysis in the hypothesis-driven framework (Penny and Holmes, 2003). At this stage, we choose to work with this simpler model and defer the development of a corresponding hierarchical model to future work.

Based on the above argument, if the set of vectors $\{m_k\}_{k=1}^K$ describes all relevant selectivity profiles in the brain system of interest, each voxel y_v^s can be thought of as an independent sample from the distribution in Eq. (4). Thus, we combine the data from several subjects to form the *group data*, i.e., $\left\{\{y_v^s\}_{v=1}^{V_s}\right\}_{s=1}^{S}$, to perform our analysis across subjects. Applying our algorithm to the group data, the resulting set of assignments $\{p(k|y_v^s)\}_{v=1}^{V_s}$ defines the spatial map of system k in subject s.

In conventional group data analysis, spatial consistency of the activation maps across subjects provides a measure for the evaluation of the results. In our method, we focus on the functional consistency of the discovered system selectivity profiles. To quantify this consistency, we define a *consistency score* (cs) for each selectivity profile found in a group analysis. Let $\left\{\{y_v^s\}_{v=1}^{V_s}\right\}_{s=1}^{S}$ be the group data including voxel profiles from S different subjects, K be the number of desired systems, and $\{m_k^G\}_{k=1}^K$ be the final set of system selectivity profiles found by the algorithm in the group data. We also apply the algorithm to the S individual subject data sets $\{y_v^s\}_{v=1}^{V_s}$ separately to find their corresponding S sets of subject-specific systems $\{m_k^s\}_{k=1}^K$. We can then match the selectivity profile of each group system to its most similar system profile in each of the S individual data sets.

Matching selectivity profiles across subjects

The matching between the group and individual selectivity profiles is equivalent to finding S one-to-one functions ω_s:{1, …, K}→{1, …, K} which assign system profile $m^s_{\omega_s(k)}$ in subject s to the group system profile m_k^G. We select the function ω_s such that it maximizes the overall similarity between the matched selectivity profiles:

$$\omega_s^*(\cdot) = \underset{\omega(\cdot)}{argmax} \sum_{k=1}^{K} \rho\left(m_k^G, m^s_{\omega(k)}\right). \tag{14}$$

Here, $\rho(\cdot,\cdot)$ denotes the correlation coefficient between two vectors. The maximization in Eq. (14) is performed over all possible one-to-one functions ω. Finding this function is an instance of graph matching problems for a bipartite graph (Diestel, 2005). The graph is composed of two sets of nodes, corresponding to the group and the individual system profiles, and the weights of the edges between the nodes are defined by the correlation coefficients. We employ the well-known Hungarian algorithm (Kuhn, 1955) to solve this problem for each subject.[1]

[1] We used the open source Matlab implementation of the Hungarian algorithm available at http://www.mathworks.com/matlabcentral/leexchange/11609.

Having matched each group system with a distinct system within each individual subject result, we compute the consistency score cs_k for group system k as the average correlation of its selectivity profile with the corresponding subject-specific system profiles:

$$cs_k = \frac{1}{S}\sum_{s=1}^{S} \rho\left(m_k^G, m^s_{\omega_s^*(k)}\right). \tag{15}$$

Consistency score values measure how closely a particular type of selectivity repeats across subjects. Clearly, cs = 1 is the most consistent case where the corresponding profile identically appears in all subjects. Because of the similarity-maximizing matching performed in the process of computing the scores, even a random data set would yield non-zero consistency score values. We employ permutation testing to establish the null hypothesis distribution for the consistency score.

Permutation test for the consistency scores

To construct the baseline distribution for the consistency scores under the null hypothesis, we make random modifications to the data in such a way that the correspondence between the components of the selectivity profiles and the experimental conditions is removed. Specifically, we randomize the condition labels before the regression step so that the individual regression coefficients do not correspond to any non-random distinctions in the task. More formally, we implement such a randomization in the linear analysis stage in Eq. (1). Each temporal block in the experiment has a category label that determines its corresponding regressor in the design matrix G. We randomly shuffle these labels and, as a result, the regressors in the design matrix include blocks of images from random categories. The resulting estimated regression coefficients do not correspond to any coherent set of stimuli. Applying our analysis to this modified data set still yields a set of group and individual system selectivity profiles and corresponding cs values. Since there is no real structure in the data, all cs values obtained in this manner can serve as samples from the desired null hypothesis.

We estimate the null distribution of the cs values by generating randomly shuffled data sets, finding selectivity profiles of the group systems, and treating the resulting consistency scores for the K selectivity profiles as samples from our null distribution. We evaluate statistical significance of the cs value of each system selectivity profile based on this null distribution. In practice, for up to 10,000 shuffled data sets, the consistency scores of most system selectivity profiles in the real data exceed *all* the cs values estimated from the shuffled data, implying the same empirical significance of $p = 10^{-4}$. To distinguish the significance of these different profiles through our p-value, we fit a Beta distribution to the null-hypothesis samples and compute the significance from the fitted distribution. Using a linear transformation to match the range $[-1, 1]$ of cs values to the support [0,1] of the Beta distribution, we obtain the pdf of the null distribution

$$f_{\text{Null}}(\text{cs}; a, b) = \frac{2}{B(a,b)}\left(\frac{1+\text{cs}}{2}\right)^{a-1}\left(\frac{1-\text{cs}}{2}\right)^{b-1}, \tag{16}$$

where $B(a,b)$ is the beta function. We find the maximum-likelihood pair (a,b) for the observed samples in the shuffled data set. We then characterize the significance of a selectivity profile with consistency score cs via its p-value, as inferred from the parametric fit to our simulated null-hypothesis distribution: $p = \int_{\text{cs}}^{1} f_{\text{Null}}(u; a, b)du$.

Other validation procedures

In the previous section, we presented our procedure for assessment of the cross-subject consistency of the discovered selectivity profiles. An alternative way to assess the resulting system selectivity

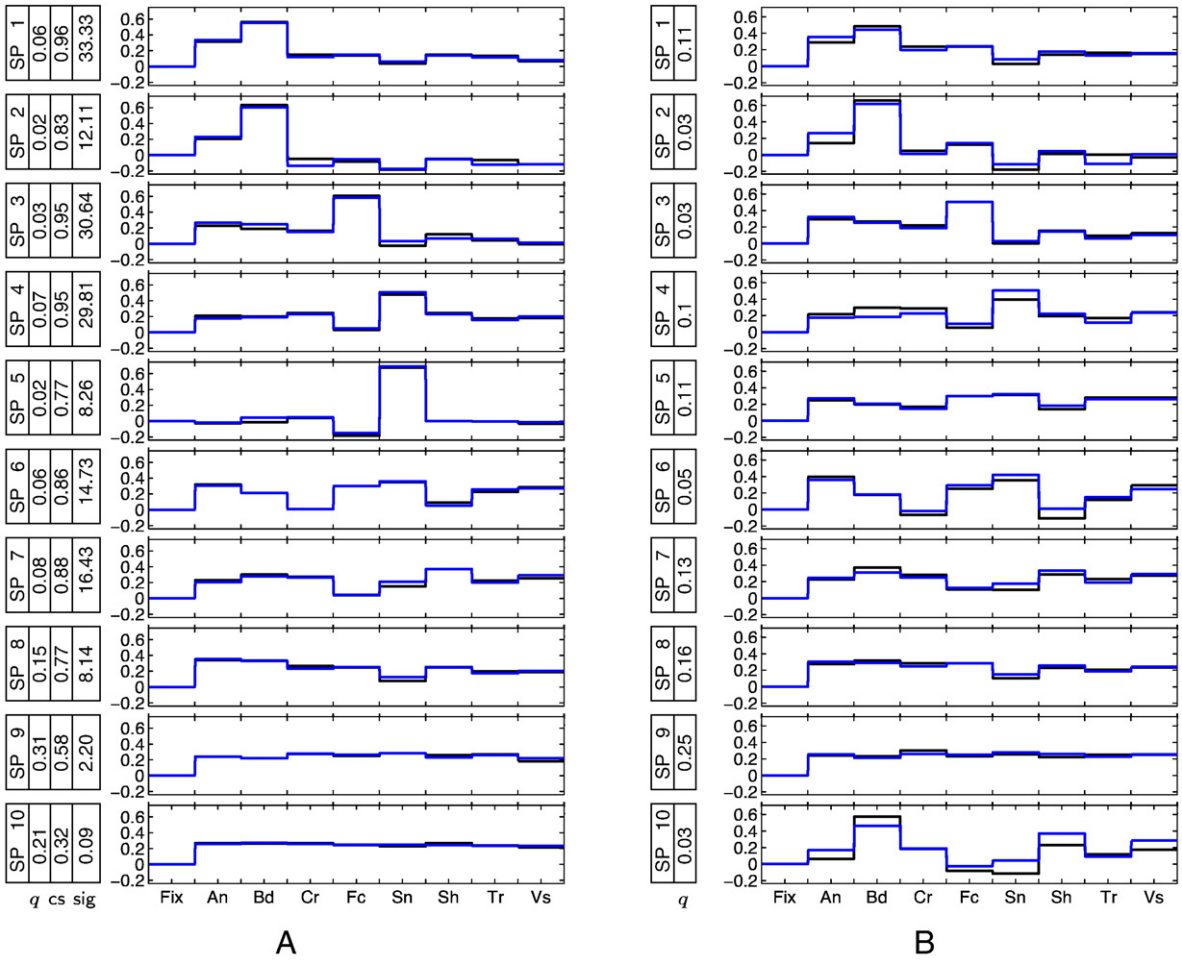


Fig. 3. (A) A set of 10 discovered group system selectivity profiles for the 16-Dimensional group data. The colors (black, blue) represent the two distinct components of the profiles corresponding to the same category. We added zero to each vector to represent Fixation. The weight q for each selectivity profile is also reported along with the consistency scores (cs) and the significance values found in the permutation test, sig $= -\log_{10} p$. (B) A set of individual system selectivity profiles in one of the 6 subjects ordered based on matching to the group profiles in (A).

profiles is to examine their consistency across repetitions of the same category in different blocks. If we define two groups of blocks that present stimuli from the same category as two distinct experimental conditions, we expect the corresponding components of a consistent system selectivity profile to be similar. We will employ this method for a qualitative study of consistency.

While our analysis is completely independent of the spatial locations associated with the selectivity profiles, we can still examine the spatial extent of the discovered systems as a way to validate the results. If the selectivity profile of a system matches a certain type of activation, i.e., demonstrates exclusive selectivity for an experimental condition, we can compare the map of its assignments with the localization map detected for that activation by the conventional method. We use this comparison to ensure that our method yields systems with spatial extents that correspond to the previously characterized selective areas in the brain.

Once we identify a system to be associated with a certain activation, we quantify the similarity between the spatial maps estimated by our method and that obtained via the standard hypothesis-driven method. We employ an asymmetric overlap measure between the spatial maps, equal to the ratio of the number of voxels in the overlapping region to the number of all voxels assigned to the system in our model. The asymmetry is included since, as we saw in the example in Space of Selectivity Profiles, being functionally more specific, our discovered systems are usually subsets of the localization maps found via the standard statistical test.

Results

We demonstrate our method on the data from a block design fMRI study of high level vision with 6 subjects. The images were acquired using a Siemens 3T scanner and a custom 32-channel coil (EPI, flip angle = 90°, TR = 2 s, in-plane resolution = 1.5 mm, slice thickness = 2 mm, 28 axial slices). The experimental protocol included 8 categories of images: Animals, Bodies, Cars, Faces, Scenes, Shoes, Trees, and Vases. For each image category, two different sets of images were presented in separate blocks. We used this setup to test that our algorithm successfully yields profiles with similar components for different images from the same category. Each block lasted 16 seconds and contained 16 images from one image set of one category. The blocks corresponding to different categories were presented in permuted fashion so that their order and temporal spacing was counter-balanced. With this design, the temporal noise structure is shared between the real data and the random permutations constructed by the procedure of Methods. For each subject, there were 16 to 29 runs of the experiment where each run contained one block from each category and three fixation blocks. We perform motion correction, spike detection, intensity normalization, and Gaussian smoothing with a kernel of 3-mm width using the standard package FsFast.[2]

[2] http://surfer.nmr.mgh.harvard.edu/fswiki/fsfast.

By modifying the condition-related part of the design matrix G in Eq. (1) and estimating the corresponding regression coefficients β, we created three different data sets for each subject:

- **8-Dimensional Data:** All blocks for one category were represented as a single experimental condition by one regressor and, accordingly, one regression coefficient. The selectivity profiles were composed of 8 components each representing one category.
- **16-Dimensional Data:** The blocks associated with different image sets were represented as distinct experimental conditions. Since we had two image sets for each category, the selectivity profiles had two components for each category.
- **32-Dimensional Data:** We split the blocks for each image set into two groups and estimated one coefficient for each split group. In this data set, the selectivity profiles were 32 dimensional and each category was represented by four components.

To discard the voxels with no visual activation, we formed contrasts comparing the response of voxels to each category versus fixation and applied the t-test to detect voxels that show significant levels of activation. The union of the detected voxels served as the mask of visually responsive voxels used in our experiment. Significance thresholds were chosen to $p=10^{-2}$, $p=10^{-4}$, and $p=10^{-6}$, for 32-Dimensional, 16-Dimensional, and 8-Dimensional data, respectively, so that the visually selective masks for different data sets are of comparable size. An alternative approach for selecting relevant voxels is to use an F-test considering all regressors corresponding to the visual stimuli (columns of matrix G in Eq. (1)). We observed empirically that the results presented here are fairly robust to the choice of the mask and other details of preprocessing.

Selectivity profiles

We apply the analysis to all three data sets. Fig. 3A and Fig. 4 show the resulting selectivity profiles of the group systems in the three data sets, where the number of clusters is $K=10$. We also report cluster weights q_k, consistency scores cs_k, and their corresponding significance values $sig=-\log_{10} p$. In all data sets, the most consistent profiles are selective of only one category, similar to the previously characterized selective systems in high level vision. Moreover, their peaks match with these known areas such that in each data set, there are selectivity profiles corresponding to EBA (body-selective), FFA (face-selective), and PPA (scene-selective). For instance, in Fig. 3A, selectivity profiles 1 and 2 show body selectivity, selectivity profile 3 is face selective, and selectivity profiles 4 and 5 are scene selective. Similar profiles appear in the case of 8-Dimensional and 32-Dimensional data as well. Comparing the selectivity profiles found for one of the individual 16-Dimensional data sets with those of the group data in Figs. 3A and B shows that the more consistent group profiles resemble their individual counterparts.

In each data set, our method detects systems with rather flat profiles over the entire set of presented categories. These profiles match the functional definition of early areas in the visual cortex,

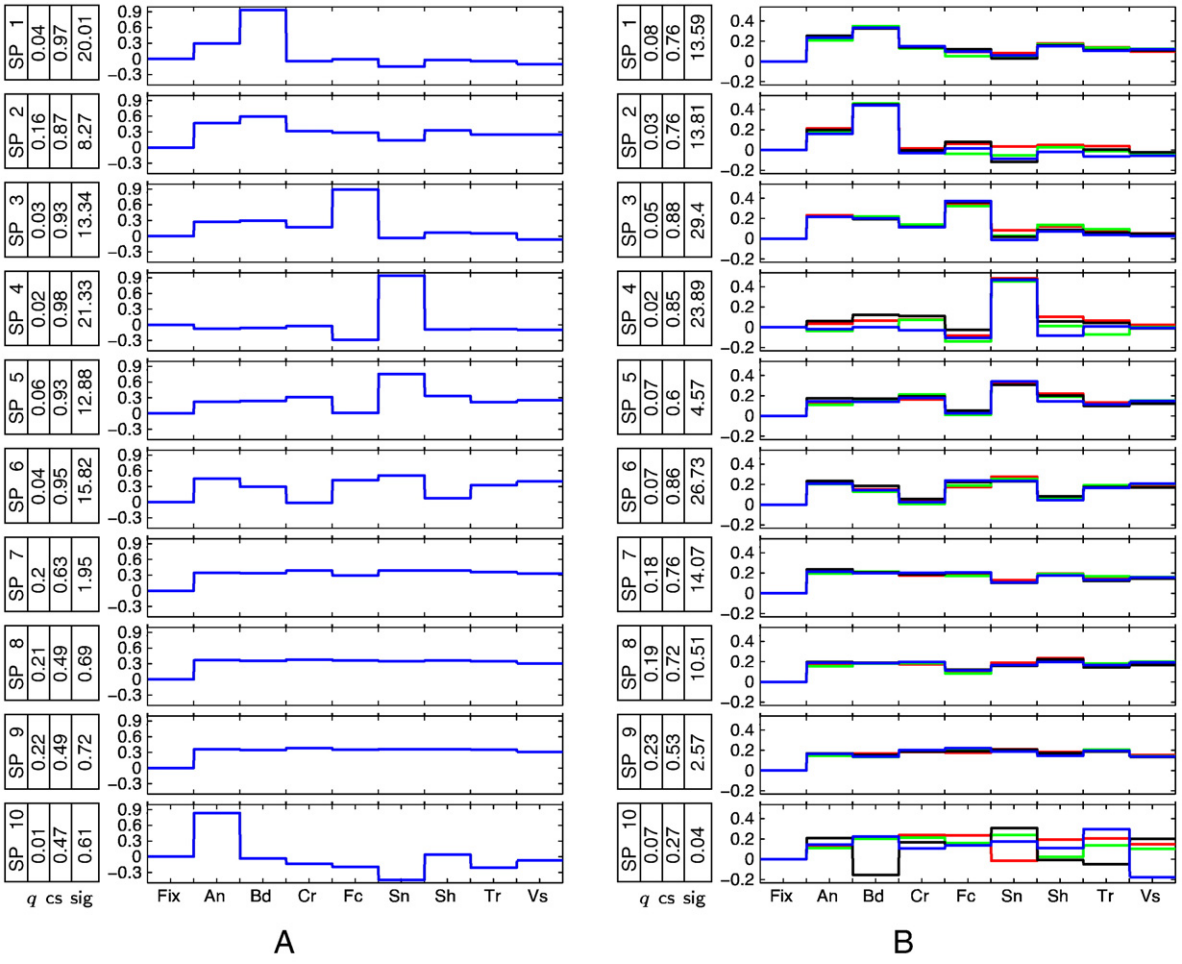


Fig. 4. Sets of 10 discovered group system selectivity profiles for (A) 8-dimensional, and (B) 32-dimensional data. Different colors (blue, black, green, red) represent different components of the profiles corresponding to the same category. We added zero to each vector to represent Fixation. The weight q for each selectivity profile is also reported along with the consistency score (cs) and the significance value.

selective to lower level features in the visual field. Not surprisingly, there is a large number of voxels associated with these systems, as suggested by their estimated weights.

The 16-Dimensional and 32-Dimensional data sets allows us to examine the consistency of the discovered profiles across different image sets and different runs. Different components of the selectivity profiles that correspond to the same category of images, illustrated with different colors in Fig. 3, have nearly identical values. This demonstrates consistency of the estimated profiles across experimental runs and image sets. The improvement in consistency from the individual data in Fig. 3B to the group data of Fig. 3A further justifies our argument for fusing data from different subjects.

To examine the robustness of the discovered selectivity profiles to the change in the number of clusters, we ran the same analysis on the 16-Dimensional data for 8 and 12 clusters. Comparing the results in Fig. 5 with those of Fig. 3A, we conclude that selectivity properties of the more consistent selectivity profiles remain relatively stable. In general, running the algorithm for many different values of K, we observed that increasing the number of clusters usually results in the split of some of the clusters but does not significantly alter the pattern of discovered profiles and their maps.

In order to find the significance of consistency scores achieved by each of these selectivity profiles, we performed a permutation test as described in Methods. For each data set, we generated 10,000 permuted data sets by randomly shuffling labels of different experimental blocks. The resulting null hypothesis distributions are shown in Fig. 6 for different data sets. Using these distributions, we compute the statistical significance of the consistency scores presented for the selectivity profiles in Figs. 3, 4, and 5.

Spatial maps

We also examine the spatial maps associated with each system. Fig. 7 shows the standard localization map for the face selective areas FFA and OFA in blue. This map is found by applying the t-test to identify voxels with higher response to faces when compared to objects, with a threshold $p = 10^{-4}$, in one of the subjects. For comparison, Fig. 7 also shows the voxels in the same slices assigned by our method to the system with the selectivity profile 3 in Fig. 4A that exhibits face selectivity (red). The assignments found by our method represent probabilities over cluster labels. To generate the map, we assign each voxel to its corresponding maximum a posteriori

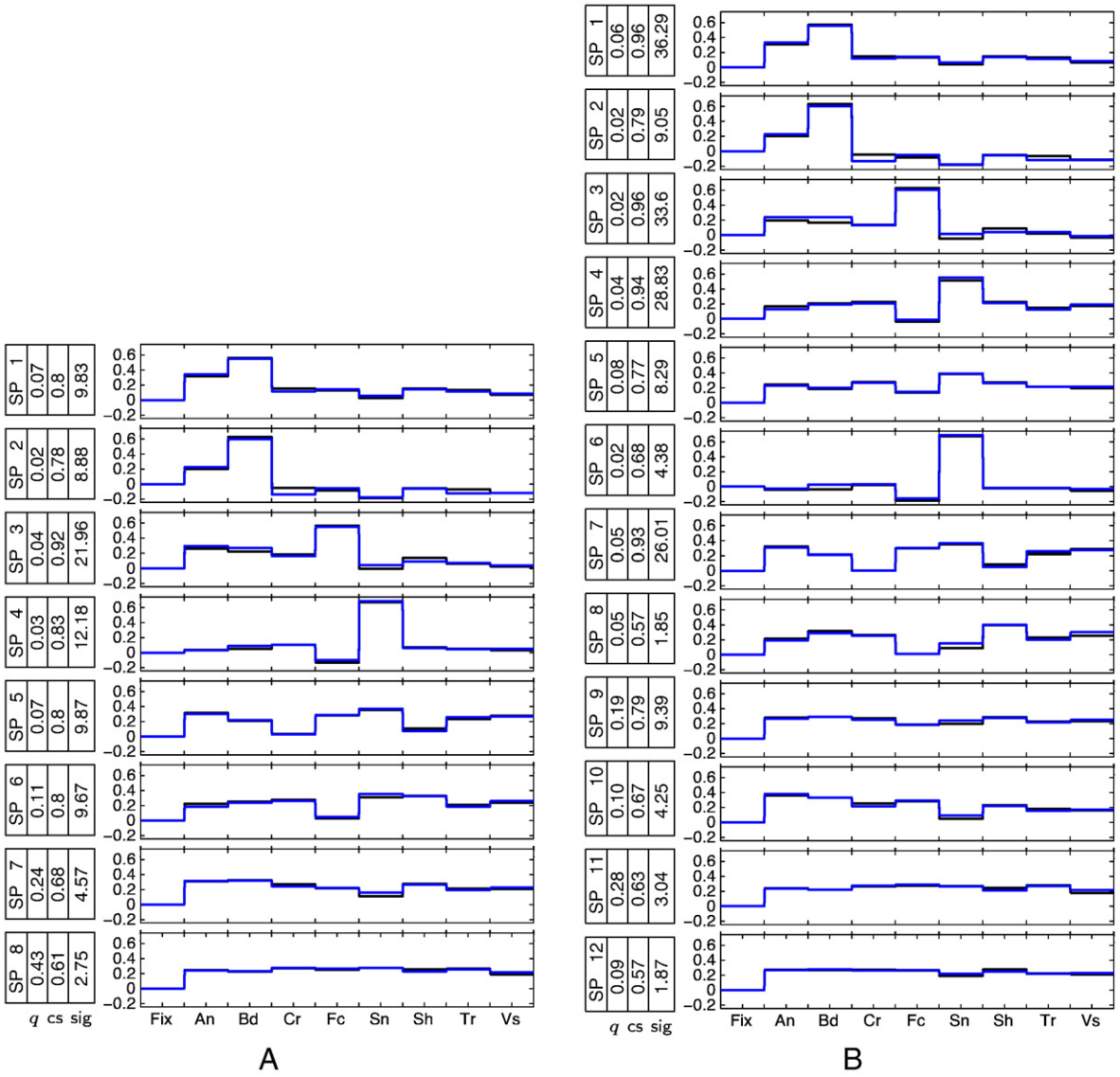


Fig. 5. Group system selectivity profiles in the 16-Dimensional data for (A) 8, and (B) 12 clusters. The colors (blue, black) represent the two distinct components of the profiles corresponding to the same category, and the weight q for each system is also indicated along with the consistency score (cs) and its significance value found in the permutation test.

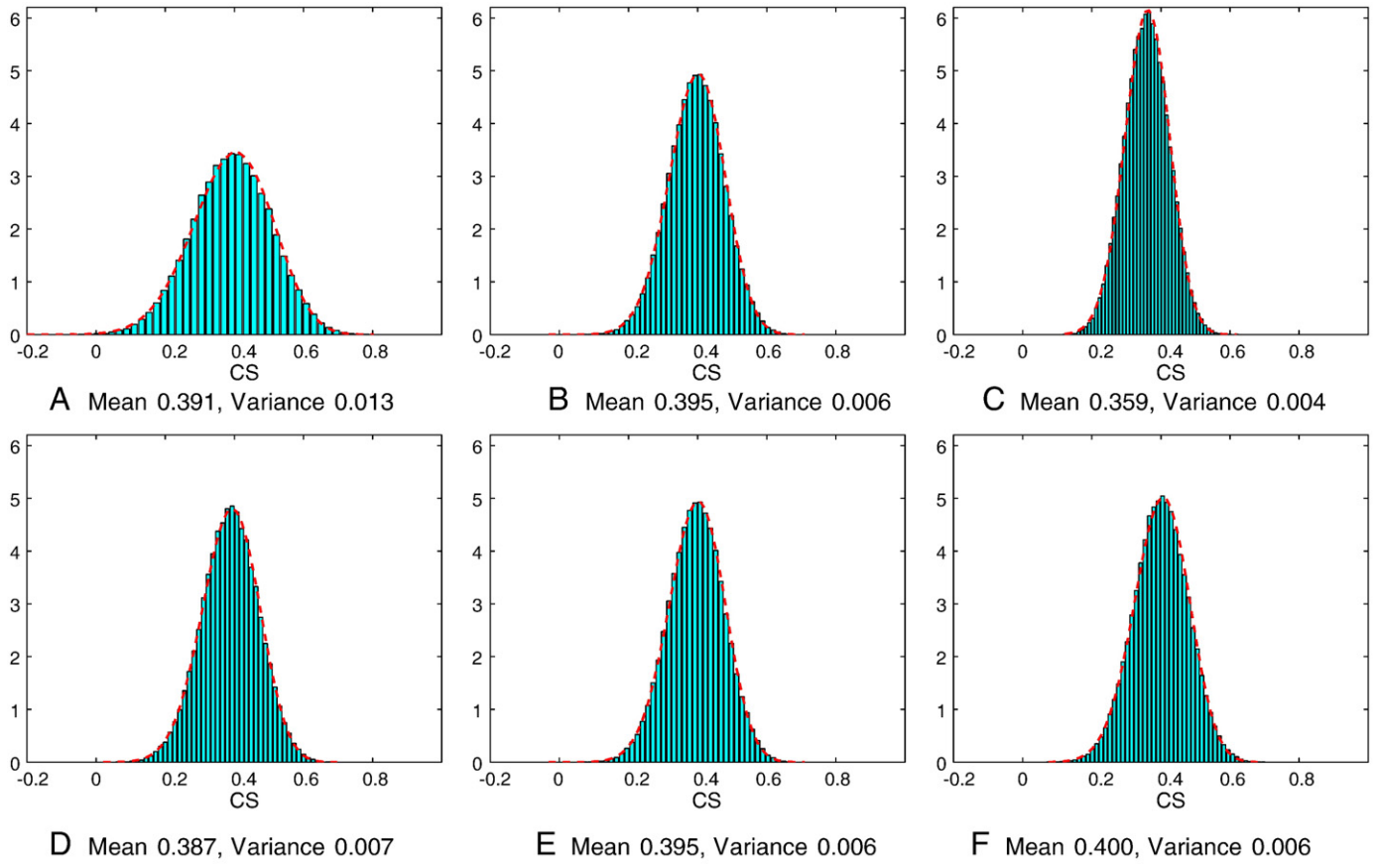


Fig. 6. Null hypothesis distributions for the consistency score values, computed from 10,000 random permutations of the data. Histograms A, B and C show the results for 8-, 16-, and 32-dimensional data with 10 clusters, respectively. Histograms D, E, and F correspond to 8, 10, and 12 clusters in 16-dimensional data (B and E are identical). We normalized the counts by the product of bin size and the overall number of samples so that they could be compared with the estimated Beta distribution, indicated by the dashed red line.

(MAP) cluster label. We have identified on these maps the approximate locations of the two known face-selective regions, FFA and OFA, based on the result of the significance map, as it is common in the field. Fig. 7 illustrates that, although the two maps are derived with very different assumptions, they mostly agree, especially in the more interesting areas where we expect to find face selectivity. As mentioned in Space of Selectivity Profiles, the conventional method identifies a much larger region as face selective, including parts in the higher slices of Fig. 7 which we expect to be in the non-selective V1 area. Our map, on the contrary, does not include these voxels.

We compute three localization maps for face, scene, and body selective regions by applying statistical tests comparing response of each voxel to faces, scenes, and bodies, respectively, with objects, and thresholding them at $p = 10^{-4}$. To define selective systems in our results, we employ the conventional definition requiring the response to the preferred category to be at least twice the value of the response to other stimuli. We observe that the largest cluster always has a flat profile with no selectivity, e.g., Figs. 4A and B. We form the map associated with the largest system as another case for comparison and call it the non-selective profile. Table 1 shows the resulting values of our overlap measure averaged across all subjects for $K = 8$, 10, and 12. We first note that the overlap between the functionally related regions is significantly higher than that of the unrelated pairs. Moreover, these results are qualitatively stable with changes in the number of clusters.

In the table, we also present the results of the algorithm applied to the data of each individual subject separately. We notice higher average overlap measures and lower standard deviations for the group data. This is due to the fact that fusing data from a cohort of subjects improves the accuracy of our estimates of the category selective profiles. As a result, we discover highly selective profiles whose response to the preferred stimuli satisfies the condition for being more than twice the other categories. On the other hand, in the results from the individual data for some noisier subjects, even the selective system does not satisfy this definition. For these subjects, no system is identified as exclusively selective of that category, degrading the average overlap measure. The improved robustness of the selectivity profile estimates in the group data prevents this effect and leads to better agreement in the spatial maps.

Discussion

Our new model rediscovers selectivity for faces, places, and bodies, now without assuming spatial contiguity of functionally similar voxels and without even assuming that category selectivity itself is a dominant property of the visual pathway.

Interestingly, we do not detect other possible profiles not known to have a corresponding selective area based on the high level vision literature. Prior work has shown that categories such as cars, shoes, and animals do not correspond to selective regions as robust and consistent as EBA, PPA, and FFA (Downing et al., 2006). Our results clearly agree with these findings but now from a completely data-driven perspective. This finding shows that the method allows us to explore the space of selectivity profiles in a less biased fashion and to search for types of selectivity that have not been hypothesized before. A possible example is the selectivity profile 6 in Fig. 3A that consistently appears with the same shape in all individual subjects, despite not being exclusively selective of a single category.

The cross-subject consistency scores show qualitative correlation with consistency across runs and different image sets. For example,

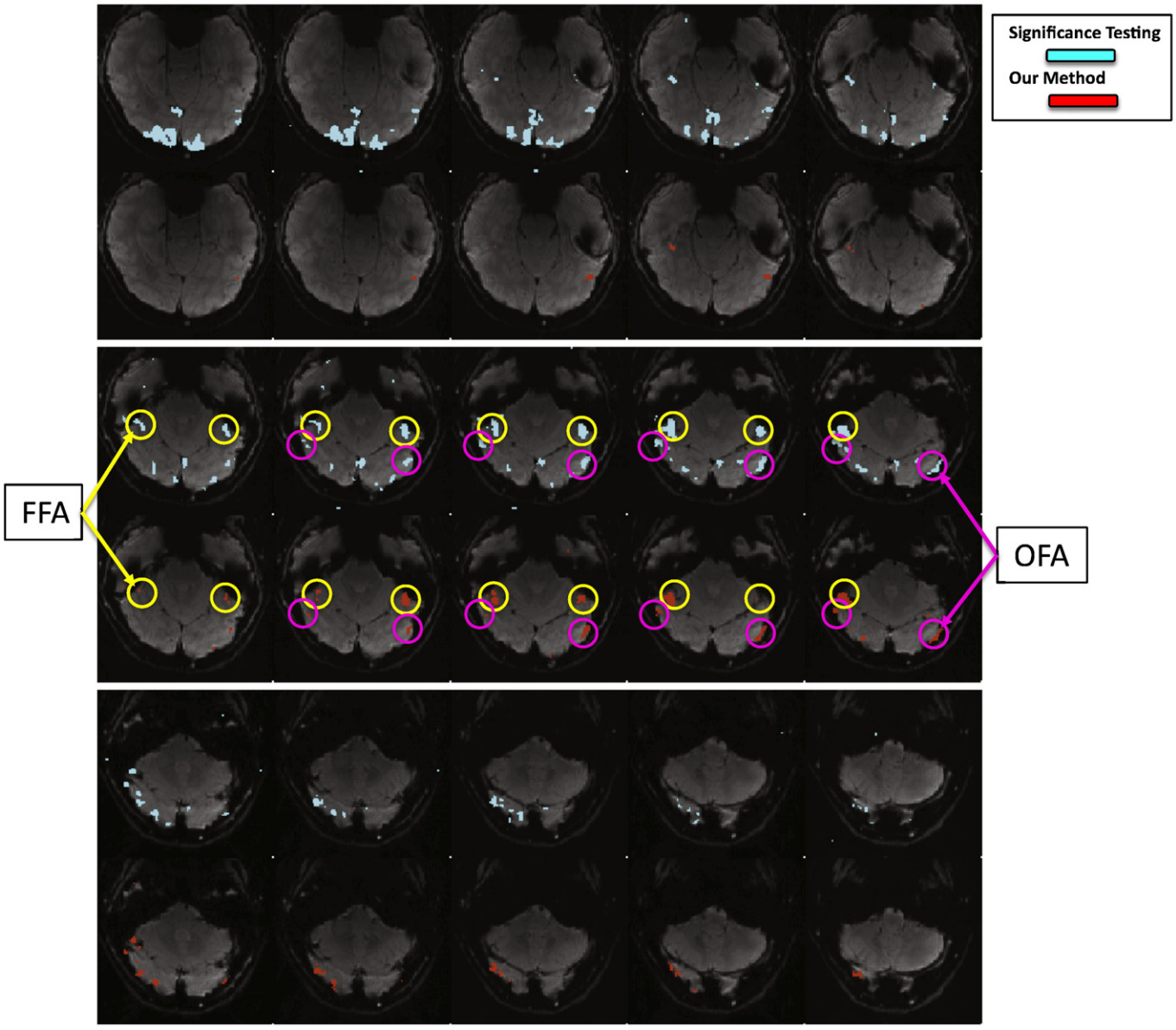


Fig. 7. Spatial maps of the face selective regions found by the significance test (light blue) and the mixture model (red). Slices from the each map are presented in alternating rows for comparison. The approximate locations of the two face-selective regions FFA and OFA are shown with yellow and purple circles, respectively.

consider the 10th selectivity profile in Fig. 4B that has low consistency score and whose components show considerable variability across repetitions of the category. By its definition, cross-subject consistency does not have to necessarily predict consistency of a profile across image sets and experimental runs. But if selectivity profiles are true signatures of the types of category selectivity existing in all subjects,

Table 1
Asymmetric overlap measures between the spatial maps corresponding to our method and the conventional hypothesis-driven approach. The exclusively selective systems for the three categories of Bodies, Faces, and Scenes, and the non-selective system (rows) are compared with the localization maps detected via traditional contrasts (columns). Values are averaged across all 6 subjects in the experiment.

Gr. $K=10$	Face	Body	Place	Indiv. $K=10$	Face	Body	Place
Face	**0.78** ± 0.08	0.14 ± 0.11	0	Face	**0.28** ± 0.44	0.05 ± 0.11	0.00 ± 0.01
Body	0.07 ± 0.06	**0.94** ± 0.01	0.01 ± 0.02	Body	0.01 ± 0.09	**0.65** ± 0.51	0.01 ± 0.02
Scene	0.01 ± 0.01	0.04 ± 0.04	**0.57** ± 0.19	Scene	0.01 ± 0.01	0.05 ± 0.08	**0.61** ± 0.47
Non-selective	0.06 ± 0.03	0.02 ± 0.02	0.1 ± 0.05	Non-selective	0.09 ± 0.09	0.09 ± 0.08	0.13 ± 0.13
Gr. $K=8$	Face	Body	Place	Gr. $K=12$	Face	Body	Place
Face	**0.72** ± 0.08	0.16 ± 0.11	0	Face	**0.83** ± 0.06	0.15 ± 0.12	0.0 ± 0.01
Body	0.07 ± 0.06	**0.94** ± 0.1	0.01 ± 0.02	Body	0.07 ± 0.06	**0.94** ± 0.09	0.01 ± 0.02
Scene	0.02 ± 0.03	0.04 ± 0.06	**0.79** ± 0.19	Scene	0.01 ± 0.01	0.05 ± 0.05	**0.66** ± 0.19
Non-selective	0.05 ± 0.02	0.02 ± 0.02	0.09 ± 0.04	Non-selective	0.08 ± 0.04	0.03 ± 0.03	0.09 ± 0.05

we expect them to exhibit consistent patterns both across subjects and across different samples of the same categories of visual stimuli.

We discussed in Spatial Maps that the spatial maps of our systems have considerable overlap with the corresponding thresholded significance maps found by the conventional method. Note however that we do not expect a perfect overlap between the two methods. The experts commonly identify some subsections of the significance maps as the selective regions based on their prior anatomical knowledge. Therefore, the degree of overlap between our systems and the localization maps does not necessarily yield a quantitative measure of accuracy of our results. Rather, it acts as an argument for relative agreement of these two different definitions of selectivity.

Several dimensions remain for further extension and improvements of the current model. As with many other learning methods, our approach so far does not offer a systematic way to choose the number of clusters (model components, K). Although our results show relative robustness to changes in K, it is more desirable to have a method for automatic selection of the number of systems. By employing nonparametric approaches, such as Dirichlet processes (Teh et al., 2006), that integrate estimation of component number within the modeling framework, we can design appropriate infinite mixture models suited to our application (Rasmussen, 2000). Another possible extension involves introducing an explicit model of inter-subject variability. The model, as it stands now, does not distinguish the individual selectivity profiles from the group profiles, effectively assuming identical structure at the group level and at the individual subject level. We are developing a hierarchical model to address this point.

We also aim to extend the model to include a clustering structure in the space of experimental conditions. In the context of high level vision, for instance, this will enable discovery of the categories of visual stimuli intrinsic to the brain's visual representation. As a result, we can present distinct images to the subjects in the experiment and search for groupings of the images suggested by the data. We are currently performing event-related visual fMRI experiments with richer sets of images than presented here to examine the possibility of discovering novel patterns of selectivity using our new approach to data analysis.

To conclude, we presented an exploratory method that enables automatic discovery of the patterns of selectivity in fMRI experiments with numerous conditions. Our method is based on the idea of selectivity profiles that characterize the specificity of response in a variety of experimental conditions. The mixture model based on the selectivity profiles yields algorithms for discovery of brain systems with coherent patterns of selectivity that robustly appear across subjects. Defining these systems on a purely functional basis opens up the possibility of investigating likely functional systems with significant anatomical variability. The method further allows us to bypass the difficult procedure of spatial normalization for group analysis.

Acknowledgments

This work was supported in part by the McGovern Institute Neurotechnology Program, NIHgrants NIBIB NAMIC U54-EB005149 and NCRR NAC P41-RR13218 to PG, NEI grant 13455 to NGK, NSF grant CAREER 0642971 to PG and IIS/CRCNS 0904625 to PG and NGK, in addition to an MIT HST Catalyst grant and an NDSEG fellowship to EV.

Appendix A. Derivation of the EM update rules

We let $\Theta = \{\{q_k, m_k\}_{k=1}^K, \lambda\}$ be the full set of parameters and derive the EM algorithm for maximizing the log-likelihood function

$$L(\Theta) = \sum_v^V \log p(y_v;\Theta) \tag{A.1}$$

for a mixture model $p(y;\Theta) = \Sigma_{k=1}^K q_k f(y;m_k,\lambda)$. The EM algorithm (Dempster et al., 1977) assumes a hidden random variable k that represents the assignment of each data point to its corresponding component in the model. This suggests a model in the joint space of observed and hidden variables:

$$p(y_v,k;\Theta) = q_k f(y_v;m_k,\lambda), \tag{A.2}$$

where $k \in \{1,\ldots,K\}$, and the likelihood of observed data is simply

$$p(y_v;\Theta) = \sum_{k=1}^K p(y_v,k;\Theta). \tag{A.3}$$

With a given set of parameters $\Theta^{(t)}$ in step t, the E-step involves computing the posterior distribution of the hidden variable given the observed data. Since the data for each voxel is assumed to be an *i.i.d.* sample from the joint distribution (A.2), the posterior distribution for the assignment of all voxels can also be factored into terms for each voxel:

$$\begin{aligned} p^{(t)}(k|y_v) \triangleq p\left(k|y_v;\Theta^{(t)}\right) &= \frac{p\left(k,y_v;\Theta^{(t)}\right)}{p\left(y_v;\Theta^{(t)}\right)} \\ &= \frac{q_k^{(t)} f\left(y_v;m_k^{(t)},\lambda^{(t)}\right)}{\sum_{k'=1}^K q_{k'}^{(t)} f\left(y_v;m_{k'}^{(t)},\lambda^{(t)}\right)} \\ &= \frac{q_k^{(t)} e^{\lambda^{(t)}<y_v,m_k^{(t)}>}}{\sum_{k'=1}^K q_{k'}^{(t)} e^{\lambda^{(t)}<y_v,m_{k'}^{(t)}>}}. \end{aligned} \tag{A.4}$$

Using this distribution, we can express the target function of the M-step:

$$\begin{aligned} L\left(\Theta;\Theta^{(t)}\right) &= \sum_{v=1}^V E_{k|y_v;\Theta^{(t)}}[\log p(y_v,k;\Theta)] \\ &= \sum_{v=1}^V \sum_{k=1}^K p^{(t)}(k|y_v)\log p(y_v,k;\Theta) \\ &= \sum_{v=1}^V \sum_{k=1}^K p^{(t)}(k|y_v)\log[\,q_k f(y_v;m_k,\lambda)\,] \\ &= \sum_{v=1}^V \sum_{k=1}^K p^{(t)}(k|y_v)[\log q_k + \log C_D(\lambda) + \lambda\langle m_k,y_v\rangle]. \end{aligned} \tag{A.5}$$

Taking the derivative of this function along with the appropriate Lagrange multipliers yields the update rules for the model parameters in iteration $(t+1)$. For the cluster centers m_k, we have

$$\begin{aligned} 0 &= \frac{\partial}{\partial m_k}\left[\sum_{v=1}^V \sum_{k'=1}^K p^{(t)}(k'|y_v)\langle m_{k'},y_v\rangle - \sum_{k'=1}^K \gamma_{k'}(\langle m_{k'},m_{k'}\rangle - 1)\right] \\ &= \sum_{v=1}^V y_v p^{(t)}(k|y_v) - 2\gamma_k m_k, \end{aligned} \tag{A.6}$$

which implies the update rule $m_k^{(t+1)} = \frac{1}{2\gamma_k}\sum_{v=1}^V y_v p^{(t)}(k|y_v)$. The Lagrange multiplier ensures that m_k is a unit vector, i.e.,

$$\gamma_k = \frac{1}{2}\left\|\sum_{v=1}^V y_v p^{(t)}(k|y_v)\right\|.$$

Similarly, we find the concentration parameter λ:

$$\begin{aligned} 0 &= \frac{1}{V}\frac{\partial}{\partial\lambda}\left[V\log C_D(\lambda) + \lambda\sum_{v=1}^V \sum_{k=1}^K p^{(t)}(k|y_v)\langle m_k,y_v\rangle\right], \\ &= \frac{\partial}{\partial\lambda}\left[\left(\frac{D}{2}-1\right)\log\lambda - \log I_{D/2-1}(\lambda) + \frac{\lambda}{V}I^{(t+1)}\right], \\ &= \frac{D/2-1}{\lambda} - \frac{I'_{D/2-1}(\lambda)}{I_{D/2-1}(\lambda)} + \frac{I^{(t+1)}}{V}, \\ &= -\frac{I_{D/2}(\lambda)}{I_{D/2-1}(\lambda)} + \frac{I^{(t+1)}}{V}, \end{aligned} \tag{A.7}$$

where we have substituted $m_k^{(t+1)}$ in the first line, used the definition of Eq. (12) in the second line, and the last equality follows from the properties of the modified Bessel functions. It follows then that $A_D(\lambda^{(t+1)}) = \Gamma^{(t+1)}$. Finally, for the cluster weights q_k, adding the Lagrange multiplier to guarantee that the weights sum to 1, we find

$$\begin{aligned} 0 &= \frac{\partial}{\partial q_k}\left[\sum_{v=1}^{V}\sum_{k'=1}^{K} p^{(t)}(k'|y_v)\log q_{k'} - \zeta\left(\sum_{k'=1}^{K} q_{k'} - 1\right)\right] \\ &= \frac{1}{q_k}\sum_{v=1}^{V} p^{(t)}(k|y_v) - \zeta, \end{aligned} \tag{A.8}$$

which together with the normalization condition results in the update $q_k^{(t+1)} = \frac{1}{V}\sum_{v=1}^{V} p^{(t)}(k|y_v)$.

Appendix B. Estimation of concentration parameter

In order to update the concentration parameter in the M-step using (11), we need to solve for λ in the equation

$$A_D(\lambda) = \frac{I_{D/2+1}(\lambda)}{I_{D/2}(\lambda)} = \Gamma. \tag{B.1}$$

Fig. B.1 shows the plot of function $A_D(\cdot)$ for several values of D. This function is smooth and monotonically increasing, taking values in the interval [0,1). An approximate solution to (B.1) has been suggested in (Banerjee et al., 2006) but the proposed expression does not yield accurate values in our range of interest for D. Therefore, we derive a different approximation using the inequality

$$\frac{x}{\gamma+\frac{1}{2}+\sqrt{x^2+\left(\gamma+\frac{3}{2}\right)^2}} \le \frac{I_{\gamma+1}(x)}{I_\gamma(x)} \le \frac{x}{\gamma+\frac{1}{2}+\sqrt{x^2+\left(\gamma+\frac{1}{2}\right)^2}} \tag{B.2}$$

proved in (Amos, 1974). Defining $u = \frac{D-1}{2\lambda}$, it follows from Eq. (B.1) and Inequality (B.2) that

$$\frac{1}{u+\sqrt{1+\left(1+\frac{2}{D-1}\right)^2 u^2}} \le \Gamma \le \frac{1}{u+\sqrt{1+u^2}}. \tag{B.3}$$

Due to continuity of $\left(u+\sqrt{1+\alpha^2u^2}\right)^{-1}$ as a function of $\alpha \ge 1$, this expression equals Γ for at least one value in the interval $1 \le \alpha \le 1+\frac{2}{D-1}$. For this value of α, we have

$$\begin{aligned} &\left(\alpha^2-1\right)u^2 + \frac{2}{\Gamma}u - \left(\frac{1}{\Gamma^2}-1\right) = 0 \\ &\Rightarrow u = \frac{\sqrt{1+(\alpha^2-1)(1-\Gamma^2)}-1}{(\alpha^2-1)\Gamma}. \end{aligned} \tag{B.4}$$

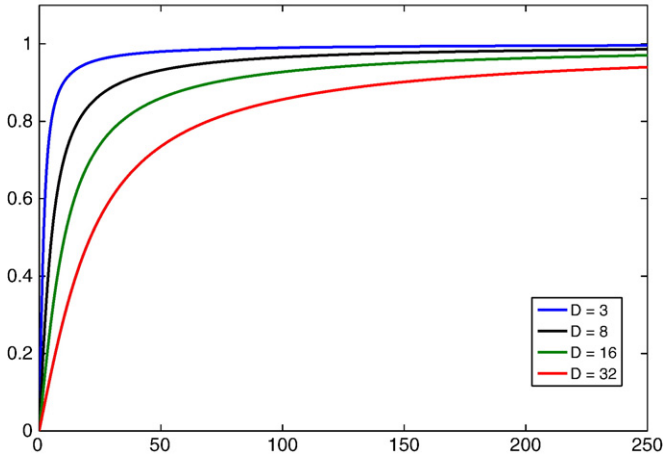


Fig. B.1. Plot of function $A_D(\cdot)$ for different values of D.

The expression for u is a monotonically decreasing function of $\alpha^2 - 1$ where $0 < \alpha^2 - 1 \le \frac{4D}{(D-1)^2}$; therefore, we find

$$\frac{(D-1)^2}{4D\Gamma}\left(\sqrt{1+\frac{4D(1-\Gamma^2)}{(D-1)^2}}-1\right) \le u \le \frac{1-\Gamma^2}{2\Gamma}. \tag{B.5}$$

Now, using the inequality $\sqrt{1+x} \ge 1+\frac{1}{2}x-\frac{1}{8}x^2$, we find a simpler expression for the lower bound

$$\frac{1-\Gamma^2}{2\Gamma}\left(1-\frac{D\left(1-\Gamma^2\right)}{(D-1)^2}\right) \le u \le \frac{1-\Gamma^2}{2\Gamma}. \tag{B.6}$$

Finally, the parameter can be bounded by

$$\frac{(D-1)\Gamma}{1-\Gamma^2} \le \lambda \le \frac{(D-1)\Gamma}{1-\Gamma^2}\left(1-\frac{D\left(1-\Gamma^2\right)}{(D-1)^2}\right)^{-1}. \tag{B.7}$$

Because of the monotonicity of $A_D(\cdot)$, starting from the average of the two bounds and taking a few Newton steps towards zero of Eq. (B.1), we easily reach a good solution. However, when Γ is too close to 1 and, hence, λ is large, evaluation of the function $A_D(\cdot)$ becomes challenging due to the exponential behavior of the Bessel functions. In this case, when Γ is large enough such that $\frac{D\left(1-\Gamma^2\right)}{(D-1)^2} \ll 1$ holds, we can approximate the second term in the upper bound of Eq. (B.7) as $\left(1+\frac{D\left(1-\Gamma^2\right)}{(D-1)^2}\right)$, reaching the final approximation

$$\lambda \approx \frac{(D-1)C}{1-C^2} + \frac{DC}{2(D-1)}. \tag{B.8}$$

References

Aguirre, G., Zarahn, E., D'Esposito, M., 1997. Empirical analyses of BOLD fMRI statistics II. spatially smoothed data collected under null-hypothesis and experimental conditions. Neuroimage 5 (3), 199–212.

Aguirre, G., Zarahn, E., D'Esposito, M., 1998. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. Neuron 21 (2), 373–383.

Amos, D., 1974. Computation of modified Bessel functions and their ratios. Math. Comput. 28, 239–251.

Baker, C., Liu, J., Wald, L., Kwong, K., Benner, T., Kanwisher, N., 2007. Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. Proc. Natl. Acad. Sci. 104 (21), 9087–9092.

Banerjee, A., Dhillon, I., Ghosh, J., Sra, S., 2006. Clustering on the unit hypersphere using von Mises–Fisher distributions. J. Mach. Learn. Res. 6 (2), 1345–1382.

Baumgartner, R., Scarth, G., Teichtmeister, C., Somorjai, R., Moser, E., 1997. Fuzzy clustering of gradient-echo functional MRI in the human visual cortex. Part I: reproducibility. J. Magn. Reson. Imaging 7 (6), 1094–1108.

Baumgartner, R., Windischberger, C., Moser, E., 1998. Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis. Magn. Reson. Imaging 16 (2), 115–125.

Beckmann, C., Smith, S., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23 (2), 137–152.

Beckmann, C., Smith, S., 2005. Tensorial extensions of independent component analysis for multisubject FMRI analysis. Neuroimage 25 (1), 294–311.

Brett, M., Johnsrude, I., Owen, A., 2002. The problem of functional localization in the human brain. Nat. Rev. Neurosci. 3 (3), 243–249.

Burgess, N., Maguire, E., Spiers, H., OKeefe, J., Epstein, R., Harris, A., Stanley, D., Kanwisher, N., 1999. The parahippocampal place area: recognition, navigation, or encoding. Neuron 23, 115–125.

Calhoun, V., Adali, T., McGinty, V., Pekar, J., Watson, T., Pearlson, G., 2001a. fMRI activation in a visual-perception task: network of areas detected using the general linear model and independent components analysis. Neuroimage 14 (5), 1080–1088.

Calhoun, V., Adali, T., Pearlson, G., Pekar, J., 2001b. A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp. 14 (3), 140–151.

Dempster, A., Laird, N., Rubin, D., 1977. Maximum likelihood from incomplete data via the em algorithm. J. R. Stat. Soc. B Methodol. 1–38.

Diestel, R., 2005. Graph Theory. Springer-Verlag, New York.

Downing, P., Chan, A.-Y., Peelen, M., Dodds, C., Kanwisher, N., 2006. Domain specificity in visual cortex. Cereb. Cortex 16 (10), 1453–1461.

Downing, P., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. Science 293 (5539), 2470–2473.

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392 (6676), 598–601.

Fadili, M., Ruan, S., Bloyet, D., Mazoyer, B., 2000. A multistep unsupervised fuzzy clustering analysis of fMRI time series. Hum. Brain Mapp. 10 (4), 160–178.
Filzmoser, P., Baumgartner, R., Moser, E., 1999. A hierarchical clustering method for analyzing functional MR images. Magn. Reson. Imaging 17 (6), 817–826.
Friston, K., Ashburner, J., Kiebel, S., Nichols, T., Penny, W. (Eds.), 2007. Statistical parametric mapping: the analysis of functional brain images. Academic Press, Elsevier.
Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C., Frackowiak, R., et al., 1994. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2 (4), 189–210.
Gee, J., Alsop, D., Aguirre, G., 1997. Effect of spatial normalization on analysis of functional data. Proceedings of SPIE, Medical Imaging, 3034, pp. 550–560.
Golay, X., Kollias, S., Stoll, G., Meier, D., Valavanis, A., Boesiger, P., 1998. A new correlation-based fuzzy logic clustering algorithm for fMRI. Magn. Reson. Med. 40 (2), 249–260.
Golland, P., Golland, Y., Malach, R., 2007. Detection of spatial activation patterns as unsupervised segmentation of fMRI data. Proceedings of MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention. Vol. 4791 of LNCS. Springer, pp. 110–118.
Goutte, C., Toft, P., Rostrup, E., Nielsen, F., Hansen, L., 1999. On clustering fMRI time series. Neuroimage 9 (3), 298–310.
Goutte, C., Hansen, L., Liptrot, M., Rostrup, E., 2001. Feature-space clustering for fMRI meta-analysis. Hum. Brain Mapp. 13 (3), 165–183.
Grill-Spector, K., Malach, R., 2004. The human visual cortex. Ann. Rev. Neurosci. 27, 649–677.
Kanwisher, N., 2003. The ventral visual object pathway in humans: evidence from fMRI. In: Chalupa, L., Wener, J. (Eds.), The Visual Neurosciences. MIT Press, pp. 1179–1189.
Kanwisher, N., McDermott, J., Chun, M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17 (11), 4302–4311.
Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. R. Soc. London B Biol. Sci. 361 (1476), 2109–2128.
Kriegeskorte, N., Mur, M., Ruff, D., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P., 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60 (6), 1126–1141.
Kuhn, H., 1955. The Hungarian Method for the assignment problem. Nav. Res. Logist. Q. 2, 83–97.
Malach, R., Reppas, J., Benson, R., Kwong, K., Jiang, H., Kennedy, W., Ledden, P., Brady, T., Rosen, B., Tootell, R., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. 92 (18), 8135–8139.
Mardia, K., 1975. Statistics of directional data. J. R. Stat. Soc. B. Methodol. 349–393.
McCarthy, G., Puce, A., Gore, J., Allison, T., 1997. Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. 9 (5), 605–610.
McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., Sejnowski, T., 1998. Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp. 6 (3), 160–188.
McLachlan, G., Peel, D., 2000. Finite Mixture Models. Wiley, New York.
Moser, E., Diemling, M., Baumgartner, R., 1997. Fuzzy clustering of gradient-echo functional MRI in the human visual cortex. Part II: quantification. J. Magn. Reson. Imaging 7 (6).
Op de Beeck, H., Haushofer, J., Kanwisher, N., 2008. Interpreting fMRI data: maps, modules and dimensions. Nat. Rev. Neurosci. 9 (2), 123–135.
Peelen, M., Downing, P., 2007. The neural basis of visual body perception. Nat. Rev. Neurosci. 8 (8), 636–648.
Penny, W., Holmes, A., 2003. Random effects analysis. In: R.S.J., F., K.J., F., C.D., F. (Ed.), Human Brain Function II. Elsevier, Oxford, pp. 843–850.
Rasmussen, C., 2000. The infinite Gaussian mixture model. Adv. Neural Inf. Process. Syst. 12, 554–560.
Rossion, B., Caldara, R., Seghier, M., Schuller, A., Lazeyras, F., Mayer, E., 2003. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain 126 (11), 2381–2395.
Schwarzlose, R., Baker, C., Kanwisher, N., 2005. Separate face and body selectivity on the fusiform gyrus. J. Neurosci. 25 (47), 11055–11059.
Spiridon, M., Fischl, B., Kanwisher, N., 2006. Location and spatial profile of category-specific regions in human extrastriate cortex. Hum. Brain Mapp. 27 (1).
Teh, Y., Jordan, M., Beal, M., Blei, D., 2006. Hierarchical dirichlet processes. J. Am. Stat. Assoc. 101 (476), 1566–1581.
Thirion, B., Faugeras, O., 2003. Feature detection in fMRI data: the information bottleneck approach. Proceedings of MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention. Vol. 2879 of LNCS. Springer, pp. 83–91.
Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., Poline, J., 2006. Dealing with the shortcomings of spatial normalization: multi-subject parcellation of fMRI datasets. Hum. Brain Mapp. 27 (8), 678–693.
Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., Poline, J., 2007a. Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. Neuroimage 35 (1), 105–120.
Thirion, B., Pinel, P., Tucholka, A., Roche, A., Ciuciu, P., Mangin, J., Poline, J.-B., 2007b. Structural analysis of fMRI data revisited: Improving the sensitivity and reliability of fMRI group studies. IEEE Trans. Med. Imaging 26 (9), 1256–1269.
Woolrich, M., Ripley, B., Brady, M., Smith, S., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14 (6), 1370–1386.