NeuroImage 56 (2011) 593–600

Contents lists available at ScienceDirect

NeuroImage

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

A comparison of volume-based and surface-based multi-voxel pattern analysis

Nikolaas N. Oosterhof [a,*], Tobias Wiestler [a,b], Paul E. Downing [a], Jörn Diedrichsen [a,b]

[a] School of Psychology, Bangor University, UK
[b] Institute of Cognitive Neuroscience, University College London, UK

ARTICLE INFO

Article history:
Received 4 December 2009
Revised 13 April 2010
Accepted 30 April 2010
Available online 4 June 2010

ABSTRACT

For functional magnetic resonance imaging (fMRI), multi-voxel pattern analysis (MVPA) has been shown to be a sensitive method to detect areas that encode certain stimulus dimensions. By moving a searchlight through the volume of the brain, one can continuously map the information content about the experimental conditions of interest to the brain.

Traditionally, the searchlight is defined as a volume sphere that does not take into account the anatomy of the cortical surface. Here we present a method that uses a cortical surface reconstruction to guide voxel selection for information mapping. This approach differs in two important aspects from a volume-based searchlight definition. First, it uses only voxels that are classified as grey matter based on an anatomical scan. Second, it uses a surface-based geodesic distance metric to define neighbourhoods of voxels, and does not select voxels across a sulcus. We study here the influence of these two factors onto classification accuracy and onto the spatial specificity of the resulting information map.

In our example data set, participants pressed one of four fingers while undergoing fMRI. We used MVPA to identify regions in which local fMRI patterns can successfully discriminate which finger was moved. We show that surface-based information mapping is a more sensitive measure of local information content, and provides better spatial selectivity. This makes surface-based information mapping a useful technique for a data-driven analysis of information representation in the cerebral cortex.

Introduction

Conventional fMRI studies use univariate (voxel-by-voxel) analyses (Friston et al., 1995) to identify brain regions that respond more to one experimental condition than another (e.g. seeing faces vs. houses). In contrast, multi-voxel pattern analysis (MVPA) uses multiple voxels and can be more sensitive to distinguish between experimental conditions because it considers patterns across a group of voxels that may respond weakly but consistently differently between conditions (Haynes and Rees, 2005b; Norman et al., 2006). Examples of topics where MVPA has been applied successfully are distinguishing between observing different object categories (Haxby et al., 2001; Kriegeskorte et al., 2008; Cox and Savoy, 2003), invisible differences between line orientations (Haynes and Rees, 2005a), intentions (Haynes et al., 2007), and observed natural scenes (Kay et al., 2008; Naselaris et al., 2009). In addition it has been used to make inferences about the underlying neural representations (Peelen et al., 2006; Haxby et al., 2001). MVPA therefore allows researchers to map the information content of distributed patterns of brain activity.

Some MVPA studies use a region-of-interest (ROI) based approach, where first a group of voxels is selected based on anatomical criteria (for example, coordinates described in the literature, or anatomical landmarks) or functional criteria (for example, a separate localizer scan). For the pattern of activity across this set of voxels, MVPA then measures how much information is present for distinguishing between experimental conditions (e.g. the presentation of different natural scenes). This information content can be quantified either by considering correlations between patterns, or by using pattern classification methods; see Mur et al. (2009), Haynes and Rees (2005b) and Norman et al. (2006) for an overview.

While this approach is suitable for hypothesis-driven questions, it requires researchers to define a-priori the boundaries of the region. To address this, Kriegeskorte et al. (2006) introduced 'information mapping', which – similar to univariate methods – yields a whole-brain map. Information mapping can be seen as the repeated application of ROI-based MVPA. Traditionally, each voxel in the brain is taken as the center of an ROI, which means that there are as many ROIs as voxels in the brain. Considering a certain center voxel, the ROI consists of a sphere-shaped region – the searchlight – around that voxel, with a certain a-priori defined radius. MVPA is applied to all voxels in the ROI and the resulting statistics (typically correlation differences or pattern classification accuracy) represents the information content of that region and is assigned to the center voxel.

This procedure is repeated for every voxel in the brain, i.e. each voxel serves as center voxel once. As in univariate analyses, the resulting information maps can be spatially normalized and then submitted to a

* Corresponding author.
E-mail address: n.oosterhof@bangor.ac.uk (N.N. Oosterhof).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.04.270

second level (group) analysis. The interpretation is different from univariate maps however; voxel intensity does not represent (change in) measured activity, but how much information the distributed pattern of voxels surrounding that voxel represents jointly.

Recent conventional univariate fMRI studies have employed surface-based analysis for visualization and inter-subject alignment. First a wire-frame model of the cortical surface is reconstructed using a high-resolution anatomical image and subsequently functional data from the grey matter is projected on this surface (Essen and Drury, 1997; Essen et al., 2001; Dale et al., 1999; Essen and Dierker, 2007). Further analysis proceeds then using only the data on the surface. Inter-subject alignment is achieved after the surface has been inflated to a sphere, thereby removing inter-subject variability that stems from different folding patterns of the cortical surface. For univariate analyses it has been reported that for the group level, surface-based spherical alignment techniques are superior to volume-based techniques (Fischl et al., 1999). Thus, compared to volume-based approaches, surface-based techniques allow for a better inter-subject alignment based on surface features, easier visualization, and results in a dramatically reduced search region for statistical inferences as it only includes grey matter voxels (Fischl et al., 1999).

It seems a sensible idea therefore to combine the advantages of information mapping and univariate surface-based activation analyses. The advantage of surface-based analysis, better visualization and improved inter-subject alignment, should generalize to multivariate analysis. However, surface representations become relevant for MVPA in an additional aspect that is not present for univariate analyses: the definition of the searchlight region around a center voxel.

In volume-based voxel selection, the searchlight is defined as a sphere around a center voxel without regard to the underlying folding structure of the neocortex. Surface-based information mapping should allow for better selection of informative voxels. It differs from volume-based voxel selection in two ways: (1) voxels for MVPA are restricted to the grey matter only, which reduces noise and enhances classification performance, and (2) a more appropriate distance metric that takes into account the folded nature of the cortical sheet. This results in a more neurologically plausible measure of information content and better spatial selectivity, which should make, at least in theory, a surface-based searchlight superior to volume-based approaches.

While a few researchers have reported the use of surface-based information mapping approaches (Wiestler et al., 2009; Soon et al., 2009; Oosterhof et al., 2009), to our knowledge there are no reports that show that for real data this approach has advantages compared to volume-based approaches. In the present paper, we compare the volume-based and surface-based approaches in an example dataset. We focus here on the comparison of the volume- and surface-based searchlight definition. For inter-subject alignment we used a spherical technique for all analyses. We show how the surface-based voxel selection improves both spatial selectivity and (when correcting for the number of voxels) also classification performance. Importantly, the surface-based approach results in a better dissociation of information content between two spatially neighbouring regions than the volume-based approach. We illustrate this in our dataset by showing that primary somatosensory (S1) represents single digit finger presses better than primary motor cortex (M1), a result that does not become apparent in a similar volume-based analysis.

Altogether, our results demonstrate that surface-based information mapping has significant advantages over volume-based approaches, which makes it a useful tool for answering data-driven questions about the involvement of regions that process or represent information in the human brain.

Surface-based information mapping

Information mapping, in general, consists of two steps: voxel selection and computing information content (correlations or classification accuracies). While the latter is studied extensively in the field of machine learning (Michie and Spiegelhalter, 1994 for example), the focus of this paper is to study the influence of different methods of voxel selection onto information-based mapping.

By voxel selection we mean the process of choosing the neighbourhood voxels around a center voxel. The computation of the information content for the center voxel is then based on these voxels. We distinguish between the 'seed domain', the set of locations that serve as a center for the searchlight, and a 'data domain', the set of voxels that are used for MVPA.

In the traditional searchlight, the seed domain and target domain are identical. This relationship is not necessary, for example one could decide to take voxels in the seed domain from a volume with a different spatial resolution and/or space (for example, original versus Talairach template space) than the volume in the data domain. This distinction is important for the case of surface representations of the cortex, where the elements in the seed domain are not voxels but nodes on the surface. Here, we followed this approach by using a wire-frame model from the cortical surface, where for a center node the neighbouring nodes are selected, and then the voxels adjacent to these nodes are used for MVPA.

Conceptually, our surface-based searchlight approach differs from the traditional volume-based version in two ways, both of which may improve the quality of the information mapping: grey matter selection and a surface-based distance measure. As mentioned earlier, in traditional information mapping (see Fig. 1a), a center position is chosen and all voxels surrounding the center are selected for MPVA. This includes not only grey matter – which produces the BOLD signal we are interested in – but also white matter, cerebral spinal fluid, and other tissue whose measured signal is not expected to correlate with the experimental conditions of interest. Our first step is therefore to inclusively mask grey matter voxels (Fig. 1b), which should improve MPVA in informative regions because features (voxels) that carry no information are excluded from the analysis.

However, this method does not take into account the folded nature of the cortical sheet. Some voxels will be close to the center voxel with a standard Euclidean distance measure but far away with a geodesic distance measure (i.e., distance measured along the cortical surface). Where Euclidean distance is used to select voxels, the spatial selectivity for estimating the information content of a local region on the cortical sheet will be reduced. This occurs especially when the center position is located inside a sulcus, which causes voxels on the opposite side of the sulcus to be selected as well. This will dilute the measure local information content, because functionally different regions will be considered in the same classifier. Thus, on the above grounds, our second step involves applying a geodesic distance measure that respects the curvature of the cortex in computing the distance from center position to voxels (Fig. 1c). We note that geodesic distance measures that are unbiased with respect to the local surface curvature (concavity or convexity) require an intermediate surface that lies in between the pial–grey matter and grey matter–white matter boundaries. Furthermore, the difference between Euclidean and geodesic distances crucially depends on the local geometry of the cortical sheet and becomes more prominent as the searchlight radius increases.

To compare surface-based searchlights with the volume-based variant, we used a dataset from seven participants who performed single digit movements of the right hand while they were scanned with fMRI. In both variants, MVPA was used to classify which of the four fingers they pressed during each run.

First, we compared the traditional volume approach with the geodesic surface approach for a whole-brain analysis. Both approaches show that regions in the primary motor (M1) and somatosensory (S1) cortices, near the hand area, represent most information content about which finger was pressed. Second, we compared an anatomically-defined ROI that contains the hand area in M1 and S1. Our results show

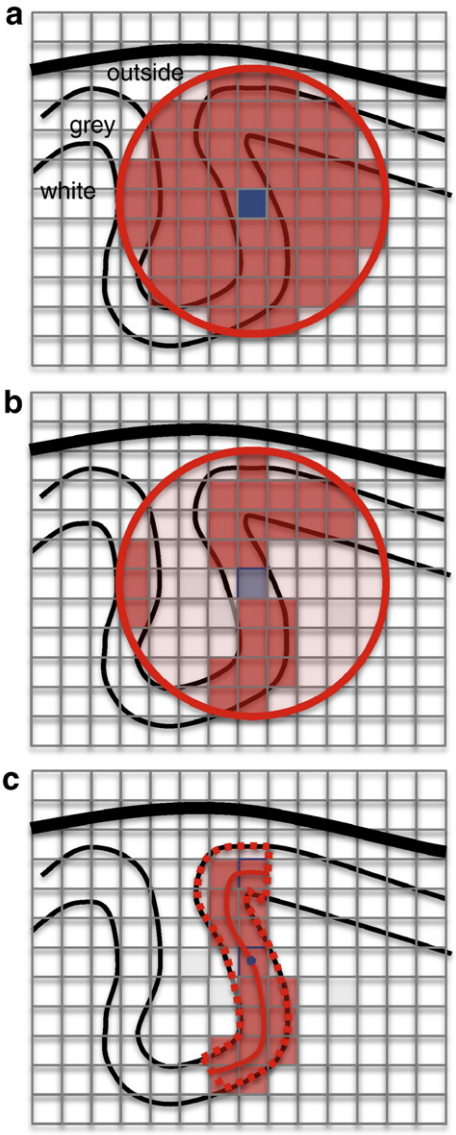

Fig. 1. Comparison of voxel selection methods. (a) Schematic representation of a brain slice, with white matter, grey matter, and cerebro-spinal fluid indicated. The lines represent the white matter/grey matter boundary, the grey matter/pial surface boundary, and the skull. With the traditional volume-based voxel selection method, a voxel (blue) is taken as the center of a sphere (red; represented by a circle), and all voxels within the sphere are selected for further pattern analysis. (b) An improvement over (a), in that only grey matter voxels are selected. The grey matter can either be defined using a probability map, or the cortical surface reconstruction. A limitation however is that voxels that are close in Euclidean distance but far in geodesic distance (i.e. measured along the cortical surface), are included in the selection, as illustrated by the three voxels on the left. (c) Using surface reconstruction, the white matter–grey matter and grey matter–pial surfaces are averaged, resulting in an intermediate surface that is used to measure geodesic distances. A node on the intermediate surface (blue) is taken as the center of a circle (red; represented by a solid line), and corresponding voxels in between the white matter and pial surfaces are selected.

that, when corrected for the number of voxels, the surface-based method shows higher classification accuracies. More importantly, we show that the volume-based approach shows artifactual regions of high accuracy in both M1 and S1. In contrast, the surface-based approach shows that the difference between finger-activation pattern is more pronounced (operationalized by higher classification accuracies) in S1 than in M1. This indicates that the surface-based approach has better spatial selectivity.

Method

Participants

Two female and five male neurological healthy volunteers participated in the study. All participants were right handed and their age ranged from 20 to 22 years. The ethics committee of the School of Psychology, Bangor University, approved all procedures of the study.

Apparatus

To study the representation of digit movements on fingertips with fMRI we used a non-magnetic finger box. The box had five piano-style keys and the forces applied to the keys were recorded by quantum tunnelling composite pills (QTC-pills). Because we controlled the finger box from outside the scanning room, a filter panel in the wall prevented radiofrequency leakage. The visual instructions and feedback were projected from outside the scanner room on a screen, which was viewed by the participants through a mirror.

Scan acquisition

The imaging data were acquired on a Phillips Achieva 3 T scanner (Philips, Best Netherlands). For the functional scans we used an echo plane imaging sequence (EPI) with a voxel size of $2.0 \times 1.8 \times 1.8$ mm^3. We achieved coverage of the superior part of the cortex with 38 axial slices (no gaps), using sensitivity encoding (SENSE) with a factor of 2 and a TR of 2 s. A scan sequence started with 4 dummy scans followed by 128 data images. T1 weighted structural images were acquired with a volumetric MPRAGE sequence using a voxel size of $1 \times 1 \times 1$ mm^3.

Procedure

Before the scan acquisition all participants underwent a training session of 4 runs in which they were familiarized with the task. During the scan participants performed 7 runs. All runs consisted of 16 trials (see Fig. 2). Each trial started with a red letter on the screen to announce the digit and participants made 5 isometric key presses with the indicated fingers. The presses were paced by the appearance of a white asterisk on the screen (every 1.35 s) and turned green if participants pressed the correct finger and red otherwise. Trials lasted for 8.1 s and each finger was repeated 4 times in random order per condition in a run.

Imaging data analysis

We analysed the functional image data using the SPM5 toolbox (http://www.fil.ion.ucl.ac.uk/spm/) and custom written routines in Matlab. We first realigned the slices temporarily to correct for the

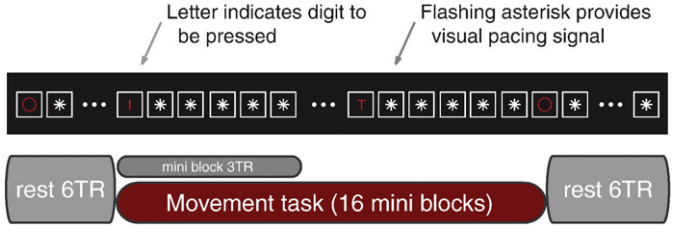

Fig. 2. Illustration of experimental procedure. Participants were scanned with fMRI while they got visual instructions to press one of their fingers (thumb, index, middle, or pinky). Each run started with a 6 TR (12 s) rest period, followed by 16 trials of 3 TRs (6 s each), and ended with a 6 TR (12 s) rest period. Finger presses were recorded with a non-magnetic finger box with five piano-style keys.

ascending order of slice acquisition. Afterwards the images were spatial realigned to the first functional image of the scan using a six-parameter rigid-body transformation. To remove slow varying trends in the data we made use of a high pass filter with a cut off frequency of 1/128 s. Finally we fitted the data of each voxel using a General Linear Model, with regressors that represented the digit responses for single fingers during a run. These regressors were boxcar functions (length 8.1 s) that indicated the 4 trials of each finger within each run and then were convolved with a standard hemodynamic response function.

Anatomical preprocessing

Cortical surface models were reconstructed using Freesurfer (Fischl et al., 1999), which provides an automated reconstruction procedure that yields an outer surface (the pial surface–grey matter boundary) and an inner surface (the grey matter–white matter boundary). These surfaces consist of vertices, edges and surfaces (the topology) and a set of coordinates that refer to the vertices. Because the two surfaces have the same topology, the coordinates of two surfaces can be simply averaged to obtain an intermediate surface along which geodesic distances are measured. Further analyses proceeded with AFNI and SUMA (Cox, 1996; Saad et al., 2004). To avoid interpolation of the functional data, the surfaces are brought in correspondence to the functional data by estimating the required affine transformation. For group analysis, the surfaces were inflated to a sphere, aligned to a standard sphere, and resampled to a standard topology (an icosehedron in which each of the twenty triangles is subdivided in 10,000 triangles, using AFNI's MapIcosehedron). This ensured that each node on the standardized surfaces represented a corresponding surface location across participants; therefore, group analyses could be conducted using a node-by-node analysis without interpolation of functional data or classification accuracy scores. For all analyses even for the volume-based searchlight we used spherical alignment to average results between different individuals. The intermediate and inflated surfaces were transformedinto Talairach template space (Talairach and Tournoux, 1988) to correct for overall brain size in distance measures, and averaged across participants for presentation of group analysis results.

Voxel selection

We used different searchlight radii of 4, 6, 8, 10, and 12 mm. The volume-based algorithm was similar to work reported earlier. For the surface-based algorithm, a center node on the intermediate surface was chosen. Using Peyré's (2008) toolbox that implements Kimmel and Sethian's (1998) method of computing geodesic distances efficiently, all nodes within the searchlight radius are selected. Grey matter voxels surrounding the selected nodes are selected as follows: for each selected node on the intermediate surface a corresponding line is constructed that connects corresponding points on the outer and inner surface. (The length of this line equals the local cortical thickness). For a given line, ten equidistant steps were taken from the outer to the inner surface, and after each step the voxel that contains the current position is selected. This procedure was repeated for each line, yielding a set of selected voxels (where duplicates are removed) for the ROI that corresponds to the center node. This procedure was repeated for each center node and stored for further processing (MVPA).

Multi-voxel pattern analysis

To investigate the representation of digit movements in the cortex, we used multi-voxel-pattern-classification with a Linear Discriminant Analysis classifier. Inputs for the classifier consisted of the 4 (digits) $\times$ 7 (runs) β-estimates from a set of selected voxels. To normalize the β-estimates, for each run we centered the β-estimates around zero, i.e. subtracted the mean of the β-estimates across the run. To avoid circular analysis (double dipping) in the classification (Kriegeskorte et al., 2009), we used cross-validation where the classifier was tested on the four β-estimates from one run after it was trained on the other six runs. This was repeated seven times, where each time the classifier was tested on a different run. Because typically the number of voxels in selected regions was larger than the number of β-estimates from the GLM, the estimate of the covariance matrix is rank deficient. We therefore regularized the matrix by adding the identity matrix scaled by one percent of the mean of the diagonal elements. We note that we have also used a Support Vector Machine (not reported here) that yielded qualitatively similar classification results.

Group analysis

All whole-brain group analyses were performed on the standardized surface. We transformed the classification accuracies for each node to z-scores using the normal approximation of the binominal distribution.[1] These z-statistics were then used in the group statistic and tested using a t-test against 0 (corresponding to chance performance). The t-statistics were thresholded at $t_6 = 2.45, p = 0.05$ two-tailed, uncorrected for multiple comparisons. Using Monte Carlo simulations, we then determined the critical cluster size, such that the family-wise error-rate was $\alpha = 0.05$. As accuracies from neighbouring nodes share many voxels, they are not independent but spatially autocorrelated (resulting in a smooth accuracy map). To account for this, the smoothness of the residual accuracy maps was estimated and applied to random Gaussian data. For the surface-based analyses we used a Matlab toolbox that implements the estimation of smoothness along the surface (Chung et al., 2005).

Results

The whole-brain analysis (Fig. 3; intended for a qualitative comparison) showed that both approaches were well able to locate M1/S1 as the most prominent region with robust classification performance. Both methods also identified a smaller cluster in the supplementary motor and secondary sensory areas. Focusing on the M1/S1 region, the surface-based map shows the most reliable information representation in a single region on the posterior wall of the central sulcus. In comparison, the volume-based map shows more similar information representation in the anterior and posterior wall of the central sulcus, and in the posterior and anterior wall of the postcentral gyrus. If we assume that most information is indeed represented in the anterior wall of the postcentral sulcus – as suggested by the geodesic method – this is exactly the pattern we would expect: for the volume-based map, spheres whose center are at the anterior banks of the central and postcentral sulcus do also contain voxels in the posterior wall of the central sulcus, and therefore show more similar classification accuracies. In other words, the volume-based results seem due to 'cross-talk' across the sulci and gyri and are an artifact of the volume-based voxel selection method. Altogether, this suggests that a surface-based searchlight has better spatial specificity than the volume-based variant.

As mentioned earlier (Fig. 1), voxel selection methods for volume- and surface-based information mapping differ in two aspects: (1) the former includes both grey and non-grey matter voxels, while the latter includes grey matter voxels only; and (2) surface-based methods use a distance metric that respects the folding of the cortical

[1] We note that for small numbers of trials or low accuracies such approximation may deviate significantly from the exact z-score.

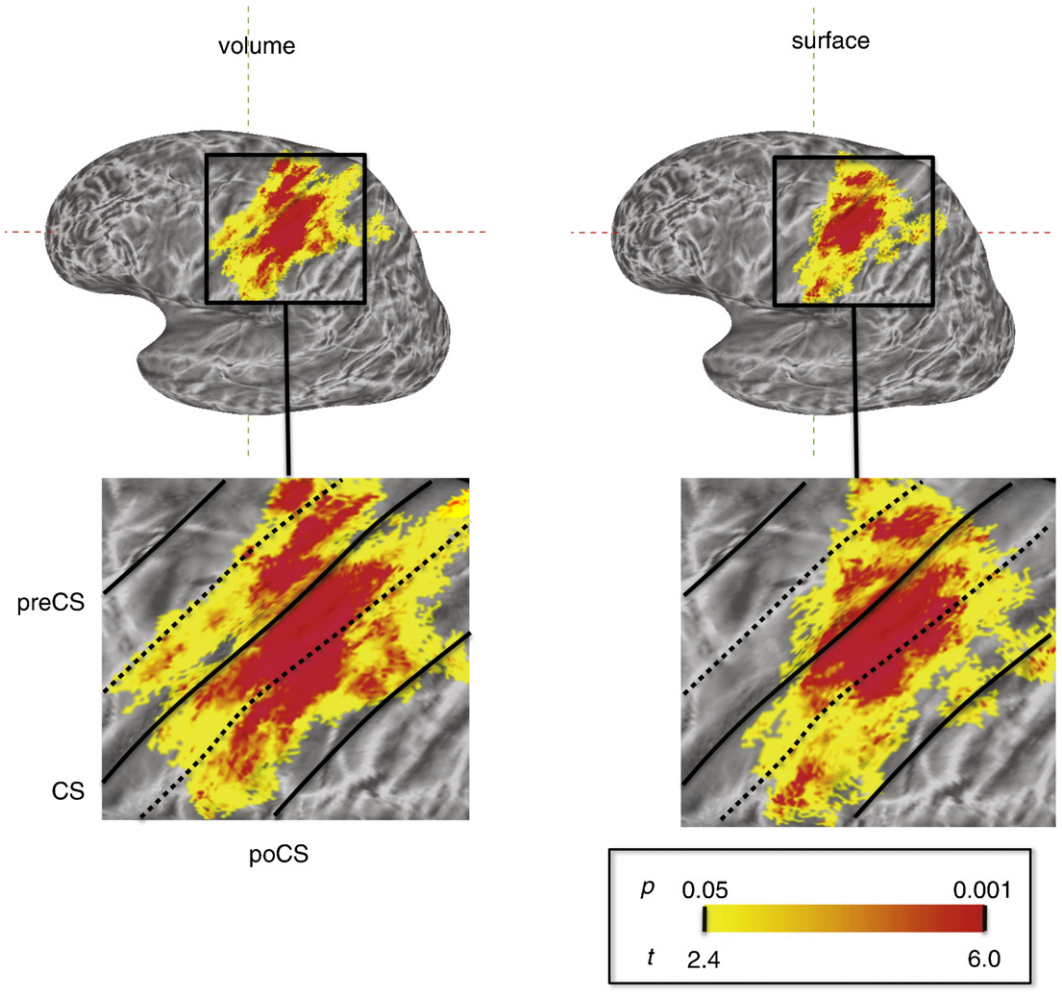

Fig. 3. Whole-brain group analysis results. Group analysis results from finger press classification using traditional volume-based voxel selection with Euclidean distance metric and no voxel masking (left; see Fig. 1a), and surface-based voxel selection with grey matter voxel masking (right; see Fig. 1c). Colors indicate *t* values for classification accuracies across participants. Maps are thresholded at $p = 0.05$ uncorrected and $p = 0.05$ cluster-size corrected (see Methods). Finger presses could most reliably be decoded from the primary motor and somatosensory regions. Both maps are averaged across individuals using surface-based normalization. The volume-based results (left) show similar patterns across the anterior and posterior wall of the central sulcus and the anterior wall of the postcentral sulcus. The surface-based results (right) show most reliable pattern classification in a single locus on the anterior wall of the postcentral sulcus. This suggests that the similarity across sulcus walls in the volume-based analysis are due to an artifactual 'cross-talk' effect from voxel selection across the sulci and gyri, and that the surface-based method has better spatial selectivity. Insets: more detailed view, with labeling of the major sulcus and gyri. Abbreviations: CS, central sulcus; preCS, precentral sulcus; poCS, postcentral sulcus.

sheet. Different selection methods also lead to different numbers of voxels selected, which influences accuracy in a non-linear manner.

To investigate the effect of different voxel selection methods quantitatively, we compared three voxel selection methods: Euclidean distance in the volume without voxel masking ('volume'; Fig. 1a), Euclidean distance in the volume with grey matter voxel masking ('Euclidean'; Fig. 1b), and geodesic distance along the surface with grey matter voxel masking ('geodesic'; Fig. 1c). We used different searchlight sizes, with radii of 4, 6, 8, 10, and 12 mm for all voxel selection methods, and also radii of 14, 16, 18, 20, and 22 mm for the geodesic method because, for a given radius, this method selects much fewer voxels than the other voxel selection methods. We selected a region that covered approximately equal areas of M1 and S1 (Fig. 4a). For easier comparison, the seed domain, the centers of the spheres/circles, was the set of nodes on the intermediate surface for all methods. This means that in the 'volume' method, distances were computed between center nodes and the centers of voxels in the volume; in the 'Euclidean' and 'geodesic' methods, distances were computed between center nodes and other nodes on the surface.

Fig. 4b shows the mean classification accuracies (across all nodes in the seed region) as a function of the number of voxels selected in the searchlight region for the different radii and the three voxel selection methods. First, we tested the effect of grey matter voxel masking. A repeated-measures ANOVA for radii of $4, 6, \ldots, 12$ mm showed that the Euclidean had higher accuracies than the volume method, $F_{1,6} = 40, p < 0.001$, which means that grey matter voxel selection improves classification performance.

Generally the accuracies increase with a larger searchlight radius (i.e. more voxels), $F_{4,24} = 75, p < 0.001$, as reported earlier (Li et al., 2007). Because the geodesic method uses less voxels – especially for large radii – than the other methods, a fair comparison of distance metric (Euclidean versus geodesic) requires correcting for the number of voxels.

To compare the effect of distance metric, corrected for the number of voxels, we fitted exponential functions ($x \mapsto c_1 + c_2 \cdot \exp(x - c_3)$) to the accuracies for each selection method and each participant. These functions explained on average 98.5% of the variance. We interpolated the fitted curves for searchlight sizes of $100, 200, \ldots, 800$ voxels, and found that for all searchlight sizes grey matter voxel masking yield

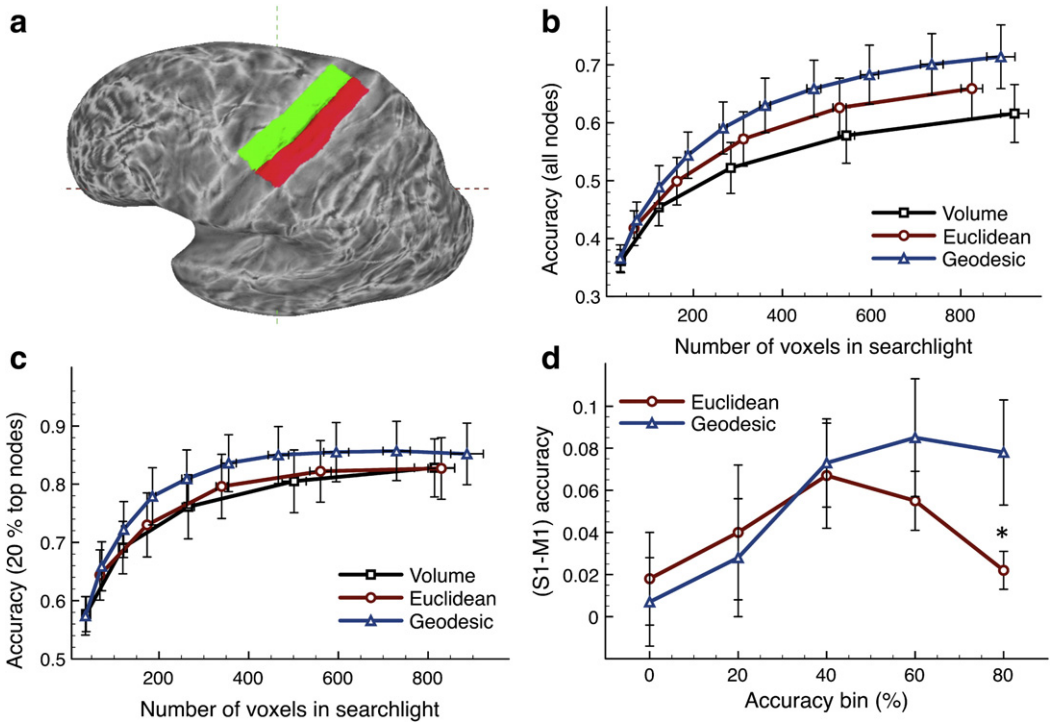

Fig. 4. Region-of-interest comparison of voxel selection accuracies. (a) Surface nodes in region covering the hand representation in primary motor (M1; green) and somatosensory (S1; red) cortex is selected. (b) Classification accuracies for three voxel selection methods: Euclidean distance metric without voxel masking ('volume'), Euclidean distance metric with grey matter voxel masking ('Euclidean'), and geodesic distance metric with grey matter voxel masking ('geodesic'). The horizontal axis shows the average number of voxels in the searchlight, and they vertical axis the average classification accuracies across the M1 and S1 region together. For the volume and Euclidean voxel selection method, the measures correspond to searchlight radii of 4,6,…,12 mm; for the geodesic method, to radii of 4,6,…,22 mm. When corrected for the number of voxels, classification accuracies are lower for the volume and Euclidean than for the geodesic voxel selection method. (c) The same figure as (b), but with the top 20% nodes (sorted on accuracy) selected. (d) Comparison of informative regions in M1 and S1. Accuracies across nodes are sorted and binned separately in M1 and S1 in five 20% wide bins (horizontal axis), and the difference between the average accuracies (vertical axis) in M1 and S1 is computed for each bin. Compared to the Euclidean method, the geodesic method yields a larger difference in accuracies between S1 and M1 for the top 20% accuracy bin, which shows that the geodesic method provides better spatial specificity.

higher classification accuracies than no masking, and also that the geodesic distance measure yields higher classification accuracies than the Euclidean distance measure, $\min(t_6) = 3.08$, $ps < 0.02$ for all searchlight sizes.

Because we were interested in the most informative sub-regions within the ROI, we considered the center nodes with the top 20% highest accuracies for each method and radius (Fig. 4c). The increase in accuracies with larger searchlight size seem to saturate at around 600 voxels. Grey matter masking did not improve classification accuracies, $\max(t_6) < 1.4, ps > 0.1$ for searchlight sizes $\geq$300voxels, which might be due to the fact that searchlights associated with the nodes with the top 20% classification accuracies had relatively few non-grey matter voxels selected. Conversely, application of a geodesic distance measure showed higher classification accuracies than the Euclidean distance metric, $\text{in}(t_6) = 1.96, ps < 0.05$ for all searchlight sizes.

To investigate the spatial separation between M1 and S1 (see whole-brain analysis, above) in more detail, we compared the Euclidean and geodesic distance metric for the 10 mm searchlight radius data (using grey matter masking in both methods). Based on the group analysis (Fig. 3), we assumed that regions in S1 hold more information for finger presses than M1. If the geodesic measure separates M1 and S1 better than the Euclidean measure, then the geodesic approach should find clusters with higher accuracies in S1 than in M1. For the Euclidean approach these difference should be diluted.

For each participant, distance metric, and subregion (M1 and S1), the set of nodes were sorted by accuracy and assigned to one of five 20%-percentile bins. For each of these five bins, we computed the difference between the mean accuracy in M1 and S1. As shown in Fig. 4d, bins with lower accuracies (0–80%) showed no difference between the two distance metrics. Importantly, the planned comparison in the bin with the top 20% accuracies showed that this difference was significant, $t_6 = 2.06, p = 0.043$. Post-hoc analysis showed that this was reliable for different sizes of the top bin, $ps < 0.01$ for selecting up to the top 7% of the nodes and $ps < 0.05$ for the top 21%. This corroborates the observations from the whole-brain map, which suggested that information about finger presses is most precisely represented in S1. When looking at the areas of highest performance, the difference between these two regions (i.e. their functional specialization) was observed better using a geodesic approach.

Discussion

In this paper we have compared volume- and surface-based information mapping of fMRI data. The surface-based methods inherits a number of advantages from univariate surface-based activation mapping—easier visualization, and less correction for multiple comparisons because inferences are only drawn on grey matter voxels. Furthermore, surface-based inter-subject alignment improves the spatial specificity of group analyses (Fischl et al., 1999). Here, we focus on a further difference between the two approaches: the method of voxel selection that determines the shape of the searchlight. Specifically, we looked at two aspects. First, surface-based information mapping should lead to a better selection of informative voxels in a searchlight region by inclusive masking of grey matter

only. Second, it employs a distance metric that takes into account the folded nature of the cortical sheet.

Using an exemplary data set, we showed that surface-based voxel selection indeed improves sensitivity for informative patterns, both through the restriction to grey matter voxels, and through a surface-based distance metric. We show that classification accuracies are generally higher with grey matter selection than without. This shows that there was much more information contained in grey matter voxels than in non-grey matter such as cerebral spinal fluid or white matter. There has been some debate about the nature (such as the role of functional organization of vasculature in the cortex; Gardner, 2010; Thompson et al., 2010; Kriegeskorte et al., 2009a) and spatial scale (such as the effects of data smoothing; de Beeck, 2010) of the informative signal used in MVPA. We note that our results do not resolve these matters beyond the observation that grey matter contains more information than non-grey matter.

We also showed that within a selected region-of-interest, application of a geodesic, surface-based distance measure led to higher accuracies than when selecting the same number of voxels using a Euclidean, volume-based distance measure. We did not specifically predict this result, because a Euclidean distance measure allows the mixture of voxels from both side of a sulcus (Fig. 1b). This should make performance worse when the seed is located in area that contains a lot of information, but at the same time should improve performance when the seed is located outside such a region but close (with Euclidean distance) to neighbouring regions that contain information. The differences between the voxel selection methods may be explained by the fact that our ROI contained mostly well-performing regions.

Most importantly, however, we showed better spatial specificity for surface-based compared to volume-based voxel selection. Specifically, the surface-based method allowed us to distinguish between information about finger presses in the primary motor and somatosensory cortex, whereas in the volume-based approach the contributions of these two are less distinguishable due to 'cross-talk' of common voxels selected across the sulci and gyri (Figs. 3 and 4d). The results cannot be explained by differences in co-registration, because we used spherical alignment for all analyses.

We note that irrespective of the voxel selection method used, classification accuracies increase with an increasing searchlight radius for all the radii tested (Fig. 4b). When the top 20% of the nodes with highest classification accuracy are selected, however, we found a saturation effect for larger radii where classification accuracies remain stable (Fig. 4c). If there are focal regions with high information content, an increase in searchlight size will also tend to select more less-informative voxels, and it is possible that for larger radii than tested here, classification accuracies will decrease with an increase in searchlight radius. Furthermore, while our findings are consistent with prior suggestions that accuracies stop to increase for large radii (Mur et al., 2009), it is clear that a larger searchlight radius also means less spatial selectivity. For example, a 12 mm radius searchlight may yield higher classification accuracies than a 6 mm one, but the interpretation of a significant node is that the distributed pattern of measured activity in the voxels in a 12 mm radius (rather than 6 mm) around that node contain information about the trial conditions of interest. Therefore, a searchlight radius entails a compromise between spatial selectivity and classification accuracy. This loss of specificity, however, may be less pronounced using surface-based voxel selection.

In sum, these results indicate that surface-based voxel selection improves classification accuracies and spatial selectivity, compared to the traditional volume-based method. This makes surface-based information mapping a useful technique for a data-driven analysis of information representation in the cerebral cortex.

• We have implemented surface-based voxel selection in a new matlab toolbox, available from http://surfing.sourceforge.net.

Acknowledgments

We would like to thank Marius Peelen and Martijn van Koningsbruggen for the helpful comments on an earlier draft of this manuscript. This research was supported by the National Science Foundation (BSC 0726685), the ESRC, and the Wales Institute of Cognitive Neuroscience. NNO was supported by a fellowship awarded by the Boehringer Ingelheim Fonds.

References

Chung, M., Robbins, S., Dalton, K., Davidson, R., Alexander, A., Evans, A., 2005. Cortical thickness analysis in autism with heat kernel smoothing. Neuroimage 25 (4), 1256–1265.
Cox, R., 1996. Afni: software for analysis and visualization of functional magnetic resonance neuroimages.
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.
Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. i. Segmentation and surface reconstruction. Neuroimage 9 (2), 179–194.
de Beeck, H.P.O., 2010. Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fmri analyses? Neuroimage 49 (3), 1943–1948.
Essen, D.V., Dierker, D., 2007. Surface-based and probabilistic atlases of primate cerebral cortex. Neuron 56 (2), 209–225.
Essen, D.V., Drury, H., 1997. Structural and functional analyses of human cerebral cortex using a surface-based atlas. Journal of Neuroscience 17 (18), 7079.
Essen, D.V., Lewis, J., Drury, H., Hadjikhani, N., Tootell, R., Bakircioglu, M., Miller, M., 2001. Mapping visual cortex in monkeys and humans using surface-based atlases. Vision Research 41 (10–11), 1359–1378.
Fischl, B., Sereno, M., Tootell, R., Dale, A., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Human Brain Mapping 8 (4), 272–284.
Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C., Frackowiak, R., et al., 1995. Statistical parametric maps in functional imaging: a general linear approach. Human Brain Mapping 2 (4), 189–210.
Gardner, J.L., 2010. Is cortical vasculature functionally organized? Neuroimage 49 (3), 1953–1956 URL http://dx.doi.org/10.1016/j.neuroimage.2009.07.004.
Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293 (5539), 2425–2430.
Haynes, J.-D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nature Neuroscience 8 (5), 686–691.
Haynes, J.-D., Rees, G., 2005b. Predicting the stream of consciousness from activity in human visual cortex Jul Current Biology 15 (14), 1301–1307.
Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., Passingham, R., 2007. Reading hidden intentions in the human brain. Current Biology 17 (4), 323–328.
Kay, K.N., Naselaris, T., Prenger, R.J., Gallant, J.L., 2008. Identifying natural images from human brain activity. Nature 452 (7185), 352–355 Mar.
Kimmel, R., Sethian, J., 1998. Computing geodesic paths on manifolds. Proceedings of the National Academy of Science of the USA 95 (15), 8431–8435.
Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. Proceedings of the National Academy of Science of the USA 103 (10), 3863–3868.
Kriegeskorte, N., Mur, M., Ruff, D., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P., 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60 (6), 1126–1141.
Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. Nature Neuroscience 12 (5), 535–540 May.
Kriegeskorte, N., Cusack, R., Bandettini, P., 2009a. How does an fmri voxel sample the neuronal activity pattern: Compact-kernel or complex-spatiotemporal filter? Neuroimage 1–12 URL http://dx.doi.org/10.1016/j.neuroimage.2009.09.059.
Li, S., Ostwald, D., Giese, M., Kourtzi, Z., 2007. Flexible coding for categorical decisions in the human brain. Journal of Neuroscience 27 (45), 12321–12330.
Michie, D., Spiegelhalter, D.J., 1994. Machine learning, neural and statistical classification.
Mur, M., Bandettini, P.A., Kriegeskorte, N., 2009. Revealing representational content with pattern-information fmri—an introductory guide. Social Cognitive and Affective Neuroscience 4 (1), 101–109.
Naselaris, T., Prenger, R.J., Kay, K.N., Oliver, M., Gallant, J.L., 2009. Bayesian reconstruction of natural images from human brain activity. Neuron 63 (6), 902–915.
Norman, K., Polyn, S., Detre, G., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fmri data. Trends in Cognitive Science 10 (9), 424–430.
Oosterhof, N.N., Wiggett, A.J., Diedrichsen, J., Tipper, S.P., Downing, P.E., 2009. Multivoxel pattern analysis reveals mirror-like responses in human parietal and occipitotemporal cortex. Poster Presented at the Annual Meeting of the Society of Neuroscience, Chicago, IL, USA.
Peelen, M.V., Wiggett, A., Downing, P.E., 2006. Patterns of fmri activity dissociate overlapping functional brain areas that respond to biological motion Jan Neuron.

Peyré, G., 2008. Toolbox fast marching — a toolbox for fast marching and level sets computations. retrieved on 16 may 2009 from http://www.ceremade.dauphine.fr/peyre/matlab/fast-marching/content.html.

Saad, Z., Reynolds, R., Argall, B., Japee, S., Cox, R., 2004. Suma. 2nd IEEE International Symposium on Biomedical Imaging Macro to Nano, 2, pp. 1510–1513.

Soon, C., Namburi, P., Goh, C., Chee, M., Haynes, J., 2009. Surface-based information detection from cortical activity. NeuroImage 47 (Supplement 1), S79 organization for Human Brain Mapping 2009 Annual Meeting.

Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Thompson, R., Correia, M., Cusack, R., 2010. Vascular contributions to pattern analysis: comparing gradient and spin echo fmri at 3t Neuroimage 1–44 URL http://dx.doi.org/10.1016/j.neuroimage.2010.03.061.

Wiestler, T., McGonigle, D., Diedrichsen, J., 2009. Sensory and motor representations of single digits in the cerebellum. Poster Presented at the Annual Meeting of the Society of Neuroscience, Chicago, IL, USA.