



## UvA-DARE (Digital Academic Repository)

### Identifying distributed and overlapping clusters of hemodynamic synchrony in fMRI data sets

Ghebreab, S.; Smeulders, A.W.M.

**DOI**

[10.1007/s10044-010-0186-6](https://doi.org/10.1007/s10044-010-0186-6)

**Publication date**

2011

**Document Version**

Author accepted manuscript

**Published in**

Pattern Analysis and Applications

[Link to publication](#)

**Citation for published version (APA):**

Ghebreab, S., & Smeulders, A. W. M. (2011). Identifying distributed and overlapping clusters of hemodynamic synchrony in fMRI data sets. *Pattern Analysis and Applications*, 14(2), 175-192. <https://doi.org/10.1007/s10044-010-0186-6>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)*

# Identifying distributed and overlapping clusters of hemodynamic synchrony in fMRI data sets

Sennay Ghebream · Arnold W. M. Smeulders

Received: 25 September 2009 / Accepted: 22 August 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Natural sensory stimuli elicit complex brain responses that manifest in fMRI as widely distributed and overlapping clusters of hemodynamic responses. We propose a statistical signal processing method for finding synchronous hemodynamic activity that directly or transiently reflects information about the experimental condition. When applied to fMRI data, the method searches for voxels with activation patterns exhibiting high coherence and simultaneously high variance across brain scans. The crux of the method is functional principal component analysis (fPCA) of activation patterns stored in a two-dimensional data matrix, with rows and columns representing voxels and scans, respectively. Without external information, fPCA is performed directly on this data matrix. Otherwise, the data matrix is first transformed to highlight a specific source of variation, enabling fully or partially supervised fPCA with a single parameter determining the degree of supervision. We evaluated our method on a public benchmark of fMRI scans of subjects viewing natural movies. Our method turns out to be very suitable for flexibly uncovering distributed and overlapping hemodynamic patterns that distinguish well between experimental conditions or cognitive states.

**Keywords** Natural perception · Brain activity · fMRI · Functional data analysis · Semi-supervised models

## 1 Introduction

Methods for functional magnetic resonance image (fMRI) analysis can be broadly divided into model-based analysis and data-driven analysis. The difference between the two is not absolute but rather indicates the point of departure. Model-driven methods, such as the common general linear model [5, 12, 14, 27, 37], assume an explicit temporal hemodynamic model based upon the experimental condition. These methods have proven to be useful for spatial localization of covariate-related brain responses. The a priori model, however, is limited in dealing with hemodynamic variations across subjects, brain regions, and even cortical layers [1, 16]. As an alternative, data-driven methods group brain responses by temporal similarity [2, 7, 30, 24] or distinguish brain response from various noise sources by data decomposition [4, 11, 27]. These methods are powerful in revealing multivariate patterns of brain activity independent of experimental conditions. The interpretation of such patterns, however, is often problematic due to the presence of many confounding sources of brain activity. Hence, the effectiveness of either data-driven and model-based methods partially resolves the fMRI data analysis problem.

A new class of methods [17, 22, 23] combines the simplicity of model-based methods with the flexibility of data-driven methods. These methods take advantage of similarities in hemodynamic patterns among subjects. Each subject's hemodynamic time course is voxel-wise correlated with every other subject's hemodynamic time courses. Intersubject correlation matrices are then constructed for all voxels to measure hemodynamic consistency given a specific task. In a post-processing step, voxels with similar temporal patterns are clustered for further examination. Intersubject similarity-based methods work well for the

---

S. Ghebream (✉) · A. W. M. Smeulders  
ISLA Lab, Informatics Institute, Faculty of Science,  
University of Amsterdam, Science Park 107,  
1098 XG Amsterdam, The Netherlands  
e-mail: s.ghebream@uva.nl

A. W. M. Smeulders  
e-mail: a.w.m.smeulders@uva.nl

identification of brain activity for such tasks as the auditory odd ball task [22]. They also work well for uncovering new brain areas responding to complex visual stimuli [17]. The versatility of these methods, however, is limited by the exclusion of valuable information from external sources. It is therefore natural, as we are pursuing here, to incorporate information about experimental conditions in the data analysis without compromising the flexibility of similarity based methods.

We propose a method that leniently uses information about the experimental condition to discover synchrony in hemodynamics. The method searches for voxels whose activation pattern exhibits high coherence and simultaneously high variance across brain scans. The crux of the method is functional principal component analysis of activation patterns stored in a two-dimensional data matrix with rows and columns representing voxels and scans, respectively. There are three modes of operation. Without external information, principal component analysis is performed on the original data matrix. Otherwise, the data matrix is first transformed to highlight specific sources of variation using stimulus data, group labels or any other coded information. The transformed data matrix is subsequently subjected to fully or partially supervised principal component analysis, with a single parameter determining the degree of supervision. Principal component analysis is performed on the rows of the data matrix in an incremental way. At each step, rows with low principal component scores are removed from the data matrix, resulting in nested voxel clusters with synchronous activity patterns. Optimal voxel clusters are subsequently determined from Gap statistics.

The underlying principle of our method comes from the popular gene shaving method (see [18]), which has been widely used in bioinformatics to find biologically relevant patterns of variations across genes, samples, and outcome measurements. Our motivation for extending the gene shaving method to fMRI data analysis is the inability of conventional fMRI data analysis to unravel the complex brain activity that natural sensory stimuli elicit [20]. Such complex brain activities often manifest in fMRI as spatially widely distributed and overlapping clusters of hemodynamic responses [19]. This type of nested clusters is the target of the method we propose here. Specifically, our fMRI data analysis method aims to detect distributed and overlapping voxel clusters with synchronous hemodynamic responses, when onsets and identities of their underlying processes are either fully known or unknown.

The difference between gene shaving and our method is that the first operates on discrete measurements (gene expression) while our method operates on signal data from EEG, fMRI or any other modality. The external source of variation may be signal data too. Here, we specifically

focus on hemodynamics in fMRI data, calling our method *voxel sieving* as it incrementally separates out voxels with asynchronous activation patterns. We evaluate voxel sieving on simulated fMRI data and on an international fMRI test benchmark involving natural movie stimuli. We explore the correspondence between voxel cluster detections and known functional specialization. In addition, we compare our method's ability to decode cognitive states with that of other state-of-the-art multivariate fMRI data analysis methods.

## 2 Materials

Stimulus and brain response data have been obtained from a publicly available benchmark for testing and comparing brain activity interpretation methods (see [32] for more detail and references). The benchmark has been extensively used in an international brain reading competition, providing the possibility to objectively compare our method's performance with that of others.

### 2.1 Data

The brain response data involve fMRI data associated with passive viewing of Home Improvement sitcom movies for approximately 20 min. This TV video provided long shots and a repeating use of a small number of actors in a small number of sets that allows common elements to reoccur. Also, the materials (character types, settings, events, objects) are typical of what the subjects would be expected to have experience with [32]. The 20-min movies contained five interruptions where no video was present but only a white fixation cross on a black background. Three subjects watched the same three movies while undergoing functional brain imaging. Neuroimage data were collected on a Siemens Allegra 3T scanner. The structural neuroimage data were acquired with 1 mm spatial resolution. The functional scans produced volumes with approximately  $V = 36,000$  brain voxels, each approximately  $3.28 \text{ mm} \times 3.28 \text{ mm} \times 3.5 \text{ mm}$ , with one volume produced every 1.75 s. These scans were preprocessed (motion correction, slice time correction, linear trend removal) and spatially normalized (non-linear registration) to the Montreal Neurological Institute brain atlas [26].

After fMRI scanning, the three subjects watched the three movies again to rate 30 movie features at time intervals corresponding to the fMRI scan rate. The extensive behavioral time vector ratings included the coding of categories such as *faces*, *motion*, and *emotional states* at multiple levels of hierarchy (i.e. *faces* versus individual *actors*). All three subjects generated ratings for each feature in each movie by moving a slider that controls a line

on a screen showing the current value of the slider. Each rating was done on a 4-point scale. For the feature *faces*, for example, 0 indicates no faces, 1 faces somewhere in the picture, 2 faces at between 25 and 50% of the image, and 3 faces seen at more than 50% of the image. Each vector-valued rating pattern was subsequently convolved with a double-gamma hemodynamic response function to define the stimulus signal. A complete description of features and generation of feature vectors can be found in [32].

We use data associated with movies 1 and 2, as data associated with movie 3 has not been made public for objective on-site evaluation purpose. As we are interested in finding continuous hemodynamics caused by the content of the movies, we exclude parts of the data corresponding to video presentations of a white fixation cross on a black background. Taking into account the hemodynamic lag, we divide each fMRI scan and each subject rating into six parts corresponding with the movie on parts. The six fMRI parts differ somewhat in number of volumes: part 1 consists of 91 volumes and the other parts of 90, 115, 108, 116, and 112 volumes, respectively. For a single movie, this results in 18 fMRI scans (3 subjects  $\times$  6 movie parts) and 18 real-valued and subject-dependent movie ratings.

We denote these four-dimensional fMRI scans by  $I_s(\mathbf{x}, t)$ , where  $s = 1, \dots, S$  indicates the sample scan,  $\mathbf{x} \in \mathbb{R}^3$  is 3D discrete spatial position, and  $t$  is time point. The real-valued ratings for sample  $s$  are denoted by the vector  $\mathbf{g}_s$ , containing  $S$  real values corresponding to the strength of a movie feature at the time scan  $I_s(\mathbf{x}, t)$  was acquired.

### 2.2 Data representation

An important first step of our approach is representation of voxel activation data as signal data rather than as a collection of discrete measurements. Such a representation enhances the discovery of underlying temporal coherences in the fMRI data [35]. It comes at the expense of slightly more complex functional statistical analysis [31], but we expect it to pay off by achieving better results. Figure 1a–d provides an illustration of our data representation approach. Invariably, in this paper bold face upper case indicates a matrix of functions, e.g.  $\mathbf{F}(t)$ , or scalars, e.g.  $\mathbf{F}$  bold face lower case indicates a vector of functions, e.g.  $\mathbf{f}(t)$ , or scalars, e.g.  $\mathbf{f}$ , and regular lower case indicates a function or a scalar.

We define the functional representation of a single voxel time course  $\mathbf{f} = [f_1, \dots, f_T]$  by

$$f^*(t) = \sum_{m=1}^M B_m(t)\omega_m \tag{1}$$

where  $B_m(t)$  is the  $m$ th basis function and  $\omega_m$  the weight of that basis. In our case B-splines are used to represent the non-periodic voxel activation data in a continuous manner.

The functional representation of all  $v = 1, \dots, V$  voxel time-courses of  $I(\mathbf{x}, t)$  forms a vector  $\mathbf{f}^*(t)$  of functions

$$\mathbf{f}^*(t) = [f_1^*(t), \dots, f_V^*(t)]^T. \tag{2}$$

Robust brain responses in fMRI generally cover multiple voxels. We therefore consider spatial clusters of voxel time courses. To avoid bias toward clusters of a given size, we hierarchically cluster voxels. Clustering is performed on the 3D brain atlas to which all fMRI scans are aligned. A computationally efficient hierarchical K-means clustering [25] is performed on the 3D grid of this atlas to assign each grid point to one of  $K$  initial cluster centers distributed equidistantly in 3D space. Cluster centers are chosen to minimize the weighted within-cluster sum of squared Euclidean distances. Clustering is repeated several times with increasing number of cluster centers, corresponding to increasing levels of hierarchy. At each hierarchical level  $l \in \mathcal{L}$  voxels are grouped in one of  $K = 2^l$  clusters. Clusters at the highest level  $l = 0$  are created by clustering with  $K = 1$ , at level  $l = 1$  by clustering with  $K = 2$ , at the next level we take  $K = 4$  and so on. The number of clusters at the lowest level is equal to the number of voxels  $V$  the atlas contains. Assuming this number is a power of two, this results in a total of  $2V - 1$  clusters. By imposing a range on the levels, for example, considering higher levels of hierarchy only ( $\mathcal{L} = \{0, \dots, L\}$  with  $L < \log_2(V)$ ), the number of all clusters to be analyzed can be limited and sensitivity to noise limited. Clusters at all levels are indexed by  $c = 1, \dots, C$  with  $C = \sum_{l \in \mathcal{L}} 2^l$ .

We transform the four-dimensional fMRI data  $I(\mathbf{x}, t)$ , by the vector of average voxel time courses  $\mathbf{f}(t) = [f_1(t), \dots, f_C(t)]$ , with

$$f_c(t) = \frac{1}{|\mathcal{V}_c|} \sum_{v \in \mathcal{V}_c} f_v^*(t) \tag{3}$$

where  $\mathcal{V}_c$  denotes the set of voxels in cluster  $c$  and  $|\mathcal{V}_c|$  denotes the number of elements in that set. We refer to  $f_c(t)$  as a *supervoxel*. Supervoxels have a regularizing effect. They reduce the multiple comparison problem and alleviate the need for spatial clustering of activated voxels as required in most voxel-wise methods.

Given a collection of  $S$  fMRI scans we define a  $C \times S$  data matrix

$$\mathbf{F}(t) = [\mathbf{f}_1(t), \dots, \mathbf{f}_S(t)] \tag{4}$$

where the rows of  $\mathbf{F}(t)$  correspond to supervoxels, the columns to fMRI scans  $I_s(\mathbf{x}, t)$ , and the element  $f_{cs}(t)$  is the  $c$ th supervoxel of scan  $s$ . For example, when only supervoxels at hierarchical levels  $\mathcal{L} = \{9, 10, 11, 12\}$  are considered for the  $S = 18$  fMRI scans from the free movie viewing study, this will result in a  $7,680 \times 18$  data matrix  $\mathbf{F}(t)$ . Each row of  $\mathbf{F}(t)$  is centered to have zero mean.

### 3 Methods

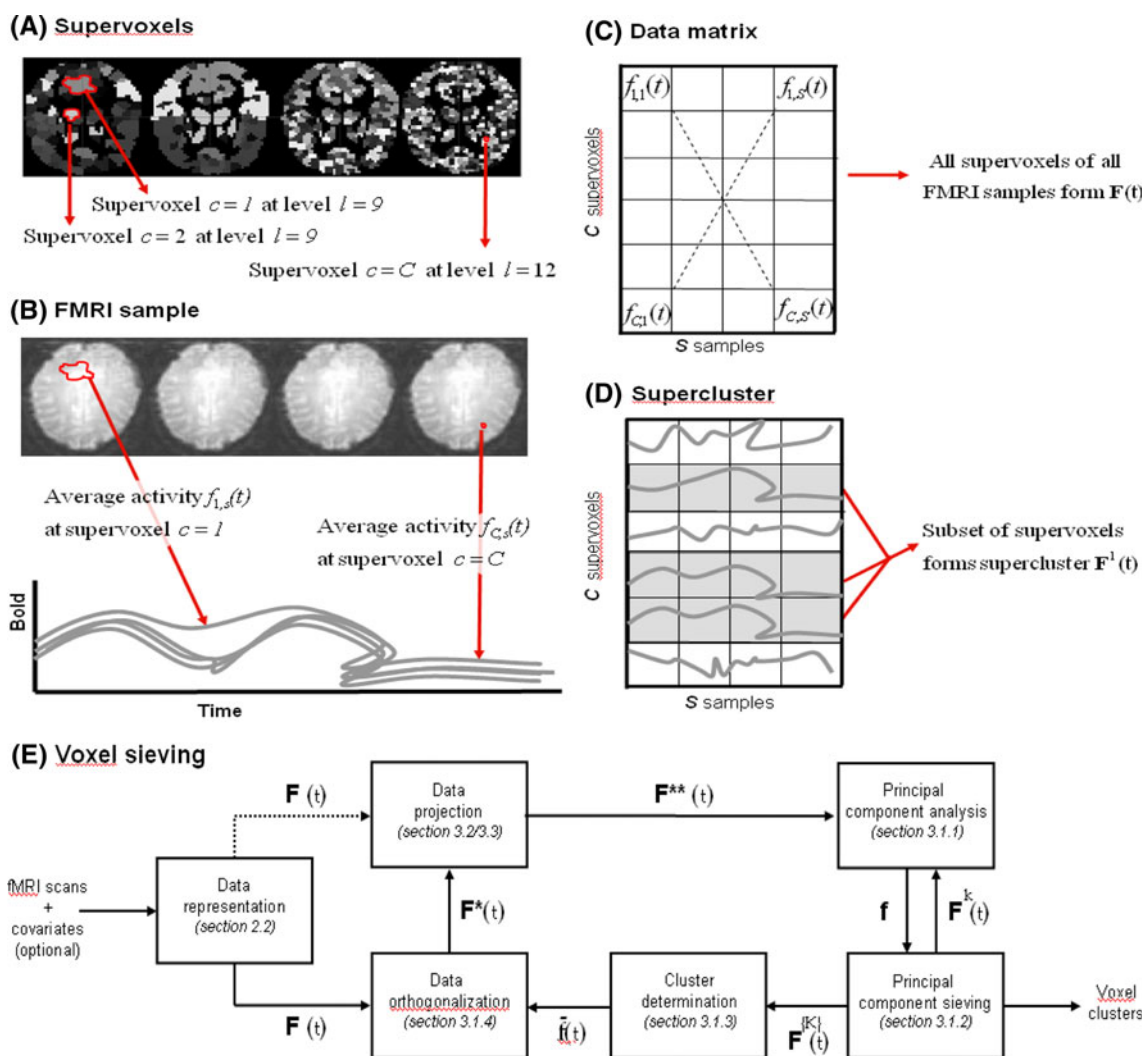
The main computational parts of the voxel sieving method are shown in Fig. 1e. Each of these components will be described in more detail in the following subsections.

#### 3.1 Unsupervised voxel sieving

Unsupervised voxel sieving operates directly on  $\mathbf{F}(t)$  (see Fig. 1c). It aims at identifying voxels with synchronous activity patterns independent of experimental conditions.

#### 3.1.1 Principal component analysis

The first task in voxel sieving is to find a subset of rows of  $\mathbf{F}(t)$  with both high column variance and high coherence between supervoxels (see Fig. 1d). A good way to accomplish this is to perform functional principal component analysis [31, 35] of  $\mathbf{F}(t)$  and to use principal component scores to identify rows of  $\mathbf{F}(t)$  that have high correlated variation. The central concept for the univariate functional data set  $\mathbf{f}(t) = [f_1(t), \dots, f_C(t)]$  is taking the linear combination



**Fig. 1** Schematic illustration of data representation and analysis. **a** Supervoxels are obtained through hierarchical spatial clustering of 3D anatomical atlas with quadratic scaling:  $2^l$ . From left to right:  $l = 9, 10, 11, 12$ . An fMRI sample  $I_s(\mathbf{x}, t)$  is represented in terms of its supervoxel’s average hemodynamic responses  $f_{c,s}(t)$ . **c** All supervoxels of all fMRI samples together form the data matrix  $\mathbf{F}(t)$ . **d** This

data matrix is subjected to voxel sieving to detect superclusters. **e** Voxel sieving is performed on  $\mathbf{F}(t)$  in an iterative way. Note that at the start fMRI data sets are required but external covariates are optional. In the absence of external covariates, fMRI data are projected onto themselves. Dashed line denotes a single step (at the start) while continuous lines indicate an iterative process

$$f_{cq} = \int_t f_c(t)\alpha_q(t)dt \tag{5}$$

where  $f_{cq}$  is the principal component score value of voxel time course  $f_c(t)$  in dimension  $q$ . Principal components  $\alpha_q(t)$ ,  $q = 1, \dots, Q$  are sought for one after the other by optimizing

$$\alpha_q(t) = \max_{\alpha_q^s(t)} \frac{1}{C} \sum_{c=1}^C f_{cq}^2 \tag{6}$$

where  $\alpha_q(t)$  is subject to the following orthonormal constraints:

$$\int_t \alpha_q(t)^2 dt = 1 \int_t \alpha_r(t)\alpha_q(t)dt = 0, \quad r \leq q. \tag{7}$$

The mapping of  $f_c(t)$  onto the subspace spanned by the  $Q$  first principal component functions results in the vector of principal component scores  $\mathbf{f}_c = [f_{c1}, \dots, f_{cQ}]$ . This mapping is very similar to local linear discriminant analysis of fMRI data (e.g. in [9, 28]). In this work, we only consider the main mode of variation, i.e. we set  $Q = 1$ .

As  $\mathbf{F}(t)$  is multivariate we need to perform multivariate functional principal component analysis (see [31]). The principal component in this case is defined by an  $S$ -vector of weight function  $\alpha = [\alpha_q^1(t), \dots, \alpha_q^S(t)]$  with  $\alpha_q^s(t)$  denoting the variation for sample  $s$ . The inner product on the space of vector functions is defined as the sum of the inner products of the  $S$  components. Hence, Eq. 5 becomes

$$f_{cq} = \sum_{s=1}^S \int_t f_{cs}(t)\alpha_q^s(t)dt. \tag{8}$$

In our case this amounts to concatenating the functional elements in each row of  $\mathbf{F}(t)$  to form a composite function. Subsequently, we perform univariate functional principal component analysis. This results in the principal component score vector  $\mathbf{f} = [f_1, \dots, f_C]$ , which is subjected to sieving.

### 3.1.2 Principal component sieving

Principal component sieving starts with the full data matrix  $\mathbf{F}(t)$ . The sieving procedure aims to remove  $\delta$  percent of the supervoxels, i.e. rows, of  $\mathbf{F}(t)$  with lowest absolute principal component scores, in order to arrive at a reduced data matrix  $\mathbf{F}^1(t)$ . The sieving parameter  $\delta$  allows to control for the graininess of sieving. When it has a low value, small clusters of voxels with strong synchronous activity can be detected (at the cost of computation). In contrast, larger voxel clusters with less hemodynamic synchrony emerge when the value of  $\delta$  is high.

We denote the set of supervoxels that survives the first sieving sequence by *supercluster*  $\mathcal{V}_1$  (note the difference between set of voxels denoted by  $\mathcal{V}$  and set of supervoxels

denoted by  $\mathcal{V}$ ). Then, functional principal component sieving is repeated on the reduced data matrix  $\mathbf{F}^1(t)$  to yield a new smaller supercluster. This process is repeated until the data matrix cannot be sieved anymore. Hence, voxel sieving results in superclusters  $\mathcal{V}_1 \supset \mathcal{V}_2, \dots, \supset \mathcal{V}_J$ , with  $I$  being the total number of sieving sequences. We denote the working matrix associated with the supercluster at sieving sequence  $j$  by  $\mathbf{F}^j(t)$ ,  $j = 1, \dots, J$ .

### 3.1.3 Cluster size determination

To distinguish real patterns from random small superclusters, we use the percentage of variance explained, R-statistic, as quality measure to select a supercluster from  $\mathcal{V}_1, \dots, \mathcal{V}_j$ . The R-statistic for a given supercluster  $\mathcal{V}_j$  is computed as the ratio between the variance  $V_B(t)$  and total variance  $V_T(t)$  defined as

$$V_B(t) = \frac{1}{S} \sum_{s=1}^S (\bar{f}_s^j(t) - \bar{f}^j(t))^2 \tag{9}$$

$$V_T(t) = \frac{1}{|\mathcal{V}_j| \times S} \sum_{c \in \mathcal{V}_j} \sum_{s=1}^S (f_{cs}^j(t) - \bar{f}^j(t))^2 \tag{10}$$

where  $\bar{f}^j(t)$  is the mean over all  $|\mathcal{V}_j| \times S$  elements of  $\mathbf{F}^j(t)$  and  $\bar{f}_s^j(t)$  is the  $s$ th column mean of  $\mathbf{F}^j(t)$ . A large value of  $R = \int (\sqrt{V_B(t)}/V_T(t))$  implies a tight cluster of coherent supervoxels.

We use Gap statistics [33] to select a reasonable cluster size based on randomization. Let  $\mathbf{F}^j(t)$  be the data matrix corresponding with sieving sequence  $j$  and  $R_j$  its R measure. To determine whether  $R_j$  is larger than expected by chance if the rows and columns of the data were independent, we permute the elements within each row of  $\mathbf{F}^j(t)$ . We perform  $P$  such permutations to obtain equally many R-measures. The Gap function is then defined by the difference between the real R-measure and the average R-measure of the randomized data

$$G(k) = R_j - \bar{R}_j^*. \tag{11}$$

The supercluster  $\mathcal{V}_j$  that produces the largest Gap is taken as the optimal cluster. The search for the next cluster is then performed on an orthogonalized version of the original matrix  $\mathbf{F}(t)$ .

### 3.1.4 Data orthogonalization

To find a new supercluster uncorrelated with the superclusters thus far, we perform orthogonalization of  $\mathbf{F}(t)$  with respect to the column average  $\bar{\mathbf{f}}(t)$  of the supercluster found in the previous step. This is equivalent to regressing each row of  $\mathbf{F}(t)$  on  $\bar{\mathbf{f}}(t)$  and replacing the rows with the regression residuals. As we are dealing with functional

data, we use a point-wise multivariate functional linear model to orthogonalize  $\mathbf{F}(t)$ . This reduces to solving

$$\mathbf{f}_c(t) = \bar{\mathbf{f}}(t)\beta(t) + \epsilon(t) \tag{12}$$

where  $\mathbf{f}_c(t)$ ,  $c = 1, \dots, C$  is a row vector of  $\mathbf{F}(t)$ ,  $\beta(t)$  is the regression function and  $\epsilon(t) = [\epsilon_1(t), \dots, \epsilon_S(t)]^T$  is the vector of residual functions. Under the assumption that the residual functions  $\epsilon(t)$  are independent and normally distributed with zero mean, the regression function is estimated by least squares minimization such that

$$\hat{\beta}(t) = \min_{\beta^*(t)} \int (\mathbf{f}_c(t) - \bar{\mathbf{f}}(t)\beta^*(t))^2 dt. \tag{13}$$

A roughness penalty is added to regularize the estimate of  $\beta(t)$ . We regularize the second derivative of  $\beta(t)$ . The estimated regression function provides the best estimate of  $\mathbf{f}_c(t)$  in least squares sense:

$$\hat{\mathbf{f}}_c(t) = \bar{\mathbf{f}}(t)\hat{\beta}(t). \tag{14}$$

The iterative voxel sieving process is then performed on the new data matrix  $\mathbf{F}^*(t)$  with rows

$$\mathbf{f}_c^*(t) = \mathbf{f}_c(t) - \hat{\mathbf{f}}_c(t). \tag{15}$$

That is, the data matrix for the next sieving operation is  $\mathbf{F}^*(t)$ . The search for the next supercluster starts with centering of the rows of  $\mathbf{F}^*(t)$ . Then all steps described in sect. 3.1 are repeated again on the new centered data matrix. This iterative procedure continues until a predefined number of superclusters has been identified. As the number of meaningful superclusters cannot be known a priori, the search for new superclusters may be stopped based on the quality of estimating voxel time course by a linear combination of supercluster averages: when adding new superclusters does not lead to increasing percent variance explained, this can be taken as a stop condition.

Note that because orthogonalization is done with respect to the average time course of a supercluster, supervoxels in different clusters can be highly correlated with one another. Moreover, one supervoxel can belong to multiple superclusters, i.e. supervoxels removed in a previous sieving step may be part of the supercluster of the next step.

### 3.2 Supervised voxel sieving

The method discussed so far has not used external information about the columns of  $\mathbf{F}(t)$  to ‘supervise’ the sieving of rows. External information such as cognitive states, subject information or stimulus patterns may be crucial in uncovering hidden hemodynamic synchrony. Here, we generalize voxel sieving to incorporate different types of external covariates such as continuously valued stimulus data or discrete class labels for the purpose of steering the discovery of hidden hemodynamic synchrony.

We consider the task of identifying synchronous brain activity directly related to a continuously valued stimulus patterns, for example, the expert movie ratings in the free movie viewing study. In a manner analogous to standard regression analysis, voxel sieving allows to search for supervoxels that best regress on expert movie ratings. To this end, we first define the stimulus function  $g_s(t)$  by fitting a B-spline to the vector-valued movie rating  $\mathbf{g}_s$ . For the task at hand, we subsequently map the supervoxels onto a subspace spanned by the movie rating data using the  $S \times S$  projection matrix

$$\mathbf{P}^1(t) = \mathbf{g}(t)\mathbf{g}^+(t) \tag{16}$$

where  $\mathbf{g}^+(t)$  is the generalized Moore-Penrose pseudo inverse of  $\mathbf{g}(t) = [g_1(t), \dots, g_S(t)]^T$ . Then, given data matrix  $\mathbf{F}(t)$  and projection matrix  $\mathbf{P}^1$  we map the supervoxels:

$$\mathbf{F}^{**}(t) = \mathbf{F}(t)\mathbf{P}^1(t). \tag{17}$$

Supervised data analysis now reduces to performing the computations described in sect. 3.1 on  $\mathbf{F}^{**}(t)$  rather than on  $\mathbf{F}(t)$ . Note that when the task at hand is to predict the stimulus from brain activity data (e.g. for brain reading tasks), we can reverse the roles of the predictor and the predictant, treating voxel activity data as the predictor and the stimulus as the response.

When alternatively the external information has discrete values or is coded with a label  $\mathcal{L}$  for each column of  $\mathbf{F}(t)$ , then an  $S \times |\mathcal{L}|$  matrix of scalars can be defined that maps the columns of  $\mathbf{F}(t)$  onto  $|\mathcal{L}|$  columns containing the class averages for each row. In the example of the three subjects watching six movie parts, we may, for example, want to identify synchronous brain activity across subjects using projection matrix

$$\mathbf{P}^2 = \begin{pmatrix} \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & 0 & 0 \\ & \ddots & \\ 0 & 0 & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} \end{pmatrix}. \tag{18}$$

Projection of  $\mathbf{F}(t)$  by  $\mathbf{P}^2$  results in an alternative  $C \times 3$  working matrix with the three columns now corresponding to the three subjects. The data analysis steps described in Sect. 3.1 are subsequently executed to identify across-subject hemodynamic synchronization.

Hence, incorporation of different types of external covariates in the voxel sieving procedure is achieved by performing a suitable data projection operation prior to the data analysis procedure of Sect. 3.1.

### 3.3 Partially supervised voxel sieving

Partially supervised sieving aims at striking a balance between supervised and unsupervised analysis so as to

encourage coherence within clusters, while allowing to exploit auxiliary information. This is particularly useful when dealing with overly aggressive supervision criteria. Given data matrix  $\mathbf{F}(t)$  and projection matrix  $\mathbf{P}$ , partially supervised data analysis is facilitated through

$$\mathbf{F}^{**}(t) = \mathbf{F}(t)\mathbf{P}^* \tag{19}$$

where  $\mathbf{P}^*$  is a weighted combination of the projection matrix  $\mathbf{P}$  and identity matrix  $\mathbf{I}$ :

$$\mathbf{P}^* = \lambda\mathbf{I} + (1 - \lambda)\mathbf{P}. \tag{20}$$

Parameter  $\lambda \in [0, \dots, 1]$  is a weight that allows to determine the extent of supervision. For  $\lambda = 1$ , the data are projected onto themselves and hence lead to unsupervised sieving. For  $\lambda = 0$ , the data are projected by  $\mathbf{P}$  only and thus analysis reduces to supervised sieving. Values between 0 and 1 enable partial supervision. Note that  $\mathbf{P}$  and  $\mathbf{I}$  become matrices of functions when the external covariate itself is functional.

#### 4 Experiments and results

We use voxel sieving to uncover distributed and overlapping patterns of fMRI activity predicative of sources underlying these patterns. Our experiments aim at exploring how well this can be achieved. All experiments are performed on a functional data representation of the fMRI data. An important motivation for using B-splines, rather than temporal smoothing with an HRF kernel, is minimization of bias. To what extent a predefined kernel smoother gives an acceptable level of bias can only be determined empirically. We choose to determine the smoother in a more objective manner by calculating smooth splines for our time courses with roughness of derivatives as a penalty [31]. We subsequently determine the minimum number of basis functions producing very similar smoothing results, to get an efficient yet accurate data representation. Note that this generally imposes some restriction on variation in fMRI scan length, repetition times, etc. The fMRI scans in our experiments are reasonably uniform in terms of number of volumes and hence can all be approximated with the same number of basis functions.

##### 4.1 Simulated fMRI data

As an initial test we apply our method to artificial fMRI data. Following [6, 8], we simulate fMRI data using three types of sources: task-related, transiently task related, and function related. The task-related source corresponds with an activation paradigm. It is periodic and slowly changing. The transiently related source closely matches the task-related source but has an activation that is more pronounced at parts

of each task cycle. The function-related source is characterized by random fluctuations. The three sources are super-Gaussian in nature; they are localized. We disregard source variations across large image areas such as motion-related sources, assuming these have been accounted for in the preprocessing step.

We convolve sources with variations of hemodynamic response functions observed across subjects [16] to mimic across-subject variation. In this way, we construct three different fMRI sets representing three scan samples ( $S = 3$ ). Each of the three fMRI data sets consists of 64 x 64 voxels and 100 time points. Approximately 22% of these 4,096 voxels has a task-related source. These voxels are clustered at three spatially distributed locations. Another 15% has a transiently task-related source, distributed over two equally large clusters. The fraction of voxels with a mixture of the aforementioned sources is 7%. Finally, a random sources is assigned to 5 percent of the voxels. We add Gaussian noise to the constructed data sets at signal-to-noise ratios (SNR): 2, 1.5, 1, 0.5, 0.25. The SNR measure we use is the standard deviation over all sources divided by the standard deviation over all noise sources. Figure 2 summarizes the sources.

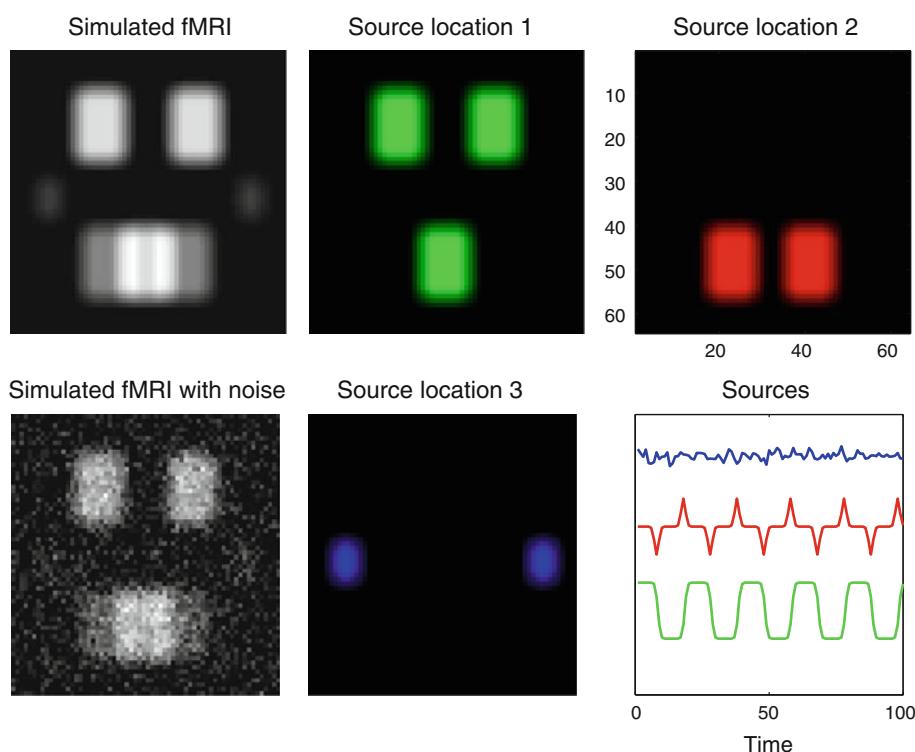
We fit a 20-coefficient B-spline to the discrete voxel time courses to obtain functional data. The voxel time courses are hierarchically clustered in space. The highest level used in hierarchical clustering was  $l = 5$ . It produces  $2^5 = 32$  voxel clusters with on average 128 voxels. We excluded higher levels because we expect these will not be informative. At the lowest level the  $2^{12} = 4,096$  individual voxels themselves are considered. The supervision weight  $\lambda$  is set to 0 (fully supervised) or 1 (fully unsupervised). Data randomization to separate real from random clusters is done on the basis of  $P = 3$  permutations.

In evaluating our method we make a distinction between relevant and irrelevant voxels. Relevant voxels have a task-related source, potentially mixed with an other source. All other voxels are irrelevant. The aim is to detect the relevant voxels precisely and completely by sieving. Detection results are confined to the first two superclusters. We use *Precision* and *Recall* to measure performance. *Precision* indicates the fraction of relevant voxels in the two detected superclusters, while *Recall* is the number of relevant voxels in the two detected superclusters divided by the number of relevant voxels in the entire fMRI volume. The harmonic mean combines these two measures into a single one [34]:

$$F_{score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{21}$$

Table 1 shows the  $F_{score}$  for supervised ( $\lambda = 1$ ) and unsupervised analysis ( $\lambda = 0$ ) of simulated fMRI data for various signal-to-noise ratios and values of sieving parameter  $\delta$ . We first discuss two noteworthy observations

**Fig. 2** Simulated fMRI data are a linear mixture of three independent sources at multiple spatially distributed locations. The task-related source (location 1) is shown in *green* (lower source signal), the transiently task-related source (location 2) in *red* (middle source signal), and the random source (location 3) in *blue* (upper source signal). Note that in the simulated fMRI slice, the *gray* level is highest where the mixing of the first two sources occurs (color figure online)



**Table 1**  $F_{score}$  of detection for: unsupervised | supervised sieving

$\delta$	SNR = 2.0	SNR = 1.5	SNR = 1.0	SNR = 0.5	SNR = 0.25
0.1	0.78   0.81	0.79   0.72	0.59   0.67	0.44   0.40	0.27   0.25
0.3	0.65   0.68	0.52   0.55	0.48   0.47	0.35   0.29	0.28   0.19
0.5	0.50   0.51	0.41   0.44	0.36   0.35	0.27   0.17	0.23   0.11
0.7	0.33   0.32	0.23   0.20	0.19   0.16	0.17   0.09	0.11   0.05
0.9	0.20   0.18	0.18   0.12	0.13   0.10	0.09   0.08	0.09   0.06

for both unsupervised and supervised sieving. First, lower values of  $\delta$ , yielding fine-grained voxel clusters, lead to better detection performance. This is to be expected as the relevant voxels are clustered in relatively small parts of the 3D space. In real fMRI scans, where sources may be spread over the entire space, larger values of  $\delta$  may perform better (as we will discuss next). Second, the decay of detection performance with decrease of SNR is lower for larger values of the sieving parameter. This can be explained by the fact that course sieving results in larger voxel clusters that tend to average out noise more vigorously.

In comparison to unsupervised sieving, supervised sieving performs better at high SNR when sieving is done in a fine-grained fashion (low  $\delta$  values). A close inspection of detected superclusters reveals the following recurring pattern. In supervised analysis, the first supercluster is large relative to the second and is almost entirely composed of task-related voxels. The second supercluster is small and includes voxels with a mix of task-related and transiently

task-related time courses. As a result *Precision* is very high. Conversely, in unsupervised analysis, the first cluster and the second supercluster are relatively large and comparable in size. Almost all voxels in the first cluster are task-related. The second supercluster also contains a considerable amount of irrelevant clusters. This leads to lower *Precision* and higher *Recall*, compared with supervised analysis. On average, detection performance reduces for unsupervised analysis. When the signal-to-noise ratio decreases and becomes more realistic, however, unsupervised sieving outperforms supervised sieving, particularly for course sieving (higher values of  $\delta$ ). Overall, these results indicate that voxel sieving is capable of identifying localized synchrony in hemodynamics at multiple levels of granularity, using covariate information in a flexible manner as a pilot.

#### 4.2 Real fMRI data

Our experiments with real data involve fMRI data acquired during a free movie viewing study involving Home Improvement sitcoms. With these experiments we aim to explore the spatial nature of detected voxel clusters under a variety of source-specific conditions. Second, we test the ability of these voxel clusters to predict natural sensory stimuli, i.e. to do brain reading. Hence, we test whether we localize brain regions containing information about the external sources, rather than testing for brain regions that activate with the external sources.

Functional data for the real fMRI data sets have been obtained by fitting a 30-coefficient B-spline to the discrete data points, both for voxel activation and stimulation data. The highest level used in hierarchical clustering was  $l = 9$  (see Fig. 1a). It produces  $2^9 = 512$  voxel clusters with an average size of 70 voxels, while the lowest level of  $l = 12$  produces  $2^{12} = 4,096$  voxel clusters with an average size of 9 voxels. In sect. 4.2.2 we discuss how we selected values of  $l = 9, \dots, 12$  to limit the search space to  $C = 2^9 + 2^{10} + 2^{11} + 2^{12} = 7,680$  supervoxels and speed up the search. Hence, the  $C \times S$  data matrix  $\mathbf{F}(t)$  consists of  $C = 7,680$  supervoxels for  $S = 18$  scan samples. The sieving parameter  $\delta$  as described above was set to 0.2. This setup requires 24 computation hours on a standard desktop computer. The parameter  $\lambda$  for controlling supervision was varied between  $[0, 1]$  depending on the experiment. Data randomization to separate real from random clusters was done on the basis of  $P = 5$  permutations.

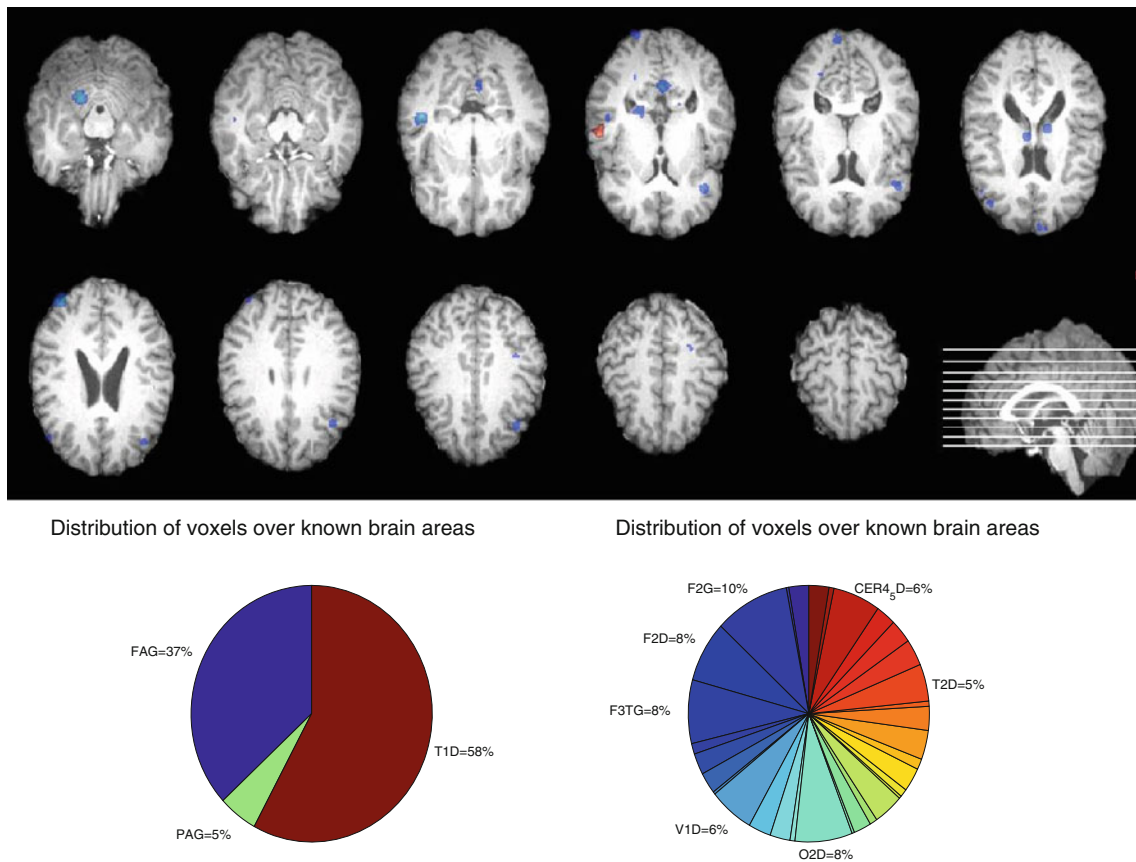
We first describe application of voxel sieving for identification of brain areas reacting in synchrony across brain scans in an unsupervised manner. Then we elaborate on

supervised voxel sieving for finding across-subject hemodynamic synchrony. An interparticipant correlation map is created to compare our findings with that of an intersubject similarity-based method ([22]).

#### 4.2.1 Interscan synchronization

Unsupervised analysis of the fMRI data implies  $\lambda = 1$ . In this case, the projection matrix  $\mathbf{P} = \mathbf{I}$ . Voxel sieving thus performs a data-driven search for voxel activity patterns with high across-scan variance and high across-voxel coherence. The resulting voxels highlight parts of the brain that act in synchrony during natural movie viewing.

The first row of Fig. 3 shows the first two superclusters overlaid over the anatomical image of subject 1. Voxels from the first supercluster are depicted in red and voxels from the second are given in blue. Voxel color value indicates the degree of match between the voxels activity pattern and the first principal component of the supercluster. The brighter the color the stronger the match between a voxel's activity pattern and the first component of its



**Fig. 3** Across scan synchronization. *First row* first two superclusters (red and blue) in two different colors overlaid over the anatomical image of subject 1. The first supercluster consists of 19 voxels and the second has 299 voxels. *Second row* functional distribution of

identified voxels in the first (left pie) and second (right pie) supercluster according to the MNI brain atlas ([26], see Fig. 10 for the labels). The percentage is only shown for 5 or higher (color figure online)

supercluster. Note that for other subjects the same voxel locations are highlighted in color, but with different intensities because of the voxel's unique activation pattern.

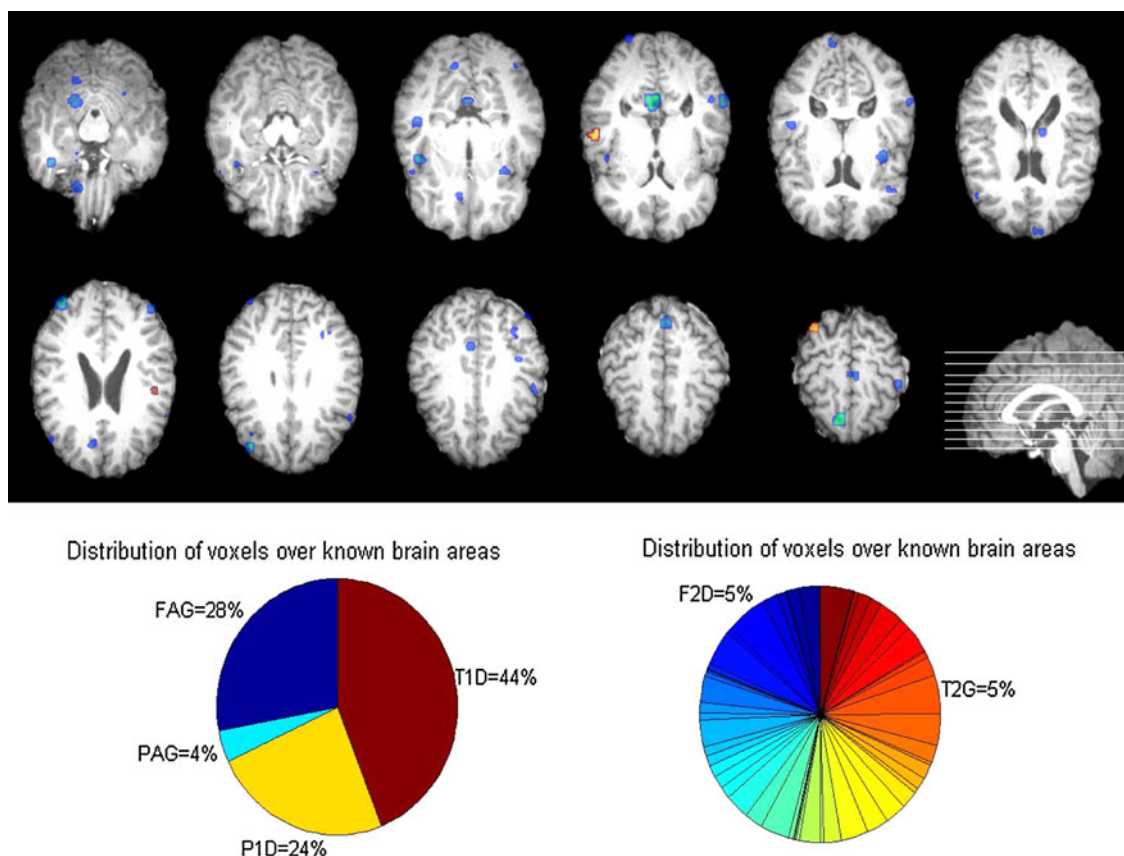
From the Gap statistics at each sieving sequence, it follows that for the first supercluster, the largest Gap occurs when only two supervoxels remain, consisting of 19 voxels. The second supercluster has 27 supervoxels with a total of 299 voxels. All supervoxels are at levels  $l = 11$  or  $l = 12$ . Note that these levels are automatically selected from the available levels by our method. We examined the spatial distribution of individual voxels over known functional areas. The pie chart in Fig. 3 shows that the voxels in the first supercluster are mostly localized in functional areas for motor and action, while voxels in the second cluster are distributed over a wide range of functional areas. We speculate that during passive movie viewing, hemodynamic synchrony is strongly present at brain areas for motor and action.

#### 4.2.2 Intersubject synchronization

Fully supervised fMRI analysis allows to identify brain areas with the strongest across-subject synchronization

during the viewing of the sitcoms. This reduces to setting  $\lambda = 0$  and consequently activating  $\mathbf{P}^2$  in Eq. 19. Projection of the data matrix  $\mathbf{F}(t)$  by  $\mathbf{P}^2$  and incrementally sieving away supervoxels identifies the voxels highlighted in the first row of Fig. 4. The first supercluster in red contains three supervoxels. Almost half of the 25 individual voxels is located at the temporal lobe where audio processing takes places. The second cluster in blue contains 54 supervoxels (578 voxels). Again all supervoxels are at levels  $l = 11$  or  $l = 12$ . Across-subject synchronization is identified at multiple areas across the entire brain. Notice that very specific brain areas are visible with a very strong synchrony rather than a widespread cortical activation pattern as reported in a similar natural movie viewing study ([17]). These specific results are typical of voxel sieving and provide additional insight into correlates of natural movie viewing.

We further compared voxel sieving to interparticipant correlation analysis [22] in terms of the spatial distribution and size of detected voxel clusters. Note that, instead of computing correlation maps on the basis of voxel activity patterns, we used supervoxels to determine such maps. This way we are less sensitive to the multiple comparison



**Fig. 4** Intersubject synchronization. *First row* first two superclusters (red and blue) in two different colors overlaid over the anatomical image of subject 1. *Second row* functional distribution of identified voxels according to MNI brain atlas ([26], see Fig. 10 for the labels) (color figure online)

problem. We computed correlation maps for each of the six movie parts, then averaged these to obtain a single average correlation map. The average interparticipant correlation map was thresholded to identify voxels with highest correlation. The threshold was selected such that approximately 500 voxels remained to ease comparison with the above voxel sieving result. These voxels were then separated into two groups with k-means clustering by temporal similarity. The first cluster contains 11 supervoxels consisting of 239 voxels, while the second has 19 supervoxels comprised of 385 voxels as shown in Fig. 5. The number and spatial distribution of identified voxels differs clearly from voxel sieving results. The voxel-by-voxel correlation analysis apparently favors larger brain areas of synchronous brain activity, whereas the multivariate clusterwise approach of voxel sieving detects specific brain areas with strong correlations and large variation.

### 4.3 Localization and prediction

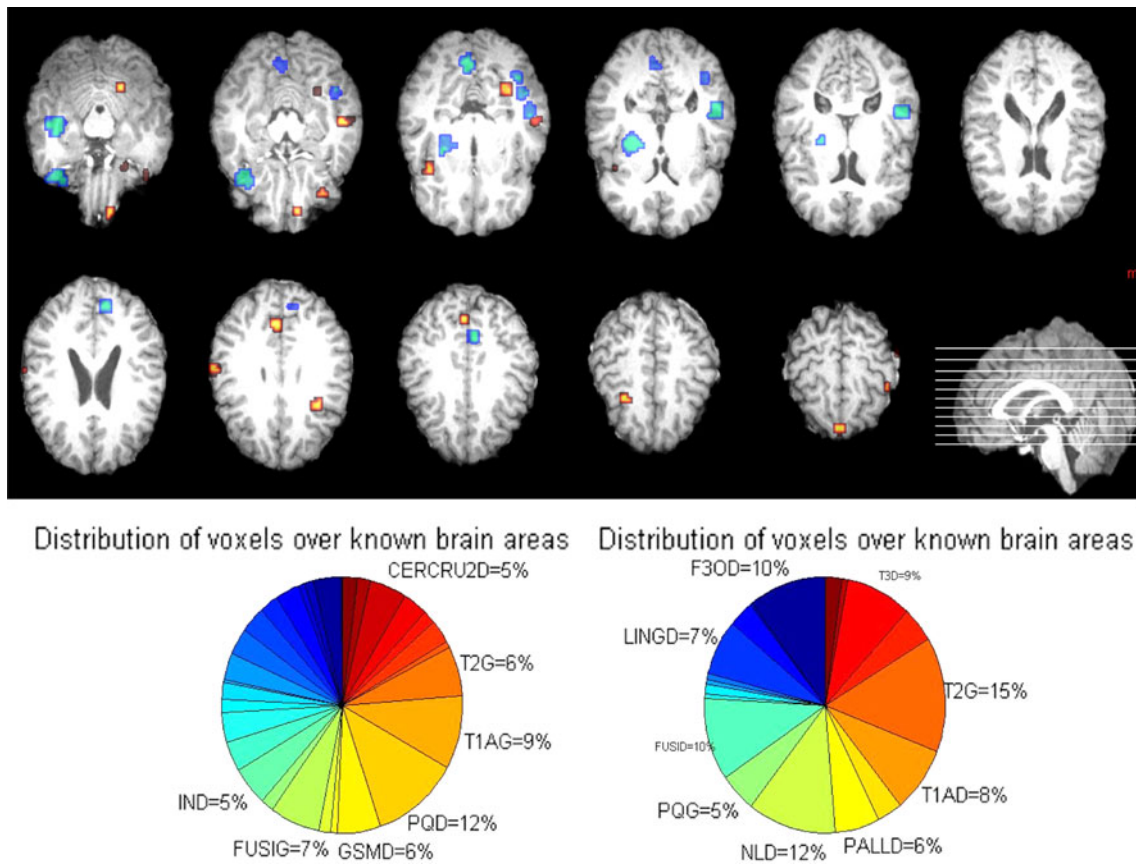
We now consider the task of localization of covariate-related brain responses. We analyze the fMRI data under

full and partial supervision. Then we concentrate on predicting external covariates on the basis of fMRI data.

#### 4.3.1 Localization

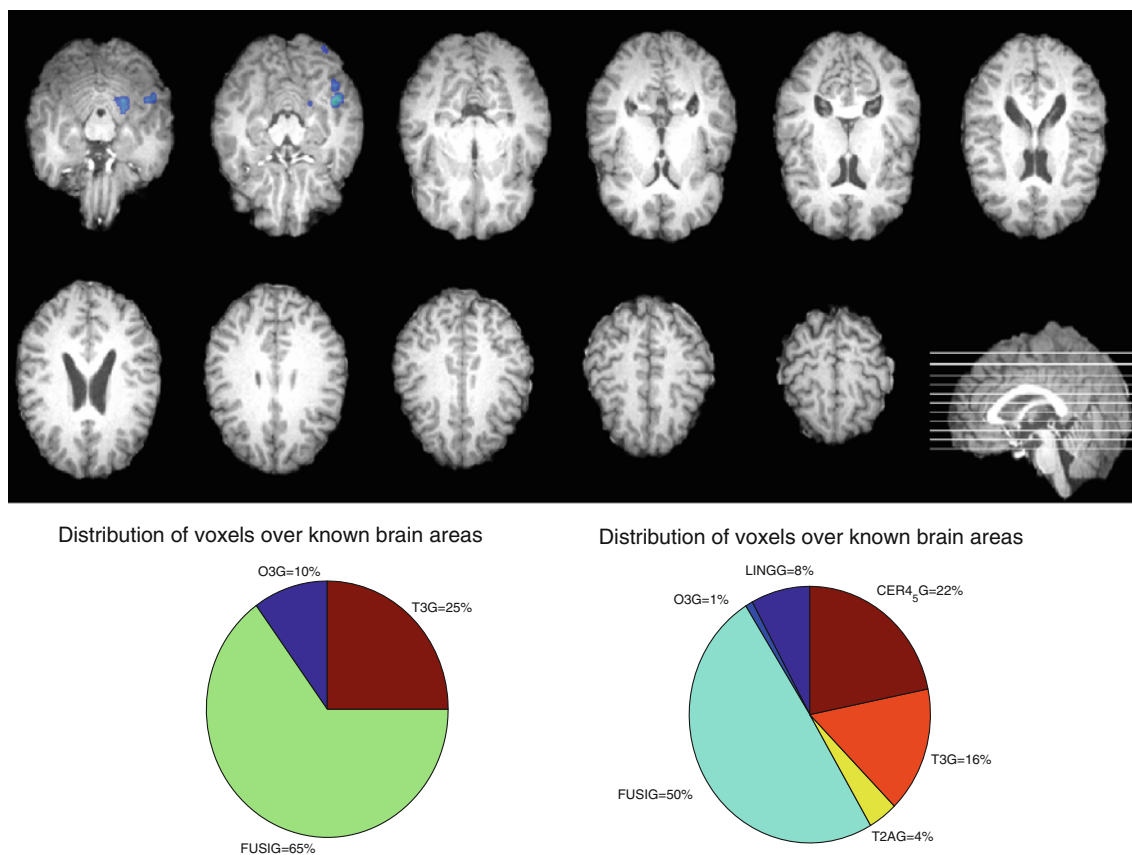
The projection matrix  $\mathbf{P}^1(t)$  in Eq. 17 forms the basis for localization of covariate-related brain responses. In a full supervision mode, i.e.  $\lambda = 0$ , sieving is performed on the matrix  $\mathbf{F}(t)\mathbf{P}^1(t)$ . This is the equivalent to standard regression analysis. The aim is to find rows of  $\mathbf{F}(t)$  with column means that best regress on the external covariates. However, rather than performing regression voxel-wise or volume-wise, it is here performed on clusters of voxels. This has the benefit of allowing to find multiple specific voxel clusters that are independently related to the stimulus. Furthermore, supervoxels eliminate the need for spatial regularization.

Figure 6 shows the first two clusters of voxels that best explain the *face* stimulus. For the first cluster the difference between the real explained variance and the randomized explained variance occurs at the last sieving sequence, corresponding to 2 supervoxels. The second supercluster



**Fig. 5** Average interparticipant correlation map superimposed on the fMRI of subject 1. *First row* two detected clusters with similar temporal similarity. *Color* values in this case correspond with the average Pearson correlation coefficient. *Second row* functional

distribution of identified voxels according to MNI brain atlas ([26], see Fig. 10 for the labels). The percentage is only shown for 5 or higher (color figure online)



**Fig. 6** Localization with  $\lambda = 0$ . *First row* first two superclusters (red and blue) in two different colors overlaid over the anatomical image of subject 1. *Second row* functional distribution of identified voxels according to MN1 brain atlas (color figure online)

contains 12 supervoxels. The majority of the individual voxels of both superclusters is located in the left fusiform area, which is known to be involved in face processing [15]. The other identified functional areas associated with the *face* stimulus are temporal inferior lobe, left cerebellum, and left lingual. Almost all of these functional areas are involved in language processing; It is conceivable that these areas activate when perceiving human faces. Note that there is a lot of (spatial) overlap between the two superclusters. The first supercluster in fact is a subset of the second, possibly indicating functional specialization.

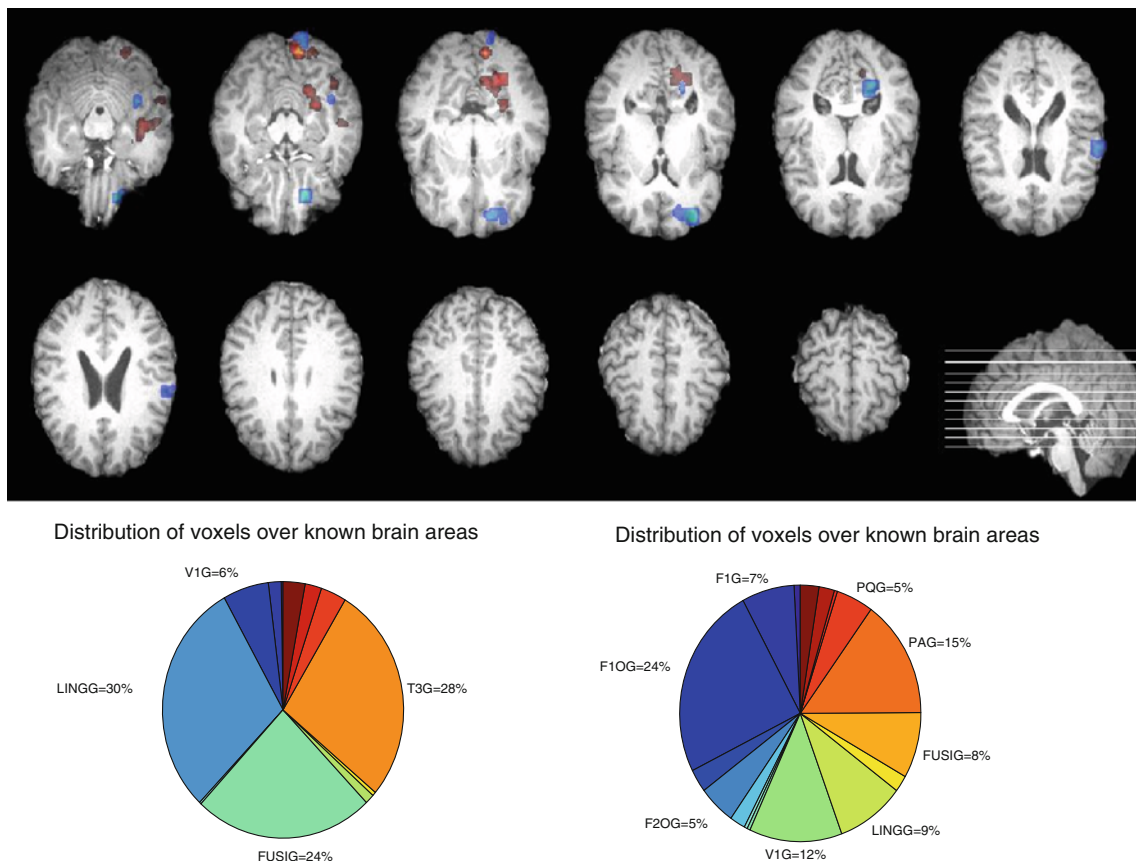
The first row of Fig. 7 shows results of partially supervised sieving with supervision weight  $\lambda = 0.5$ . In this case, the supervision criteria is less rigid, providing more room for identifying transient brain activity related to the *face* stimulus. The first supercluster contains 34 supervoxels. The second supercluster has 27 supervoxels, mostly at higher levels of hierarchy ( $l \in [10, 11]$ ). The individual voxels are found at a broader range of spatial and functional areas. Most voxels are found in the following areas: fusiform, temporal inferior lobe, left cerebellum, and left lingual. This gives reason to believe that next to voxels that

are directly related to the stimulus many more are transiently related.

#### 4.3.2 Prediction

Evaluation of detected brain responses to naturalistic stimuli, as in our case, is difficult because of lack of appropriate reference material. One way of dealing with this challenge is to invert the task from correlating external covariates with fMRI data to predicting these covariates from the fMRI data. This makes evaluation of detected brain responses more objective [21]. Here, we use partially supervised voxel sieving to uncover voxels that are predictive of the *face* stimulus in our movie data. We concentrate on the *face* stimulus because of the large body of reference material [15].

For various values of  $\lambda$  we identify two clusters that we subsequently use as predictors in a functional linear model (see [31] for more detail), with the stimulus as dependent variable and the cluster averages as independent variables, i.e. predictors. In the training phase the best model is selected: a model with one or two predictors. The trained



**Fig. 7** Localization with  $\lambda = 0.5$ . *First row* first two superclusters (red and blue) in two different colors overlaid over the anatomical image of subject 1. *Second row* functional distribution of identified voxels according to MNI brain atlas (color figure online)

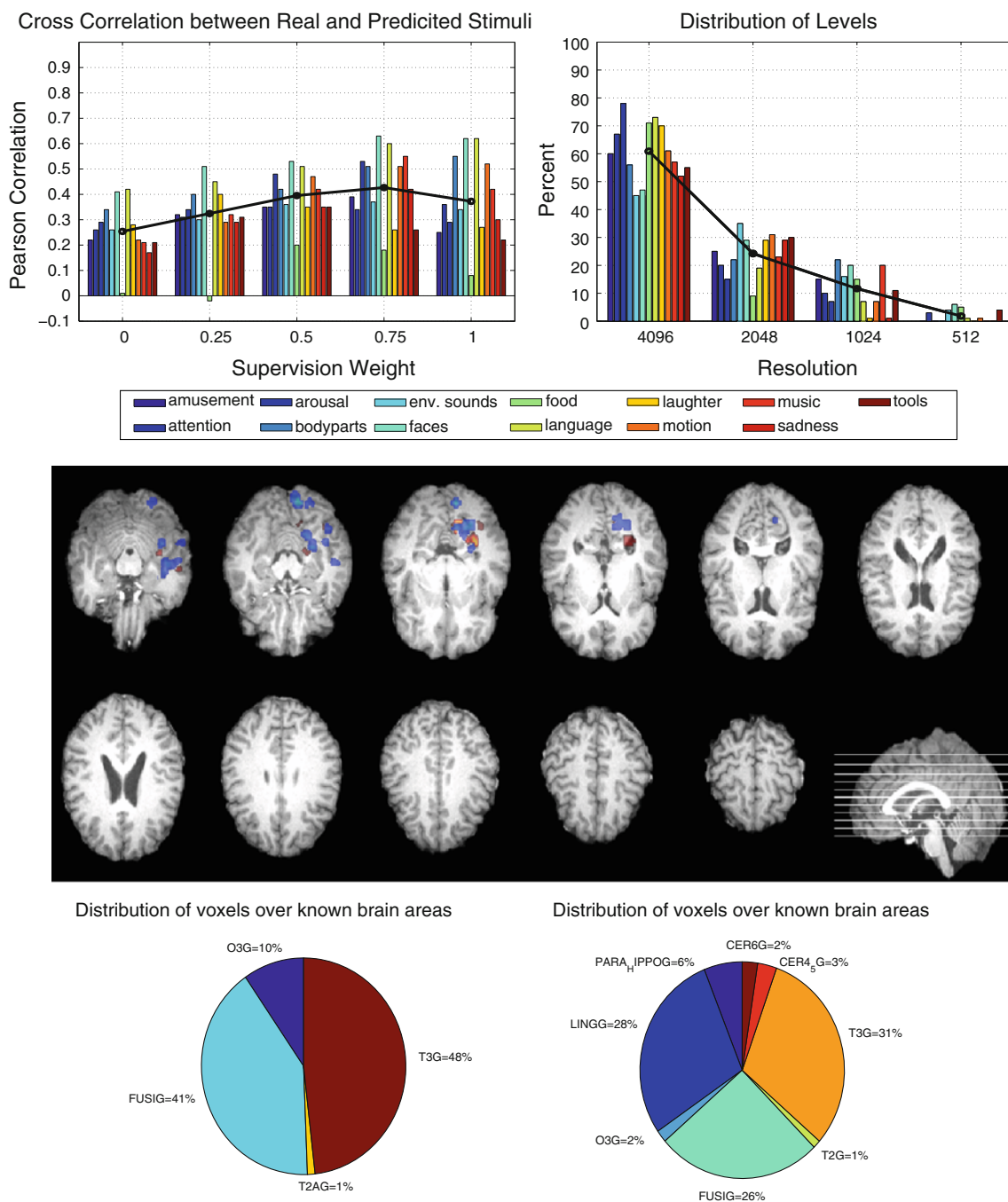
model is then applied in the testing phase on independent data to predict a feature. We use movie 1 data for training and movie 2 data for prediction, and vice versa. Pearson correlation coefficient between manual feature rating functions and the automatically predicted feature functions was used as an evaluation measure.

Prediction results are summarized in the first row of Fig. 8. Shown is the average of  $2 \times 18$  cross correlation values from cross validation for all 13 movie features with supervision weight  $\lambda$  set to 0, 0.25, 0.5, 0.75 and 1. The gross pattern of the graphs shows that prediction performance reaches a maximum around  $\lambda = 0.75$ . This indicates that brain activity patterns that are transiently related to the stimulus are relevant for prediction. The highest cross correlation value of 0.62 is for feature *faces* for  $\lambda = 0.75$ . The second row of Fig. 8 shows the voxels that have been used for prediction of this feature, with the first supercluster containing 5 supervoxels and the second supercluster 12. As expected, most voxels are localized in brain areas related directly or indirectly to face processing.

The first row of Fig. 8 also shows the distribution of cluster resolutions that were used for prediction. Most of the identified voxel clusters are at the lowest hierarchical level, i.e. have cluster size of approximately nine voxels.

Some features such as *environmental sounds*, however, also benefit from supervoxels at higher levels of hierarchy, suggesting that some features are processed more globally than other ones. We note that we restricted our supervoxels to only four hierarchical levels, as these levels performed best in a prediction experiment where we started with supervoxels at the lowest level ( $l = 12$ ,  $C = 4,690$ ) and stepwise included higher levels. Prediction performances for all features and for supervision weight  $\lambda$  set to 0, 0.25, 0.5, 0.75, 1, increased steadily up to level  $l = 9$ . Beyond this level performance first remained stable and then reduced. Hence, at least for the prediction task, a multiresolution approach pays off.

We compared voxel sieving performance with that of the three winning entries of the 2006 EBC Brain Reading competition. These entries used recurrent neural networks, ridge regression, and a dynamic Gaussian Markov Random Field modeling on the same test data benchmark, yielding across feature average cross correlations of 0.49, 0.49, and 0.47, respectively. For the voxel sieving method, the feature average cross correlation value is 0.44. This is good considering that the predictions are based on a reduced data set, while the reported results of the winning entries are based on the full data set. The fact that we have used a



**Fig. 8** fMRI-based stimuli prediction. *First row* average cross correlation values for all 13 movie features and 5 supervision weights. *Right* distribution of resolutions of supervoxels used for prediction. Note that one or a combination of predictors is used

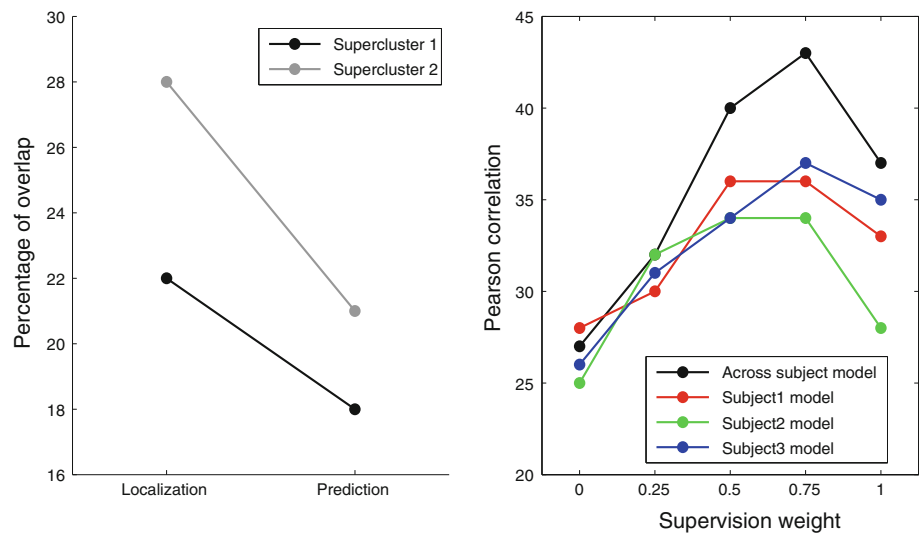
depending on the best prediction outcome. *Second row* first two superclusters (*red* and *blue*) overlaid over the anatomical image of subject 1. *Third row* functional distribution of identified voxels according to MNI brain atlas (color figure online)

smaller training set is likely to have had a negative impact on the prediction results. Note, that in the 2006 competition our entry, an initial version of the voxel sieving method, ranked first in the actor category [32]. We were able to accurately predict which actor the subjects were seeing purely based on fMRI scans [10].

### 4.3.3 Consistency

In order to check for consistency of the voxel detections across subjects, we repeated the localization ( $\lambda = 0$ ) and prediction experiments ( $\lambda = \{0, 0.25, 0.5, 0.75, 1\}$ ) three times, each time using only fMRI data of a single subject

**Fig. 9** *Left* Consistency of voxel cluster detections across subjects differs between tasks and superclusters. Consistency is reasonably high for the second supercluster in the localization task compared with consistency for the first supercluster in the prediction task. *Right* Prediction results based on fMRI data of individual subjects drops significantly for larger values of the supervision weight. This shows the importance of across subject synchrony



instead of all three. We measured the overlap in supervoxels across subjects in terms of the number of supervoxels that were in the superclusters of all three subjects relative to the total number of supervoxels. We examined consistency separately for the superclusters. Figure 9a shows the results of consistency analysis.

When we only consider supervoxels in the first supercluster, almost 22 percent of the supervoxels from subjects 1, 2, and 3 overlap in the localization task. For the second supercluster the overlap is significantly higher: 28%. We attribute this difference in consistency between the first and second superclusters to number and spatial size of supervoxels, which tend to be larger for the second supercluster. In the prediction task, we computed consistency separately for supervision weights 0, 0.25, 0.5, 0.75, 1 and subsequently averaged these. For larger values of the supervision weight, the supervoxels detected are generally few, spatially confined and variable across subjects, adversely affecting consistency of voxel detections across subjects. The amount of overlap drops for both superclusters to 21 and 18%, respectively. This, however, does not necessarily imply that a source-specific search for hemodynamic synchrony yields more consistency than unbiased probing. It might be that consistency emerges with across subject analysis as reflected in the prediction results based on fMRI data of individual subjects (Fig. 9b). Considering the large amount of supervoxels ( $C = 7,680$ ), the obtained results indicate a reasonable consistency of voxel detections across subject.

### 5 Discussion

We have introduced a statistical signal analysis method for identification of distributed and overlapping synchronous hemodynamic patterns that are directly or transiently

linked to their underlying sources. The method is applicable for brain activity from any modality and covariates of any form. We focused on fMRI data from a free natural movie viewing study, as these data generally contain complex distributed and overlapping synchronous hemodynamic patterns. Our experiments showed that voxel sieving is very effective in uncovering both anticipated (visual and auditory regions) and unexpected cortical areas involved in face processing (such as motor and action regions). The viability of voxel sieving to find established or discoverable relations also holds for the other movie stimuli in our data set. There is generally a meaningful relation between cognitive concepts from the movie stimuli and synchronously active brain areas as identified by voxel sieving. The performance of voxel sieving in fMRI-based prediction of the movie stimuli strongly supports the significance of exposed brain areas.

Voxel sieving can be conceived of as a superset of many existing fMRI data analysis methods. When a single cluster is searched for in a fully unsupervised mode ( $\lambda = 1$ ) without sieving ( $\delta = 1$ ), our method reduces to functional principal component analysis [35]. Independent component analysis [4, 11] is approximated when multiple independent clusters are searched using  $\lambda = 1$  and  $\delta = 1$ . In a fully supervised mode ( $\lambda = 0$ ) using projection matrix  $\mathbf{P}^l(t)$ , fMRI data are analyzed in a manner analogous to standard regression analysis ([13]). Other projections matrices can be used, for example, for discriminating activity between subjects, groups or conditions (similar to [3, 29, 36]). By varying the values of sieving parameter  $\delta$  and levels of hierarchy  $l$ , the method enables voxel-wise, cluster-wise, and volume-wise data analysis. In addition, as voxel sieving relies heavily on data averaging and dimension reduction on a data sets from multiple subjects or multiple conditions, it is reasonably robust against multiple testing

'FAG'	'Precentral_L'	'PARA_HIPPOD'	'ParaHippocampal_R'	'HESCHLG'	'Heschl_L'
'FAD'	'Precentral_R'	'AMYGDG'	'Amygdala_L'	'HESCHLD'	'Heschl_R'
'FIG'	'Frontal_Sup_L'	'AMYGDD'	'Amygdala_R'	'T1G'	'Temporal_Sup_L'
'F1D'	'Frontal_Sup_R'	'V1G'	'Calcarine_L'	'T1D'	'Temporal_Sup_R'
'F1OG'	'Frontal_Sup_Orb_L'	'V1D'	'Calcarine_R'	'T1AG'	'Temporal_Pole_Sup_L'
'F1OD'	'Frontal_Sup_Orb_R'	'QG'	'Cuneus_L'	'T1AD'	'Temporal_Pole_Sup_R'
'F2G'	'Frontal_Mid_L'	'QD'	'Cuneus_R'	'T2G'	'Temporal_Mid_L'
'F2D'	'Frontal_Mid_R'	'LINGG'	'Lingual_L'	'T2D'	'Temporal_Mid_R'
'F2OG'	'Frontal_Mid_Orb_L'	'LINGD'	'Lingual_R'	'T2AG'	'Temporal_Pole_Mid_L'
'F2OD'	'Frontal_Mid_Orb_R'	'O1G'	'Occipital_Sup_L'	'T2AD'	'Temporal_Pole_Mid_R'
'F3OPG'	'Frontal_Inf_Oper_L'	'O1D'	'Occipital_Sup_R'	'T3G'	'Temporal_Inf_L'
'F3OPD'	'Frontal_Inf_Oper_R'	'O2G'	'Occipital_Mid_L'	'T3D'	'Temporal_Inf_R'
'F3TG'	'Frontal_Inf_Tri_L'	'O2D'	'Occipital_Mid_R'	'CERC1G'	'Cerebehum_Crus1_L'
'F3TD'	'Frontal_Inf_Tri_R'	'O3G'	'Occipital_Inf_L'	'CERC1D'	'Cerebehum_Crus1_R'
'F3OG'	'Frontal_Inf_Orb_L'	'O3D'	'Occipital_Inf_R'	'CERC2G'	'Cerebehum_Crus2_L'
'F3OD'	'Frontal_Inf_Orb_R'	'FUSIG'	'Fusiform_L'	'CERC2D'	'Cerebehum_Crus2_R'
'ORG'	'Rolandic_Oper_L'	'FUSID'	'Fusiform_R'	'CER3G'	'Cerebehum_3_L'
'ORD'	'Rolandic_Oper_R'	'PAG'	'Postcentral_L'	'CER3D'	'Cerebehum_3_R'
'SMAG'	'Supp_Motor_Area_L'	'PAD'	'Postcentral_R'	'CER4_5G'	'Cerebehum_4_5_L'
'SMAD'	'Supp_Motor_Area_R'	'P1G'	'Parietal_Sup_L'	'CER4_5D'	'Cerebehum_4_5_R'
'COBG'	'Olfactory_L'	'P1D'	'Parietal_Sup_R'	'CER6G'	'Cerebehum_6_L'
'COBD'	'Olfactory_R'	'P2G'	'Parietal_Inf_L'	'CER6D'	'Cerebehum_6_R'
'FMG'	'Frontal_Sup_Medial_L'	'P2D'	'Parietal_Inf_R'	'CER7BG'	'Cerebehum_7b_L'
'FMD'	'Frontal_Sup_Medial_R'	'GSMG'	'SupraMarginal_L'	'CER7BD'	'Cerebehum_7b_R'
'FMOG'	'Frontal_Med_Orb_L'	'GSMD'	'SupraMarginal_R'	'CER8G'	'Cerebehum_8_L'
'FMOD'	'Frontal_Med_Orb_R'	'GAG'	'Angular_L'	'CER8D'	'Cerebehum_8_R'
'GRG'	'Rectus_L'	'GAD'	'Angular_R'	'CER9G'	'Cerebehum_9_L'
'GRD'	'Rectus_R'	'PQG'	'Precuneus_L'	'CER9D'	'Cerebehum_9_R'
'TNG'	'Insula_L'	'PQD'	'Precuneus_R'	'CER10G'	'Cerebehum_10_L'
'TND'	'Insula_R'	'LPCG'	'Paracentral_Lobule_L'	'CER10D'	'Cerebehum_10_R'
'CIAG'	'Cingulum_Ant_L'	'LPCD'	'Paracentral_Lobule_R'	'VER1_2'	'Vermis_1_2'
'CIAD'	'Cingulum_Ant_R'	'NCG'	'Caudate_L'	'VER3'	'Vermis_3'
'CINMG'	'Cingulum_Mid_L'	'NCD'	'Caudate_R'	'VER4_5'	'Vermis_4_5'
'CINMD'	'Cingulum_Mid_R'	'NLG'	'Putamen_L'	'VER6'	'Vermis_6'
'CIPG'	'Cingulum_Post_L'	'NLD'	'Putamen_R'	'VER7'	'Vermis_7'
'CIPD'	'Cingulum_Post_R'	'PALLG'	'Pallidum_L'	'VER8'	'Vermis_8'
'HIPPOG'	'Hippocampus_L'	'PALLD'	'Pallidum_R'	'VER9'	'Vermis_9'
'HIPPOD'	'Hippocampus_R'	'THAG'	'Thalamus_L'	'VER10'	'Vermis_10'
'PARA_HIPPOG'	'ParaHippocampal_L'	'THAD'	'Thalamus_R'		

**Fig. 10** The functional areas according to Montreal Neurological Institute. In three columns the abbreviations and descriptions of 42 functional areas are listed

problems. Hence, voxel sieving is generic in that it can be used for an ensemble of data analysis approaches and tasks.

More importantly, voxel sieving has the capability to uncover patterns in brain activity data that are hard to capture with existing fMRI data analysis techniques. Our method generally identifies multiple specific cortical clusters across the brain. We attribute the specificity of the results to the ability of our method to find voxel activity patterns with both high coherence and high variance, while other similar methods [17, 22, 23] focus on coherence only. Another distinguishing feature of voxel sieving is that identified voxel clusters are independent of each other. Rather than seeking for voxel clusters with similar temporal properties, the method inclines to search for distinct cluster characteristics. As a consequence, once a specific

synchrony is captured in one cluster, the same structure will no longer be captured in subsequent clusters. Overlapping voxel clusters, however, are allowed if such voxels induce clusters that uncover distinct brain processes. These aspects of our method are important and cannot be captured by fitting predefined models to voxels or by globally grouping voxels into classes, clusters or components.

In the study, we have experienced, as others have done before, that estimating the number of clusters and finding the optimal cluster size is a difficult task as there is no clear definition of a ‘cluster’. Simulation studies have demonstrated that the Gap estimate is good for identifying well-separated clusters ([18]). However, when data are not clearly separated into groups, suboptimal clusters can be identified. In this case, a more flexible procedure is needed

for the determination of the best cluster. One alternative is to select a cluster with a larger size than the optimal cluster and a Gap statistic within a small percentage of the maximal Gap statistic. We have not investigated whether our brain activity data suffer from suboptimal voxel clusters and how alternative procedures effect the performance of voxel sieving.

## 6 Conclusion

Our statistical signal analysis method identifies hemodynamic synchrony that distinguishes well between experimental conditions or cognitive states. Two important properties of these method are that it allows to conveniently specify (1) external sources of variation associated with brain activity and (2) the degree of supervision during the data analysis process. In the absence of prior or external information about brain scans, the method operates in a data-driven manner. When meta-information about brain activity is present, the method uses this for fully or partially supervised data analysis. This flexibility of our method together with its ability to identify multiple, potentially overlapping, brain areas independently of each other and in a multivariate way, makes it appropriate for finding very specific brain responses, even to complex stimuli. We have shown this in the context of a free movie viewing fMRI study, where flexible probing of functional characteristics exposed spatially localized synchronous brain activity at anticipated and less expected brain regions. The significance of these findings is supported by the excellent performance of our method on an international test benchmark for fMRI-based movie stimuli prediction. Hence, we conclude that the unique ability of our method to capture distributed and overlapping hemodynamic responses in a flexible and effective way, suitably complements existing statistical signal processing methods in neuroscience.

**Acknowledgments** We acknowledge Prof. dr. Victor Lamme, Dr. Steven Scholte and Dr. Hilde Huizenga (Department of Psychology, University of Amsterdam) for the valuable discussions. Support for this work has been generously provided by the national projects MultimediaN and VL-E, both funded by the Dutch BSIK Program.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Aguirre GK, Zarahn E, D'Esposito M (1998) The variability of human, bold hemodynamic responses. *NeuroImage* 8(4):360–369
2. Arfanakis K, Cordes D, Haughton VM, Moritz CH, Quigley MA, Meyerand ME (2000) Combining independent component analysis and correlation analysis to probe interregional connectivity in fMRI task activation datasets. *Magn Reson Imaging* 18(8):921–930
3. Beckmann C, Jenkinson M, Smith SM (2003) General multilevel linear modeling for group analysis in fmri. *NeuroImage* 20:1052–1063
4. Beckmann C, Smith SM (2004) Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* 23:137–152
5. Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human v1. *J Neurosci* 16(13):4207–4221
6. Calhoun V, Adali T (2006) Unmixing fmri with independent component analysis. *IEEE Eng Med Biol Mag* 25(2):79–90
7. Chuang K-H, Chiu M-J, Lin C-C, Huang K-M, Chiang P-J, Chen J-H (1999) Model free functional mri analysis using kohonen clustering neural network. *IEEE Trans Med Imaging* 18(12):1117–1128
8. Correa N, Li Y-O, Adali T, Calhoun V (2005) Comparison of blind source separation algorithms for fmri using a new matlab toolbox: Gift. In: *Proceedings of IEEE international conference on acoustics, speech, signal processing*, vol. 5
9. Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fmri) “brain reading”: detecting and classifying distributed patterns of fmri activity in human visual cortex. *NeuroImage* 19(2 Pt 1):261–270
10. Editorial (2006) What's on your mind. *Nat Neurosci* 7(8):523:534
11. Esposito F, Scarabino T, Hyvarinen A, Himberg J, Formisano E, Comani S, Tedeschi G, Goebel R, Seifritz E, Di Salle F (2005) Independent component analysis of fmri group studies by self-organizing clustering. *Neuroimage* 25(1):193–205
12. Friston KJ, Holmes A, Worsley K, Poline JB, Frith C, Frackowiak R (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 1:189–210
13. Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith C, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210
14. Friston KJ, Josephs O, Rees G, Turner R (1998) Nonlinear event-related responses in fmri. *Magn Reson Med* 39(1):41–52
15. Grill-Spector K, Knouf N, Kanwisher N (2004) The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* 7(5):555:562
16. Handwerker DA, Ollinger JM, D'Esposito M (2004) Variation of bold hemodynamic response function across subjects and brain regions and their effects on statistical analysis. *NeuroImage* 21(4):1639–1651
17. Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Inter-subject synchronization of cortical activity during natural vision. *Science* 303(5664):1634–1640
18. Hastie T, Tibshirani R, Eisen MB, Alizadeh A, Levy R, Staudt L, Chan W, Botstein B, Brown P (2000) Gene shaving as a method for identifying distinct sets of genes with similar expression patterns. *Genome Biol* 1(2) (Epub)
19. Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430
20. Haynes J, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8:686–691
21. Haynes J, Rees G (2006) Decoding mental states from brain activity in humans. *Nature Neuroscience* 7(7):523:534
22. Hejnar MP, Kiehl KA, Calhoun VD (2006) Interparticipant correlations: a model free fmri analysis technique. *Hum Brain Mapp*

23. Kim D, Pearlson GD, Kiehl KA, Bedrick E, Demirci O, Calhoun VD (2008) A method for multi-group inter-participant correlation: abnormal synchrony in patients with schizophrenia during auditory target detection. *NeuroImage* 39(2):1129–1141
24. Liu Y, Gao JH, Liu HL, Fox PT (2000) The temporal response of the brain after eating revealed by functional mri. *Nature* 405(6790):1058–1062
25. Marzouk YM, Ghoniem AF (2005) K-means clustering for optimal partitioning and dynamic load balancing of parallel hierarchical n-body simulations. *J Comput Phys* 207(2):493–528
26. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J (1995) A probabilistic atlas of the human brain: theory and rationale for its development. *NeuroImage* 2:89–101
27. McKeown MJ, Makeig S, Brown CG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998) Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp* 6(3):160–188
28. McKeown MJ, Li J, Huang X, Lewis MM, Rhee S, Young Truong KN, Wang ZJ (2007) Local linear discriminant analysis (lllda) for group and region of interest (roi)-based fmri analysis. *Neuroimage* 37(3):855–65
29. Mouro-Miranda J, Bokde ALW, Born C, Hampel H, Stetter M (2005) Classifying brain states and determining the discriminating activation patterns: support vector machine on functional mri data. *Neuroimage*
30. Ngan S, Hu X (1999) Analysis of functional magnetic resonance imaging data using self-organizing mapping with spatial connectivity. *Magn Reson Med* 41(5):939–946
31. Ramsay J, Silverman B (1997) *Functional data analysis*. Springer, Berlin
32. Schneider W, Bartels A, Formisano E, Haxby J, Goebel R, Mitchell T, Nichols T, Siegle G (2006) Competition: inferring experience based cognition from fmri. In: *Proceedings organization of human brain mapping*, Florence, Italy
33. Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters via the gap statistic. *J R Stat Soc B* 63(2)
34. van Rijsbergen CJ (1979) *Information retrieval*. Butterworths, Toronto
35. Viviani R, Grohn G, Spitzer M (2005) Functional principal component analysis of fmri data. *Hum Brain Mapp* 24:109–129
36. Wang Z, Childress AR, Wang J, Detre JA (2007) Support vector machine learning-based fmri data group analysis. *Neuroimage*
37. Worsley KJ, Poline JB, Friston KJ, Evans AC (1997) Characterizing the response of pet and fmri data using multivariate linear models. *Neuroimage* 6(4):305–319