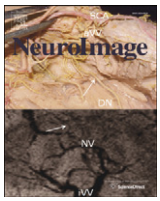
NeuroImage 56 (2011) 508–516

Contents lists available at ScienceDirect

NeuroImage

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



A Monte Carlo method for locally multivariate brain mapping

Malin Björnsdotter *, Karin Rylander, Johan Wessberg

Department of Physiology, Institute of Neuroscience and Physiology, University of Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 30 November 2009
Revised 27 May 2010
Accepted 20 July 2010
Available online 30 July 2010

ABSTRACT

Locally multivariate approaches to functional brain mapping offer a highly appealing complement to conventional statistics, but require restrictive region-of-interest hypotheses, or, in exhaustive search forms (such as the "searchlight" algorithm; Kriegeskorte et al., 2006), are excessively computer intensive. We therefore propose a non-restrictive, comparatively fast yet highly sensitive method based on Monte Carlo approximation principles where locally multivariate maps are computed by averaging across voxelwise condition-discriminative information obtained from repeated stochastic sampling of fixed-size search volumes. On simulated data containing discriminative regions of varying size and contrast-to-noise ratio (CNR), the Monte Carlo method reduced the required computer resources by as much as 75% compared to the searchlight with no reduction in mapping performance. Notably, the Monte Carlo mapping approach not only outperformed the general linear method (GLM), but also produced higher discriminative voxel detection scores than the searchlight irrespective of classifier (linear or nonlinear support vector machine), discriminative region size or CNR. The improved performance was explained by the information-average procedure, and the Monte Carlo approach yielded mapping sensitivities of a few percent lower than an information-average exhaustive search. Finally, we demonstrate the utility of the algorithm on whole-brain, multi-subject functional magnetic resonance imaging (fMRI) data from a tactile study, revealing that the central representation of gentle touch is spatially distributed in somatosensory, insular and visual regions.

Introduction

In addition to conventional general linear model (GLM; Friston et al., 1994) approaches, multivoxel pattern analysis (MVPA) has emerged as a highly appealing alternative for localization of regions involved in specific cognitive processes (as measured by functional magnetic resonance imaging; fMRI). In contrast to univariate GLM activation mapping, MVPA methods utilize the combined effect of numerous voxels in a multivariate fashion, allowing identification of spatially distributed condition-related signal patterns (see Norman et al., 2006 and Haynes and Geraint, 2006 for reviews). Such multivariate approaches allow highly sensitive differentiation of subtle brain activity changes encoded across multiple voxels which conventional univariate statistics fail to capture (Kriegeskorte et al., 2006; Björnsdotter Åberg and Wessberg, 2008b; De Martino et al., 2008; Formisano et al., 2008).

A number of MVPA techniques have been developed specifically for fMRI-based brain mapping. These include massively multivariate methods, where a classifier is applied to all voxels in an entire brain volume simultaneously and resulting classifier weights, indicative of voxel-wise contributions, are extracted (Beauchamp et al., 2009; LaConte et al., 2005; Mourão-Miranda et al., 2005; Sato et al., 2009). In addition to being limited to linear classifiers, providing a direct relationship between individual voxels and classifiers weights, these methods do not alleviate the curse of dimensionality (Bellman, 1961) and hence yield suboptimal brain state discrimination sensitivities.

In contrast, voxel selection optimization methods such as globally multivariate recursive feature elimination (RFE; De Martino et al., 2008; Formisano et al., 2008; Staeren et al., 2009) and evolutionary algorithms (Björnsdotter et al., 2009; Björnsdotter Åberg and Wessberg, 2008a), explicitly identify maximally discriminative regions of arbitrary size, shape and location. Similar to massively multivariate methods, however, RFE requires linear classifier weight rankings for iterative voxel elimination, whereas the highly flexible evolutionary approach suffers from implementation and application complexity issues (such as large number of parameters requiring specification).

Locally multivariate search-based approaches targeting regions of fixed size and shape (including region-of-interest analysis; Cox and Savoy, 2003; Haynes and Rees, 2005; Kamitani and Tong, 2005; and exhaustive searches scanning whole brain volumes; Kriegeskorte et al., 2006), in contrast, provide both appealing flexibility and simplicity of implementation. In particular, the attractively simple and

* Corresponding author. Department of Physiology, Box 432, Institute of Neuroscience and Physiology, University of Gothenburg, 405 30 Gothenburg, Sweden. Fax: +46 31 786 3512.
E-mail addresses: malin.bjornsdotter@neuro.gu.se (M. Björnsdotter), karin.rylander@neuro.gu.se (K. Rylander), wessberg@physiol.gu.se (J. Wessberg).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.07.044

intuitive "searchlight" algorithm introduced by Kriegeskorte et al. (2006) has shown great utility in numerous studies (see e.g. Kriegeskorte et al., 2007; Kriegeskorte and Bandettini, 2007; Haynes et al., 2007; Bode and Haynes, 2009; Clithero et al., 2008; Stokes et al., 2009). Whole-volume maps are produced by computing multivariate information measures across fixed-size, fixed-shape (typically spherical) search volumes sequentially centered on each voxel in the brain volume. As opposed to region-of-interest methods, the searchlight requires no a priori hypotheses on the location of activity patterns and can be used in conjunction with any arbitrary information-measure approach (including nonlinear classifiers).

A substantial drawback of the searchlight algorithm is the excessive number of times local information must be computed (once per voxel, that is, in the order of tens to hundreds of thousands or more). Initially, the simple and fast Mahalanobis distance was therefore proposed as a measure of brain activity pattern content (Kriegeskorte et al., 2006), but more recently the support vector machine (SVM; Suykens et al., 2002) has been suggested (De Martino et al., 2005; Haynes et al., 2007). The latter is a highly effective classifier, and, as such, has dominated the multivoxel literature (see e.g. Cox and Savoy, 2003; Mitchell et al., 2004; Mourão-Miranda et al., 2005; LaConte et al., 2005; Kamitani and Tong, 2005; Davatzikos et al., 2005; Mourão-Miranda et al., 2006; De Martino et al., 2008; Björnsdotter Åberg and Wessberg, 2008b; Formisano et al., 2008, 1975; Staeren et al., 2009; Beauchamp et al., 2009; Björnsdotter et al., 2009). The time requirements associated with training and testing of the classifier, however, restrict the usefulness of SVMs in conjunction with the searchlight—particularly with nonlinear kernels, high-resolution fMRI, whole-brain analysis and multiple subjects.

We, therefore, propose a Monte Carlo approximation (Hammersley and Handscomb, 1975) of the exhaustive search to reduce computational demands. Monte Carlo methods approach problems by observing properties of randomly generated instances and have been successfully utilized in various fields where the system under investigation is highly complex and exact computations are not feasible (including e.g. medicine, finance and engineering; Zaidi and Sgouros, 2002; McLeish, 2005; Amar, 2006). MVPA analysis based on such random selection principles where multivariate information is evaluated on randomly selected voxel subsets has been successful for brain mapping (Björnsdotter Åberg and Wessberg, 2008b) as well as classification (Kuncheva et al., 2010). Unrestricted global random sampling does, however, effectively ignore information encoded in local neighborhoods. Here, instead, we sample spatially restricted voxel subsets of fixed size and shape akin to the searchlight algorithm such that spatial correlations in the data are explicitly taken into account. Instead of exhaustively centering the search volume on each single voxel, information is averaged across stochastically selected locations in a Monte Carlo fashion, such that the resulting multivariate measure assigned to any given voxel is a robust estimate of the mean information contributed by that voxel in conjunction with its neighbors.

Comparing with the searchlight, we illustrate the dramatically reduced computer requirements and beneficial effect on brain mapping performance on realistically simulated data. In addition, we demonstrate the practical utility of the algorithm in localizing spatially distributed cortical representations of experimental conditions on whole-brain, multi-subject fMRI data from a gentle touch study using a standard desktop computer.

Materials and methods

Monte Carlo mapping

The Monte Carlo computational method is based on repeated stochastic sampling and has been extensively implemented in various scenarios where it is infeasible or not possible to compute exact results deterministically (Hammersley and Handscomb, 1975). In general terms, the Monte Carlo method can be described by the following procedure:

1. Identify a domain of possible inputs.
2. Generate inputs randomly from the domain.
3. Perform a computation using the inputs.
4. Combine the results of the individual computations into a final result.

The algorithm is thus iterative in nature, and an increased number of iterations results in an improved approximation of the true (exhaustive) result.

Similar to the searchlight algorithm (Kriegeskorte et al., 2006), Monte Carlo maps are obtained by evaluation of fixed-size, fixed-shape (typically spherical) brain regions in terms of information content. As demonstrated by pseudo-code in Fig. 1 and the illustration in Fig. 2A, each iteration of the algorithm stochastically partitions the fMRI volume into a number of search spheres such that each voxel is included in one (and one only) sphere and the information measure is computed for and assigned to each such sphere. A robust iteration-average voxel-wise information contribution estimate is subsequently computed, and, thus, the mean information across all constellations in which a voxel took part is assigned to that voxel.

The searchlight algorithm, in contrast, assigns each voxel the value computed when the search sphere was centered on that voxel (Fig. 2B). Thus, where the searchlight requires as many information computations as there are voxels in the brain volume in an exhaustive fashion, the Monte Carlo method reduces the required number to the total number of voxels in the volume divided by the number of voxels in the search sphere multiplied by the number of iterations. Notably, larger searchlight spheres yield increased computational requirements (since the number of computations are unchanged but each computation is performed on a larger number of voxels), whereas the contrary holds true for the Monte Carlo approach: with increased search sphere volume, fewer search spheres are required to cover the brain volume and thus the number of information computations is reduced. As per the nature of Monte Carlo schemes, iterating the algorithm to perform as many information computations as the searchlight yields an exhaustive result with the important difference that the final map reflects an average contribution across all search spheres in which each voxel was included. Note that the proposed method should not be confused with random subsampling techniques regularly used in machine learning for estimation of classifier error rates (see e.g. Efron, 1994), where classifier performances are evaluated on stochastically selected data sample—as opposed to voxel—subsets.

Only two parameters affect the Monte Carlo mapping performance and require specification: the search sphere volume (given as percentages of the total brain volume throughout this paper) and the number of iterations (see the Results section for further details).

```
BEGIN
  For (Number of iterations);
      While (Not all voxels are selected);
          Select a random hitherto unselected voxel;
          Select all other hitherto unselected voxels within the specified sphere radius;
          Compute an information measure on the voxels contained in the sphere;
      End While
  End For
  Compute the mean information per voxel across iterations;
END
```

Fig. 1. Pseudo-code describing the proposed Monte Carlo fMRI brain mapping algorithm.

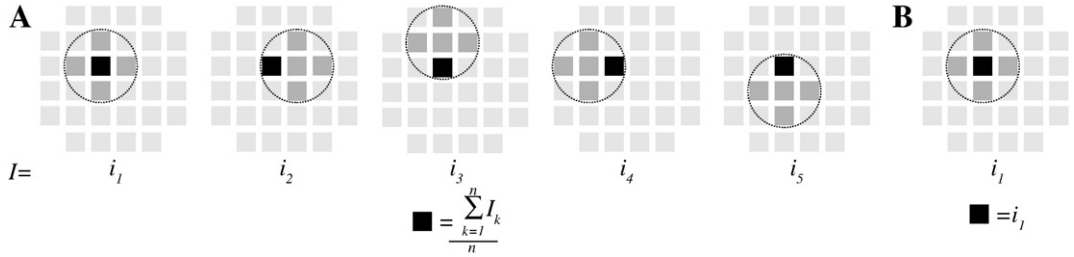


Fig. 2. A. Visualization of the proposed Monte Carlo fMRI brain mapping approach for computing the multivariate information contribution by one voxel. For each iteration *k*, the search volume (dot-line circle) is centered on a random voxel (black) and the voxels within the volume (black and dark gray) are used to compute a multivariate information measure i_k (e.g. condition classification accuracy) for that voxel. In subsequent iterations, the voxel is included in different constellations with other neighboring voxels, and a final information measure is computed as the average across all *n* iterations. Here, $n = 5$ iterations corresponds to an exhaustive search. The searchlight algorithm (B; Kriegeskorte et al., 2006), in contrast, sequentially centers the sphere on each single voxel and the corresponding (non-averaged) measure is assigned to that voxel.

Multivariate information measure

We evaluated mapping performances using SVM classifiers, with linear as well as nonlinear (radial basis function; RBF) kernels, to quantify locally multivariate information content. In each search volume, a hold-out validation with 50% training data was performed and averaged over two repetitions. The data were randomly partitioned into training and testing sets each time, and the linear and nonlinear SVMs were applied on identical partitions. The excessive computer resources required by the searchlight prevents SVM classifier parameter optimization within reasonable time, and to ensure identical performance evaluation procedures between the searchlight and Monte Carlo approach the SVM parameters were specified empirically and fixed throughout the analysis (regularization parameter = 0.1, bandwidth for RBF kernel = 1000). Similarly, the searchlight could not feasibly be used on the whole-brain, multi-subject real fMRI data and was therefore only applied to the simulated data.

We implemented the algorithms in Matlab (The Mathworks, Massachusetts, USA) and C on a standard desktop computer (3 GHz, 2 GB RAM), and the Matlab SVM package LS-SVMlab developed by the group SCD/sista in the department ESAT at the KULeuven, Belgium, was used (Suykens et al., 2002).

fMRI data

Simulated data

The performance of the algorithm was evaluated on realistic data modeled similar to that presented by De Martino et al. (2008). We simulated a block-design with two conditions for a repetition time (TR) set to 3.5 s, and a functional voxel resolution to $3 \times 3 \times 3$ mm. Eighty stimulations (equal number of each condition), lasting 10.5 s, were modeled. Blood-oxygen-level dependent (BOLD) responses were simulated by convolving the stimuli with a conventional hemodynamic response function (modeled using a double-gamma function). On a voxel-by-voxel basis, activations were added to temporally autocorrelated noise. The noise was modeled as follows:

$$R_a(t) = \rho_k R_0(t-1) + \sqrt{1-\rho_k^2} R_0(t) \quad (1)$$

where R_0 is random Gaussian noise and ρ_k: $N(0.5,0.1)$ allowed adjustment of the amount of autocorrelation at voxel *k*.

Two different datasets were synthesized. The first mimicked authentic fMRI data, and a cortex mask extracted from a human subject containing 28 502 voxels was used for realistic spatial arrangement of the simulated activity. Six active brain regions of different size and discriminative information content were simulated. The voxels in these regions were assigned to one of two populations ($\text{condition}_1 > \text{condition}_2$; $\text{condition}_2 > \text{condition}_1$) with random spatial distribution within the discriminative region. Three regions were large (142 voxels, same shape with different location) and three small (38 voxels, again same shape and different locations). For each region size, the contrast-to-noise ratio (CNR; the difference between responses compared to the standard deviation of the noise) was set to either 0.2, 0.5 or 0.8. All reported results refer to this dataset, unless otherwise specified. The second dataset was synthesized in order to investigate whether the information-averaging procedure compromises spatial resolution (i.e. the ability to resolve neighboring elements). A $20 \times 40 \times 10$ voxel cube was generated and five discriminative regions (CNR = 0.8) arranged in straight lines with spacing varying from one to four voxels were simulated in this cube. For both datasets, the variability of the BOLD response to the condition trials (percent of variability compared to the maximum response) was set to 10%. The signal-to-noise ratio (SNR; the amplitude of the response compared to the standard deviation of the noise) was fixed to 0.8.

Real fMRI data

In addition to the simulated data, the proposed Monte Carlo mapping approach was evaluated on authentic fMRI data from an experiment studying brain responses to gentle touch. The Ethical Review Board at the University of Gothenburg approved the study, and the experiments were performed in accordance with the Declaration of Helsinki. Informed consent was obtained from nine healthy subjects (aged 18–40 years, all right-handed and female).

Gentle stimulation, known to vigorously activate mechanoreceptive thin, unmyelinated C-tactile (CT) fibers as well as thick, myelinated Aβ afferents (Edin, 2001; Löken et al., 2009; Vallbo et al., 1993, 1999), was delivered using a seven centimeter wide soft artist's goat hair brush. The experimenter manually stroked the brush on the left thigh in a proximal to distal direction using an indentation force of 0.8 N. The stimuli were applied according to cues from the scanner, and all subjects were instructed to focus on the stimulus throughout the fMRI scanning session. The distance covered was 16 centimeters for a duration of 3 s. Three-volume blocks of brushing and rest were alternated. The scanning session consisted of one anatomical and four functional scans. During each functional scan, 20 blocks were obtained, totaling 80 three-volume blocks.

A 1.5 T fMRI scanner (Philips Intera) with a SENSE head coil (acceleration factor 1) was used to collect whole-brain anatomical scans using a high-resolution T1-weighted anatomical protocol (TR 10 ms, TE 4 ms, flip angle 15°, FOV 300 mm). Functional scans were acquired using a BOLD protocol and a T2*-weighted gradient-echo, echo-planar imaging (EPI) sequence (flip angle 90°, TR 3.0 s, TE 50 ms, thickness 6 mm, in-plane resolution 3.6×3.6 mm, FOV 230 mm, matrix 64×64 voxels). The scanning planes were oriented parallel to the line between the anterior and posterior commissure and

covered the brain from the top of the cortex to the base of the cerebellum. The functional data was reconstructed from an in-plane grid resolution of 64×64 to 128×128 voxels while still in frequency space.

Using the software package BrainVoyager QX (Brain Innovation, Maastricht, The Netherlands), the data was pre-processed as follows: slice-scan-time correction by cubic spline interpolation resampling, motion correction by trilinear interpolation resampling according to parameters obtained from minimizing the sum of squares of the voxel-wise intensity differences between each volume and the first volume of the run, high-pass filtering to remove temporal drifts, co-registration to anatomical images, projection into Talairach space using trilinear interpolation (with a final resolution of 3×3×3 mm), and finally, masking of non-cortex voxels (retaining 34,519 voxels in each subject). In addition to the multivoxel analysis, the real data was also subject to standard general linear modeling using BrainVoyager QX in order to obtain a univariate activation map.

Single-trial response estimation

We estimated the single-trial BOLD activity pattern for both simulated and real data by fitting a voxel-wise GLM with one predictor representing each single-trial response (obtained by convolution of a boxcar with a double-gamma hemodynamic response function; Friston et al., 1994) and one linear predictor accounting for a within-trial linear trend. The GLM was estimated in single windows of length two TRs before stimulus onset and three TRs after onset. The regressor β coefficients were subsequently obtained as an estimate of each trial response.

The final simulated datasets contained 80 volumes (half of each condition) and 28,502 or 8000 voxels, respectively, whereas the real fMRI dataset comprised a total of 160 volumes (half tactile stimulation, half rest) and 34,519 voxels.

Performance metrics

For the simulated data, the receiver operating characteristic curve (ROC; a plot of the voxel detection sensitivity versus 1-specificity for varying map thresholds) was computed for each map. As a measure of the discriminative voxel detection performance, the area under the curve (AUC), where a value of 1 corresponds to perfect discriminative voxel detection (no true negatives or false positives), was subsequently obtained. First, the AUC was computed on the dataset as a whole. Second, in order to isolate the impact of CNR and discriminative region size, the AUC was estimated for each of the individual discriminative regions by masking the voxels in the other discriminative regions. With locally multivariate methods this is feasible since mapping of any given region is not affected by the presence of discriminative voxels in other (sufficiently distant, as is the case here) regions. The obtained values were tested for statistical significance using specific AUC tests (Metz, 1984; Zhou et al., 2002), and designated software (ROCKIT, C.E. Metz, Univ. Chicago, IL, USA).

The computer resource requirements were quantified in terms of the number of required information computations (training and testing of the SVM classifier), since this measure (as opposed to e.g. time requirements) is independent of computer power. We used n Monte Carlo iterations such that the total number of information computations was less than that of the searchlight and the maximum allowed was $n=20$. For improved reliability, each algorithm was applied five times and the average resulting performance is presented.

Due to excessive time requirements for large search sphere radii, the Monte Carlo and searchlight algorithm performances were compared on search sphere volumes of 0.1%, 0.5% and 1% of the total brain volume, whereas the effect of larger radii was evaluated using only the Monte Carlo approach.

For the real fMRI data, group analysis maps reflecting regional information content (measured in proportion of correctly classified testing volumes) were formed by averaging across results obtained on the individual level. All clusters of minimum volume 50 mm^3 with a subject-mean score of at least 0.7 were subsequently identified. Although formal single-subject (e.g. using permutation testing; Good, 2004) or group-level statistical testing (e.g. using fixed or random effect models; Mourão-Miranda et al., 2005; LaConte et al., 2005; Mourão-Miranda et al., 2006; Wang et al., 2007) is outside the scope of this study, 0.7 is sufficiently above chance level (0.5) to be considered statistically sound. The Talairach coordinates of the peak voxel in each cluster are reported, as is the Brodmann area in which a majority of the voxels were located.

For data visualization, the program MRIcron (by Chris Rorden, www.sph.sc.edu/comd/rorden/mricron/) was used.

Results

Simulated data

The multivoxel approaches (searchlight and Monte Carlo) both outperformed the univariate GLM, which achieved a discriminative voxel detection performance of 0.503 (as measured in area under the ROC-curve, AUC; see Table 1).

Surprisingly, the Monte Carlo method was significantly more sensitive for all search sphere volumes and classifier kernels at a mean AUC of 0.873 (range 0.840–0.899) than the searchlight (0.826; range 0.767–0.857; paired t-test, $p<0.05$; Table 1; test for AUC differences, $p<0.001$; Metz, 1984). Similar results were obtained for all discriminative regions irrespective of size and CNR (Table 2).

Information-averaging improves voxel detection performance

In order to explore the mechanisms underlying the higher mapping performance achieved by the Monte Carlo approach, we applied the algorithm exhaustively to mimic the searchlight while still quantifying voxel-wise information content as the average across iterations (see the illustration in Fig. 2 for details). As shown in Table 1, this information-average searchlight consistently achieved higher mapping sensitivities than both the Monte Carlo approach and the standard searchlight. Hence, assigning each voxel the average information content across all constellations in which it has been included improves the mapping sensitivity and explains the higher performance of the Monte Carlo method. The information-averaging procedure improved the average performance over the searchlight for all search sphere volumes and kernel types significantly to an AUC of 0.904 (range 0.877–0.918, test for AUC differences, $p<0.001$; Metz, 1984), corresponding to an increase in discriminative voxel detection performance of 9.52% (range 7.07–14.38%).

Computational load reduction

For the kernel (RBF) and search sphere volume (0.5% of the total volume) with which all methods achieved their highest mapping sensitivities, the number of required information computations (number of times the classifier was trained and tested) was dramatically reduced from 28 502 for the searchlight to 9 705 for the Monte Carlo method, corresponding to a reduction of 66% with no loss in performance (see Table 1).

The more iterations allowed for the Monte Carlo method, the better the mapping performance, and, consequently, the computational load. As exemplified in Fig. 3A (for a search sphere volume of 0.5%), the Monte Carlo method mapping performance improves dramatically during the initial information computations (up to the order of 5000) and approximates the information-average searchlight performance at relatively few computations. For the RBF kernel and search sphere volume of 0.5% where the information-average searchlight obtained an AUC of 0.918, for example, the standard Monte Carlo method reduced the number of required information

Table 1
Table of brain mapping sensitivities for the searchlight (Kriegeskorte et al., 2006), Monte Carlo method and information-average searchlight on the simulated data, measured in area under the receiver operating curve (AUC). The best performance for each approach is denoted by *. The number of information computations for the searchlights was 28,502. The search sphere volume is expressed in percentage of total brain volume. Lin-SVM: support vector machine with a linear kernel, RBF-SVM: support vector machine with a radial basis function kernel.

	Classifier					
	Lin-SVM			RBF-SVM		
Search volume (%)	0.1	0.5	1.0	0.1	0.5	1.0
Searchlight	0.767	0.835	0.842	0.809	0.857*	0.848
MC	0.840	0.883	0.880	0.850	0.899*	0.886
MC*	0.877	0.915	0.911	0.896	0.918*	0.909
MC nr. computations	27 778	11 643	6 745	27 778	9 705	6 745
Reduction (%)	2.54	59.15	76.33	2.54	65.95	76.33

computations by 75% for a 2.80% reduction in mapping performance (Fig. 3B).

Information-averaging effect on spatial resolution

Fig. 4 shows the maps obtained with a search sphere volume of 0.5% and the RBF kernel. Despite a 66% reduction in computer resources the Monte Carlo map is strikingly similar to both the searchlight and information-average searchlight maps. Moreover, as is clear in the large discriminative cluster with $CNR = 0.8$ (shown in blue) and in stark contrast to the searchlight map, voxels with similar discriminative information content obtain homogenous values as a result of the information-averaging smoothing effect. Map values obtained with the searchlight algorithm, on the other hand, are deceptively dependent on the regional context such that any voxel value is highly sensitive to the number of discriminative voxels within the search sphere; this results in substantially higher values for voxels in the center of the discriminative cluster than for the surrounding voxels (despite the same information content). The information-averaging does not, on the other hand, appear to affect the limit at which the methods can resolve closely spaced elements. When evaluated on the second simulated dataset with a search sphere radius of two voxels, both the original and information-average searchlight failed in resolving lines spaced by two voxels or less (Fig. 5). Similarly, both methods were equally successful in resolving lines separated by three or more voxels. Whereas the search sphere size is clearly a critical factor, information-averaging appears to neither improve nor decrease the ability to resolve fine spatial details.

Effect of CNR and discriminative cluster size

The mapping performance of all methods increased linearly with CNR, as can be expected (Table 2). Importantly, the size of the discriminative cluster also has a significant impact on the mapping performance—the larger the discriminative cluster the higher the achieved performance (across search sphere volumes and classifier kernels for all methods; Wilcoxon's signed rank test, $p<0.001$). For example, the Monte Carlo method achieved an AUC of 0.938 on the large cluster, whereas a reduced AUC of 0.769 was obtained for the small cluster with the same CNR.

Effect of search sphere volume

All methods achieved highest sensitivities with small to medium sized search spheres, and the best results were obtained at a volume of 0.5% of the total brain volume (see Table 1 and Fig. 6A; but note that search volumes larger than 1% were not used in conjunction with the searchlight and information-average searchlight due to excessive time requirements). There was no substantial difference in performance between the small and medium search spheres (0.5 to 2%), despite a dramatic decrease in number of required information computations up to a volume of 2% (Fig. 6B). All methods, however, obtained better results on the large clusters using larger search spheres (volume of 1.0% vs. 0.1% and 0.5% vs. 0.1%, $p<0.01$, test for AUC differences; Metz, 1984; interaction between cluster size and search sphere volume, ANOVA, $F_{2,15} = 4.89$, $p<0.01$).

Linear vs. nonlinear kernel

Across all discriminative regions and search sphere volumes, the mapping sensitivities obtained with all methods in conjunction with the RBF kernel were significantly higher than those of the linear kernel maps ($p<0.01$, AUC test; Metz, 1984; Table 2). The difference was not substantial; the RBF kernel improved the Monte Carlo voxel detection performance from 0.883 to 0.899 on the combined cluster analysis with a search sphere volume of 0.5%, and the corresponding figures for the information-average searchlight were 0.915 and 0.918 (Table 1).

Real data

Given the results obtained on the simulated data, the Monte Carlo brain mapping algorithm was iterated 15 times with a search sphere volume of 0.9% of the total cortex volume (corresponding to between 5997 and 6007 information computations per subject, compared to a required 34,519 for any exhaustive search) on each individual tactile

Table 2
Table of separate simulated data mapping performances for each discriminative region for the searchlight, Monte Carlo method and information-average searchlight, measured in area under the receiver operating curve (AUC). The corresponding search sphere volume, expressed in percentage of total brain volume, is noted in parenthesis. For all but the small cluster with CNR = 0.2 for the Monte Carlo method, the large cluster with CNR = 0.2 for the searchlight, and the large cluster with CNR = 0.5 for the information-average searchlight, the RBF kernel obtained the best results. The number of information computations for the searchlights was 28,502.

CNR	Discriminative region size					
	Large			Small		
	0.2	0.5	0.8	0.2	0.5	0.8
Searchlight	0.807 (1.0)	0.869 (0.5)	0.988 (1.0)	0.587 (0.1)	0.569 (0.5)	0.875 (0.5)
MC	0.882 (1.0)	0.938 (0.5)	0.997 (1.0)	0.652 (0.5)	0.769 (0.5)	0.966 (1.0)
MC*	0.897 (0.5)	0.956 (0.5)	0.998 (0.5)	0.647 (0.1)	0.784 (0.5)	0.980 (0.5)
MC nr. computations	7 102	12 943	7 102	651	10 996	6388
Reduction (%)	75.08	54.59	75.08	97.72	61.42	77.59

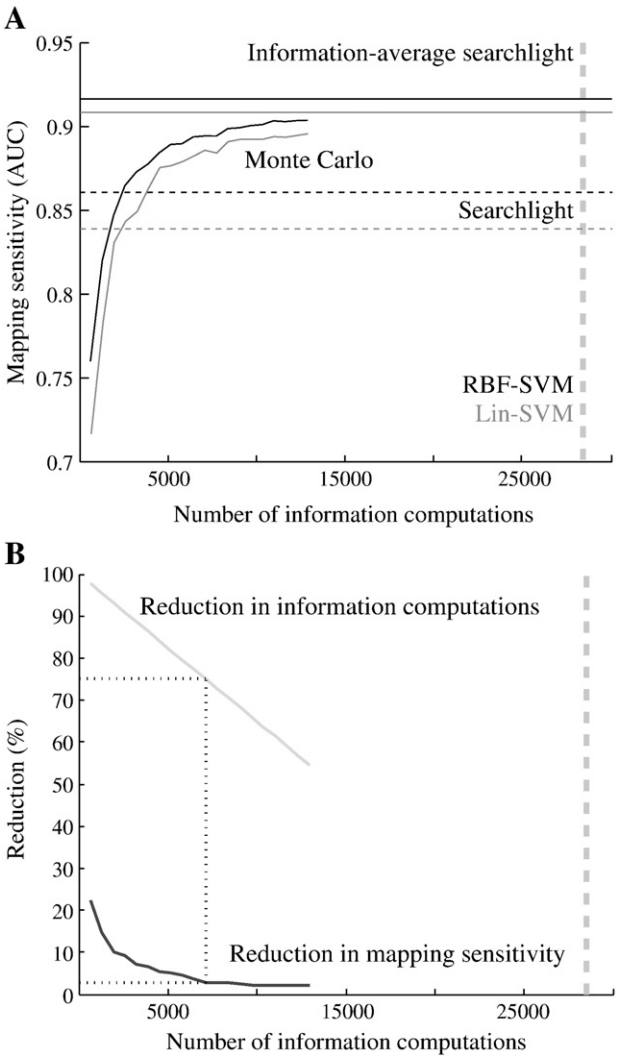


Fig. 3. A. Monte Carlo algorithm voxel detection performance (measured in area under the receiver operating characteristic curve; AUC) as a function of the number of information computations (number of times the classifier was trained and tested), for the RBF (black) and linear (gray) SVM kernel. The searchlight and information-average searchlight performances are plotted as dotted and straight lines, respectively. B. The reduction in number of information computations (gray) and voxel detection performance (black) of the Monte Carlo method compared to the information-average searchlight, for the RBF kernel. A reduction in number of information computations by 75% corresponds to a decrease in voxel detection performance of 2.80%, for example, as indicated by the dotted lines. A search sphere volume of 0.5% of the total brain volume was used in this example, and the thick gray vertical lines represent the number of information computations required for the searchlights.

fMRI dataset (requiring 13.9–14.5 min to complete per individual). The resulting group map is shown in Fig. 7, and the peak voxel Talairach coordinates are presented in Table 3.

The contralateral (right) primary and bilateral secondary somatosensory cortices (located in the post-central gyrus and parietal operculum, respectively), which are well known to be involved in the processing of touch (see e.g. Basbaum, 2007; Eickhoff et al., 2006a, b), were highly informative regarding the tactile stimulation. The secondary somatosensory cortex cluster extended well into the contralateral posterior insular cortex which we have previously demonstrated to be activated by the type of pleasant, gentle brushing stimulation applied through stimulation of C-tactile afferents (Björnsdotter et al., 2009; Olausson et al., 2002, 2008). These touch-related regions all achieved highly significant univariate BOLD response increases in the GLManalysis (Table 3). Additionally, the ipsilateral (left) post-central gyrus was also found to differentiate brushing from rest, although a closer inspection of GLM *t*-map revealed a clear decrease in signal during the stimulation compared to baseline. This observation appears to correspond to the suppression of the ipsilateral primary somatosensory cortex previously described for vibrotactile stimulation (Hlushchuk and Hari, 2006).

The inferior frontal gyrus, implied to have a role in attending to tactile stimuli (Hagen et al., 2002), was identified bilaterally, although with a significant decrease in fMRI activity according to the GLM analysis, as was the superior frontal gyrus, which has been implicated in multimodal integration of spatial representations in body-centered coordinates (Galati et al., 2001).

Agreeing with reports in the literature, visual areas also contained information about the tactile stimulation. Bilateral peaks (although with a decrease in BOLD response in the right hemisphere) were found stretching throughout the medial to superior temporal gyrus, corresponding to medial temporal visual area 5 (V5), also known as the medial temporal area (MT) or medial superior temporal area (MST), which have been extensively studied in relation to tactile stimulation and found to activate by the application of a brush stimulus (see e.g. Hagen et al., 2002; Beauchamp et al., 2007). The MST has been reported as not being able to discriminate foot stimulation in a previous multivoxel analysis of vibrotactile stimulation, supporting the idea of a role for MST/STP in eye–hand coordination (Beauchamp et al., 2009). Our results, however, do not support this theory, although this should be interpreted with some caution since we used a moving stimulation. What appears to be the visual area 3 (V3) in the extrastriate cortex was also identified bilaterally (with a decrease in BOLD response according to the GLM analysis) and, although less implicated than V5, V3 has been previously described to deactivate in relation to tactile stimuli (Merabet et al., 2007).

The GLM analysis revealed that all identified regions were highly significant also in the univariate analysis (Table 3). Although there was a significant correlation between the magnitude of stimulation/rest separation between the uni- and multivariate analyses ($R = 0.66$; $p < 0.05$), there were some inconsistencies. For example, the right postcentral gyrus (primary somatosensory cortex) achieved the lowest univariate score, but ranked among top multivariate results.

Discussion

We propose a Monte Carlo approach to locally multivariate functional brain mapping, which, as shown on simulated data, is more sensitive in detecting discriminative voxels than the searchlight (Kriegeskorte et al., 2006) and requires dramatically reduced computer resources. In addition, we demonstrated that the method is highly applicable in whole-volume, multi-subject authentic fMRI brain mapping in combination with complex classifiers, such as support vector machines, on a standard desktop computer.

Surprisingly, the Monte Carlo approach achieved better voxel detection results on the simulated data than the searchlight algorithm. The improved performance was attributed to the local information smoothing induced by the process of averaging across all constellations in which each voxel has been included, and may be related to two effects. First, where the searchlight produced artificially well-localized maps with relatively higher values for voxels with more discriminative neighbors (despite equal information content), the information-average method yielded more homogenous values for voxels of equal information content. Second, the inherent variation in performance estimation between classification attempts, a problem which is of particular concern when few fMRI volumes are available, is reduced as a result of the averaging across numerous performance evaluations. It should be noted that the improved mapping performance does not depend on the sampling method: exhaustive searches may well benefit from information-averaging, whereas the Monte Carlo stochastic sampling approach merely exploits the information-average feature to

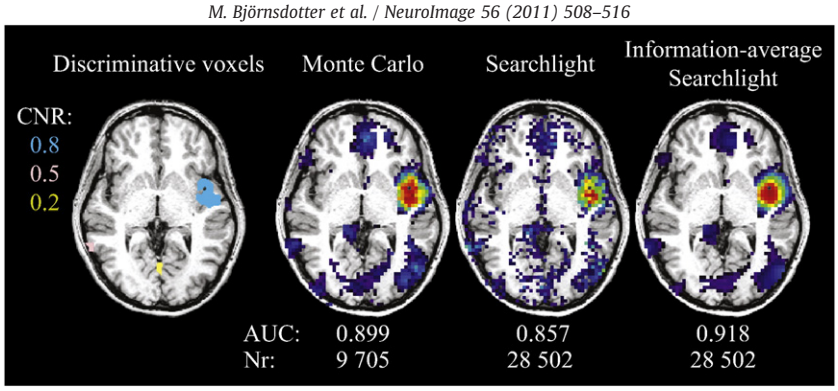


Fig. 4. Comparison of the maps produced using the Monte Carlo, searchlight and information-average searchlight algorithms on the simulated data with a search sphere volume of 0.5% and the RBF SVM kernel. The position and contrast-to-noise ratio (CNR) of the simulated discriminative voxels are shown in the leftmost panel. The Monte Carlo map is strikingly similar to the searchlight maps despite a reduction in the number of required information computations from 28 502 to 9 705. The maps are thresholded to show voxels with values above 0.5. CNR: contrast-to-noise ratio; Nr: Number of information computations (training and testing of the classifier); AUC: area under the receiver operating characteristic curve.

allow a reduction in computational load with little or no loss in mapping sensitivity compared to a non-averaged map.

In the current study we explored SVMs, which have dominated the fMRI multivoxel field (see e.g. Mourão-Miranda et al., 2005; LaConte et al., 2005; De Martino et al., 2008; Formisano et al., 2008; Haynes et al., 2007; Björnsdotter et al., 2009), to distinguish the experimental conditions. However, any arbitrary performance measure can be used with the Monte Carlo approach, including the metrics originally proposed for the searchlight (i.e. the Mahalanobis distance or average absolute *t*-map; Kriegeskorte et al., 2006), including supervised binary or multiclass classification methods, non-supervised learning approaches (such as self-organizing maps or multidimensional scaling), or techniques for abstracting brain activity patterns such that representational similarity analysis can be applied (Kriegeskorte et al., 2008). Regardless of performance metric, the Monte Carlo method requires correspondingly reduced computer resources compared to any exhaustive search.

The nonlinear SVM kernels obtained better mapping results than the linear kernels, although the difference was not substantial. The improved performance may be a result of higher classification scores obtained by the nonlinear SVM, differentiating informative voxels from noise more efficiently. Nonlinear classifiers have not been extensively studied in the context of fMRI classification (with the exception of e.g. Mørch et al., 1997; Davatzikos et al., 2005; Hanson et al., 2004; Polyn et al., 2005), partially since the improvement over linear classifiers is not conclusive (see e.g. Cox and Savoy, 2003). Discouraging results with nonlinear classifiers are likely due to limited number of training samples (and high number of voxels) which restricts the construction of complicated relationships with the voxels (Mørch et al., 1997). FMRI studies explicitly designed for pattern recognition analysis with fast, event-related paradigms yielding more samples in combination with effective voxel selection methods may increase the utility of nonlinear classifiers and possibly reveal patterns not distinguishable with linear classifiers. In such contexts, efficient mapping approaches are of fundamental importance due to the increased computational requirements of nonlinear classifiers.

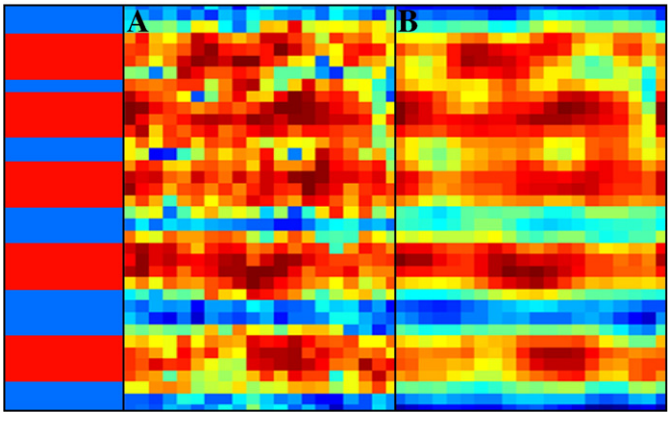


Fig. 5. Comparison of the ability of the conventional (A) and information-averaged (B) searchlight map to resolve closely spaced elements. The organization of the discriminative voxel regions (four voxels in width, separated by one, two, three and four voxels, respectively) are indicated in red in the leftmost panel. A search volume with a diameter of four voxels was used. Neither of the methods successfully resolved the regions separated by one or two voxels (top rows) whereas both clearly distinguished the lines separated by three and four voxels (bottom rows).

In the current implementation, a fixed-size spherical search volume was utilized to estimate local brain activity patterns, and the medium size search sphere of radius 9.6 mm (or 0.5% of the total brain volume) achieved the best overall voxel detection results. This is similar to the 9 mm search sphere radius used by Haynes et al. (2007), and substantially larger than the observation by Kriegeskorte et al. (2006) that a 4 mm sphere radius is near optimal. Closer investigation of the current results, however, revealed that larger discriminative regions were better identified with larger search spheres, indicating that the appropriate search volume size is linked to the spatial extent of the relevant brain area. Since any experimental condition may activate a number of differently sized brain regions, lack of an a priori hypothesis on the size of the active brain areas requires sequential iterations of the algorithm for varying search volumes—again pointing to the need for an efficient algorithm with minimal time requirements. Alternatively, the Monte Carlo algorithm may incorporate random-sized search volumes to reflect variations in brain region spatial extent and thus improve voxel detection performance. As pointed out by Hanke et al. (2009), the problem of inclusion of irrelevant voxels, such as white matter, can be addressed by applying the algorithm on data in a surface representation. Similarly, inhomogeneity effects (e.g. lower sensitivity in regions containing fewer voxels, such as close the border of the brain) will be mitigated.

Notably, the computational load decreased with increased search sphere volume for the Monte Carlo method, since fewer spheres are required to cover the brain volume (see Fig. 6B). This is in stark contrast with the exhaustive searches, where the number of information computations remained constant but the time requirements to train and test the classifier increased linearly with the number of included voxels.

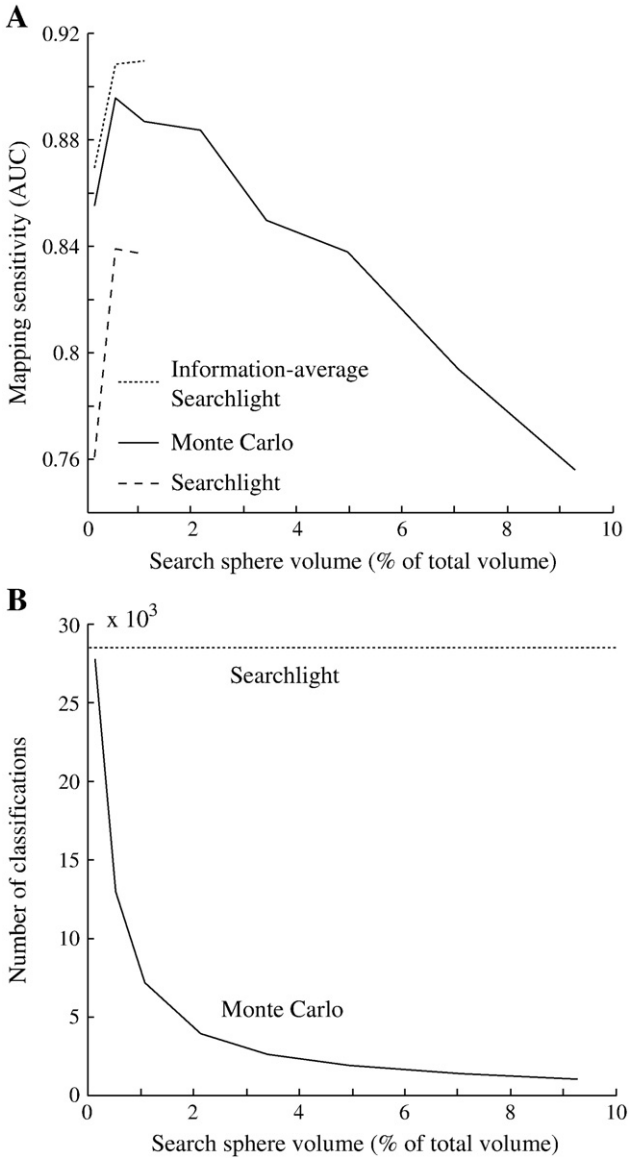


Fig. 6. The effect of search sphere volume (expressed in percentage of total brain volume) on A) the voxel detection performance (measured in area under the receiver operating characteristic curve; AUC) and B) number of required information computations for all methods and the linear support vector machine kernel. There was a dramatic decrease in number of required performance computations up to a volume of 2%. Note that for the searchlights only the three smallest search sphere volumes were investigated due to excessive time requirements for larger sizes. Similar although slightly higher results were obtained with the support vector machine with a radial basis function kernel.

Table 3
Peak group mean Monte Carlo map (average classification) values and Talairach coordinates for whole-brain activation peaks due to gentle brushing on the left thigh. BA, Brodmann area; **, GLM *t*-value $p<0.01$.

Anatomical Structure	Hemisphere	*x*	*y*	*z*	BA	Map Value	GLM*t*
Postcentral gyrus	Right	18	−40	55	3	0.76	3.39**
Postcentral gyrus	Left	−33	−28	52	3	0.74	−12.69**
Parietal operculum	Right	54	−25	19	40	0.82	14.84**
Parietal operculum	Left	−51	−22	22	40	0.77	11.74**
Inferior frontal gyrus	Left	−36	29	10	13	0.71	−4.94**
Inferior frontal gyrus	Right	39	22	12	13	0.72	−4.46**
Precentral gyrus	Left	−33	−25	49	4	0.75	−8.67**
Extrastriate cortex	Right	18	−76	37	19	0.75	−12.75**
Extrastriate cortex	Left	−18	−76	34	18	0.74	−10.55**
Middle temporal gyrus	Right	45	−47	−8	20	0.73	−7.27**
Middle temporal gyrus	Left	−51	−40	16	20	0.73	4.86**
Superior frontal gyrus	Right	15	59	22	10	0.72	−6.00**

The authentic data analyzed using the presented Monte Carlo method served to demonstrate the feasibility of the algorithm on multi-subject, whole-brain fMRI data on a standard desktop computer. Although not required in the current study, statistical map thresholds can be obtained through, for example, any permutation approach (see Nichols, 2003, for details), including the randomization test proposed by Kriegeskorte et al. (2006). Such tests involve the generation of a large number of permuted data maps, for which the improved efficiency of the Monte Carlo approximation should prove particularly appealing. Subsequently, the obtained *p*-values must be corrected for multiple comparisons, which becomes an increasingly severe problem with improved spatial resolution (i.e. in high-resolution fMRI). Kriegeskorte et al. (2007), however, points out that the multiple comparisons problem should be less severe for multivariate than univariate (e.g. GLM) maps due to the smoothing induced by local information integration (given appropriate correction methods, e.g. false-discovery rates, which accounts for such dependencies between voxels). Appropriate implementation of the algorithm in a compiled programming language, as opposed to the current scripted version, promises further computational efficiency and faster analysis times.

The often staggering time requirements for locally multivariate whole brain mapping restricts the practical utility of existing exhaustive algorithms and fast multivariate techniques are becoming increasingly acute, particularly as the functional spatial resolution improves (Kriegeskorte and Bandettini, 2007). The Monte Carlo mapping method is proposed as a time-efficient alternative without compromising voxel detection sensitivity.

Acknowledgments

This study was supported by the Swedish Research Council (grant 3548) and the Sahlgrenska University Hospital (grant ALFGBG 3161). We are grateful to Dr. Federico De Martino for assistance with the data simulation scripts.

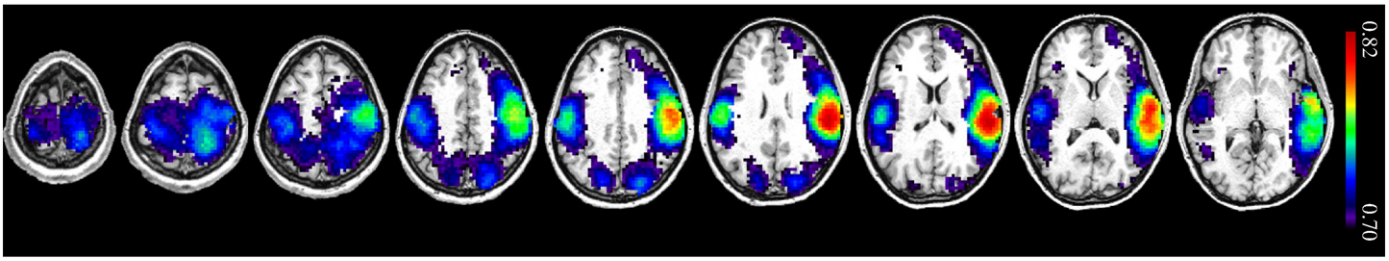


Fig. 7. Monte Carlo group analysis maps of gentle brushing on the left thigh, showing the mean classification performance across all nine subjects. A linear support vector machine was used for performance evaluation, and the algorithm was iterated 15 times with a search sphere volume of 0.9% of the total cortex volume.

References

Amar, J.G., 2006. The Monte Carlo method in science and engineering. Comput. Sci. Eng. 8 (2), 9–19.

Basbaum, A.I., 2007. The Senses. Elsevier, Amsterdam. [u.a.].

Beauchamp, M.S., Yasar, N.E., Kishan, N., Ro, T., 2007. Human MST but not MT responds to tactile stimulation. J. Neurosci. 27 (3), 8261–8267.

Beauchamp, M.S., LaConte, S., Yasar, N., 2009. Distributed representation of single touches in somatosensory and visual cortex. Hum. Brain Mapp. 30 (10), 3163–3171.

Bellman, R.E., 1961. Adaptive Control Processes. Princeton University Press, Princeton, NJ.

Björnsdotter Åberg, M., Wessberg, J., 2008a. An evolutionary approach to the identification of informative voxel clusters for brain state discrimination. IEEE J. Sel. Top. Sign. Process. 2 (6), 919–928.

Björnsdotter Åberg, M., Wessberg, J., 2008b. A multivariate approach to fmri activation detection using pattern recognition and information entropy on tactile data. 14th Annual Meeting of the Organization for Human Brain Mapping, Melbourne, Australia.

Björnsdotter, M., Löken, L., Olausson, H., Vallbo, Å., Wessberg, J., 2009. Somatotopic organization of gentle touch processing in the posterior insular cortex. J. Neurosci. 29 (29), 9314–9320.

Bode, S., Haynes, J.-D., 2009. Decoding sequential stages of task preparation in the human brain. NeuroImage 45 (2), 606–613.

Clithero, J.A., Carter, R.M., Huettel, S.A., 2008. Local pattern classification differentiates processes of economic valuation. NeuroImage 45 (4), 1329–1338.

Cox, D.D., Savoy, R.L., 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.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D.G., Acharyya, M., Loughead, J.W., Gur, R.C., Langleben, D.D., 2005. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. Neuroimage 28, 663–668.

De Martino, F., Gentile, F., Formisano, E., 2005. Local multivariate mapping of fMRI dynamics using support vector machines. 11th Human Brain Mapping Conference (Toronto, Canada).

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., Formisano, E., 2008. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. NeuroImage 43 (1), 44–58.

Edin, B., 2001. Cutaneous afferents provide information about knee joint movements in humans. J. Physiol. 531 (Pt 1), 289–297.

Efron, B., Tibshirani, R.J., 2009. An Introduction to the Bootstrap. Chapman & Hall Ltd.

Eickhoff, S.B., Amunts, K., Mohlberg, H., Zilles, K., 2006a. The human parietal operculum: II. Stereotaxic maps and correlation with functional imaging results. Cereb. Cortex 16 (2), 268–279.

Eickhoff, S.B., Schleicher, A., Zilles, K., Amunts, K., 2006b. The human parietal operculum: I. Cytoarchitectonic mapping of subdivisions. Cereb. Cortex 16 (2), 254–267.

Formisano, E., De Martino, F., Bonte, M., Goebel, R., 2008. Who is saying what? Brain-based decoding of human voice and speech. Science 5903 (322), 970–973.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J., 1994. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2 (4), 189–210.

Galati, G., Committeri, G., Sanes, J.N., Pizzamiglio, L., 2001. Spatial coding of visual and somatic sensory information in body-centred coordinates. Eur. J. Neurosci. 14 (4), 737–746.

Good, P.I., 2004. Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer Series in Statistics). Springer-Verlag New York, Inc, Secaucus, NJ, USA.

Hagen, M.C., Franzèn, O., McGlone, F., Essick, G., Dancer, C., Pardo, J.V., 2002. Tactile motion activates the human middle temporal/V5 (MT/V5) complex. Eur. J. Neurosci. 16 (5), 957–964.

Hammersley, J.M., Handscomb, D.C., 1975. Monte Carlo Methods. Methuen, London, UK.

Hanke, M., Halchenko, Y.O., Sederberg, P.B., Hanson, S.J., Haxby, J.V., Pollmann, S., 2009. PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. Neuroinformatics 7 (1).

Hanson, S.J., Matsuka, T., Haxby, J.V., 2004. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? NeuroImage 23 (1), 156–166.

Haynes, J.-D., Geraint, R., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7 (7), 523–534.

Haynes, J.-D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. 8 (5), 686–691.

Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., Passingham, R.E., 2007. Reading hidden intentions in the human brain. Curr. Biol. 17 (4), 323–328.

Hlushchuk, Y., Hari, R., 2006. Transient suppression of ipsilateral primary somatosensory cortex during tactile finger stimulation. J. Neurosci. 26 (21), 5819–5824.

Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8 (5), 679–685.

Kriegeskorte, N., Bandettini, P., 2007. Analyzing for information, not activation, to exploit high-resolution fMRI. NeuroImage 38 (4), 649–662.

Kriegeskorte, M., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. PNAS 103, 3863–3868.

Kriegeskorte, N., Formisano, E., Sorger, B., Goebel, R., 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci. 104 (51), 20600–20605.

Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis—connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2.

Kuncheva, Ludmila I., Rodriguez, Juan J., Plumpton, Catrin O., Linden, David E.J., Johnston, Stephen J., 2010. Random subspace ensembles for fmri classification. IEEE Trans. Med. Imaging 29 (2), 531–542.

LaConte, S., Strother, S., Cherkassky, V., Anderson, J., Hu, X., 2005. Support vector machines for temporal classification of block design fMRI data. Neuroimage 26 (2), 317–329.

Löken, L.S., Wessberg, J., Morrison, I., McGlone, F., Olausson, H., 2009. Coding of pleasant touch by unmyelinated afferents in humans. Nat. Neurosci. 12 (5), 547–548.

McLeish, Don L., 2005. Monte Carlo Simulation and Finance. John Wiley and Sons, New York, USA.

Merabet, L.B., Swisher, J.D., McMains, S.A., Halko, M.A., Amedi, A., Pascual-Leone, A., Somers, D.C., 2007. Combined activation and deactivation of visual cortex during tactile sensory processing. J. Neurophysiol. 97 (2), 1633–1641.

Metz, C.E., 1984. A new approach for testing the significance of differences between ROC curves measured from correlated data. Information Processing in Medical Imaging. Nijhoff, The Hague, pp. 432–445.

Mitchell, T.M., Hutchinson, R., Niculescu, R.S., Pereira, F., Wang, X., Just, M., Newman, S., 2004. Learning to decode cognitive states from brain images. Mach. Learn. 57 (1–2), 145–175.

Mørch, N., Hansen, L.K., Strother, S.C., Svarer, C., Rottenberg, D.A., Lautrup, B., Savoy, R., Paulson, O.B., 1997. Nonlinear versus linear models in functional neuroimaging: learning curves and generalization crossover. IPMI '97: Proceedings of the 15th International Conference on Information Processing in Medical Imaging. Springer-Verlag, London, UK, pp. 259–270.

Mourão-Miranda, J., Bokde, A.L., Born, C., Hampel, H., Stetter, M., 2005. Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data. NeuroImage 28 (4), 980–995.

Mourão-Miranda, J., Reynaud, E., McGlone, F., Calvert, G., Brammer, M., 2006. The impact of temporal compression and space selection on SVM analysis of single-subject and multi-subject fMRI data. Neuroimage 33 (4), 1055–1065.

Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Statistical Methods in Medical Research 12, 419–446.

Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. 10 (9), 424–430.

Olausson, H., Lamarre, Y., Backlund, H., Morin, C., Wallin, B.G., Starck, G., Ekholm, S., Strigo, I., Worsley, K., Vallbo, A.B., Bushnell, M.C., 2002. Unmyelinated tactile afferents signal touch and project to insular cortex. Nat. Neurosci. 5 (9), 900–904.

Olausson, H.W., Cole, J., Vallbo, Å., McGlone, F., Elam, M., Krämer, H.H., Rylander, K., Wessberg, J., Bushnell, M.C., 2008. Unmyelinated tactile afferents have opposite effects on insular and somatosensory cortical processing. Neurosci. Lett. 436 (2), 128–132.

Polyn, S.M., Natu, V.S., Cohen, J.D., Norman, K.A., 2005. Category-specific cortical activity precedes retrieval during memory search. Science 310 (5756), 1963–1966.

Sato, J.R., Fujita, A., Thomaz, C.E., Martin Mda, G., Mourão-Miranda, J., Brammer, M.J., Amaro Junior, E., 2009. Evaluating SVM and MLDA in the extraction of discriminant regions for mental state prediction. Neuroimage 46 (1), 105–114.

Staeren, N., Renvall, H., De Martino, F., Goebel, R., Formisano, E., 2009. Sound categories are represented as distributed patterns in the human auditory cortex. Curr. Biol. 19 (6), 498–502.

Stokes, M., Thompson, R., Cusack, R., Duncan, J., 2009. Top-down activation of shape-specific population codes in visual cortex during mental imagery. J. Neurosci. 29 (5), 1565–1572.

Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J., 2002. Least Squares Support Vector Machines. World Scientific.

Vallbo, A.B., Olausson, H., Wessberg, J., Norrsell, U., 1993. A system of unmyelinated afferents for innocuous mechanoreception in the human skin. Brain Res. 628 (310), 301–304.

Vallbo, A.B., Olausson, H., Wessberg, J., 1999. Unmyelinated afferents constitute a second system coding tactile stimuli of the human hairy skin. J. Neurophysiol. 81 (310), 2753–2763.

Wang, Z., Childress, A.R., Wang, J., Detre, J.A., 2007. Support vector machine learning-based fMRI data group analysis. Neuroimage 36 (4), 1139–1151.

Zaidi, H., Sgouros, G., 2002. Therapeutic Applications of Monte Carlo Calculations in Nuclear Medicine. Taylor and Francis Ltd., London, Great Britain.

Zhou, X.-H., Obuchowski, N.A., McClish, D.K., 2002. Statistical Methods in Diagnostic Medicine. Wiley, New York.