NeuroImage xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

Cortical surface-based searchlight decoding

Yi Chen [a,b,1], Praneeth Namburi [c,1], Lloyd T. Elliott [b,d], Jakob Heinzle [b], Chun Siong Soon [a,b,c], Michael W.L. Chee [c], John-Dylan Haynes [a,b,e,f,*]

[a] Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[b] Bernstein Center for Computational Neuroscience Berlin and Charité - Universitätsmedizin Berlin, Germany
[c] Duke-NUS Graduate Medical School, Singapore, Singapore
[d] Gatsby Computational Neuroscience Unit, University College London, UK
[e] Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany
[f] Graduate School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany

ARTICLE INFO

Article history:
Received 17 March 2010
Revised 15 July 2010
Accepted 19 July 2010
Available online xxxx

ABSTRACT

Local voxel patterns of fMRI signals contain specific information about cognitive processes ranging from basic sensory processing to high level decision making. These patterns can be detected using multivariate pattern classification, and localization of these patterns can be achieved with searchlight methods in which the information content of spherical sub-volumes of the fMRI signal is assessed. The only assumption made by this approach is that the patterns are spatially local. We present a cortical surface-based searchlight approach to pattern localization. Voxels are grouped according to distance along the cortical surface—the intrinsic metric of cortical anatomy—rather than Euclidean distance as in volumetric searchlights. Using a paradigm in which the category of visually presented objects is decoded, we compare the surface-based method to a standard volumetric searchlight technique. Group analyses of accuracy maps produced by both methods show similar distributions of informative regions. The surface-based method achieves a finer spatial specificity with comparable peak values of significance, while the volumetric method appears to be more sensitive to small informative regions and might also capture information not located directly within the gray matter. Furthermore, our findings show that a surface centered in the middle of the gray matter contains more information than to the white–gray boundary or the pial surface.

Introduction

Cortical processing relies on computations within local and global networks of neurons. Functional magnetic resonance imaging (fMRI) indirectly measures such neural activations via the blood oxygen level dependent (BOLD) signal (Logothetis, 2008; Logothetis and Wandell, 2004). It is common practice to detect 'mass univariate' changes in the BOLD signal within rather large regions (on the order of centimeters). However, in recent years it has been shown that the local, distributed fine-scale patterns of voxel activations (on the order of millimeters) contain information beyond the univariate standard analysis. Multi-voxel pattern analysis (MVPA) has been successfully used in many neuroimaging studies (Haynes and Rees, 2006; Norman et al., 2006). MVPA allows assessing the information in a set of voxels considered jointly. The sensitivity and statistical stability offered by MVPA methods have been testified by numerous studies in various topics (Cox and Savoy, 2003; Eger et al., 2008; Formisano et al., 2008; Haxby et al., 2001; Haynes et al., 2005; Haynes and Rees, 2005, 2006; Kamitani and Tong, 2005; Knops et al., 2009; Macevoy and Epstein, 2009; Williams et al., 2008).

To further our understanding of neuronal information processing, we must investigate not only the extent to which distributed patterns in BOLD signal inform us about experimental conditions but also where in the brain such information is processed. A promising approach to pattern localization is to use multivariate searchlights to construct information-based maps of the brain (Haynes et al., 2007; Kriegeskorte et al., 2006). In this method, a 3-dimensional (3D) spherical volume is defined around each voxel and all voxels situated within the sphere are jointly analyzed with MVPA to extract information about some experimental condition or cognitive state of the subject. This volumetric searchlight method provides a spatially unbiased estimate of the information contained in local patterns of activity around every voxel location in the volume and as such can be fashioned as a tool to guide detailed investigation into local patterns. The unbiased estimation of information produced by the searchlight method distinguishes it from other biased approaches that use regions of interest (ROIs) (Haynes and Rees, 2006; Pereira et al., 2009). The efficacy of multivariate searchlight methods has been

* Corresponding author. Charité - Universitätsmedizin Berlin, Bernstein Center for Computational Neuroscience, Philippstrasse 13, Haus 6, 10115 Berlin, Germany.
E-mail address: haynes@bccn-berlin.de (J.-D. Haynes).
[1] Yi Chen and Praneeth Namburi are co-first authors.

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.07.035

successfully demonstrated in many recent studies (Bode and Haynes, 2009; Haynes et al., 2007; Kriegeskorte et al., 2006; Soon et al., 2008).

The use of the spherical searchlight as a technique for evaluating information contained in local patterns rests on the assumption that the patterns are distributed volumetrically in the brain. However, cortical neurons are located within the two highly convoluted sheets of gray matter varying in thickness from 1 to 4.5 mm (Fischl and Dale, 2000). In fact, most of the signals measured with fMRI reside within the gray matter (Jin and Kim, 2008). Its thickness is approximately the same as the standard side length of the voxels typically measured in fMRI. Thus, a 3D-spherical searchlight might sample information from cortical regions that are close in Euclidian space, but relatively far apart with respect to the distance induced by geodesics on the cortical surface. Searchlights located near the longitudinal fissure might even sample information from both hemispheres. This implies that the volumetric searchlight technique may include voxels from different anatomical structures such as regions across spatially non-contiguous cortices. This is likely to deteriorate the spatial specificity of the results. Moreover, the inclusion of many non-gray matter voxels in a volumetric searchlight is likely to increase the noise in the fMRI pattern. Due to its limited temporal resolution, statistical inference about fMRI data is made based on a relatively small number of samples. Thus, any increase in noise reduces the power of MVPA techniques.

Motivated by the problem of pattern localization and preservation of structural topology, we propose the use of 2-dimensional (2D), i.e. surface-based, searchlight methods: 2-dimensional disks with a given radius are centered at and extended from each point on the cortical surface. The disks are deformed to match the cortical curvature, preserving the distance metric induced by the surface geodesic. The set of voxels intersecting each disk are then analyzed jointly by MVPA techniques. This extraction of voxels from the 2-dimensional cortical surface assures that voxels included in each such surface-based searchlight are close to each other with respect to anatomical structure. In addition, the number of non-gray matter voxels in each searchlight is reduced. We illustrate this method on an fMRI data set in which we decode the category of visually presented rotating objects. Compared to a volumetric searchlight, the surface-based searchlight has a similar statistical reliability, but it improves the spatial specificity of multivariate analyses.

Methods and materials

We describe the experimental procedures followed by a detailed account of the surface-based and volumetric searchlight approaches.

Experimental design and visual stimulation

Twelve healthy subjects (seven females, five males, mean age 25.2) participated in the study and gave written informed consent to the test procedure. All subjects were right-handed and had normal or corrected to normal visual acuity. The experiment was approved by the local ethics committee and was conducted according to the Declaration of Helsinki.

Each experimental run consisted of 60 trials. In each trial, a rotating object was dynamically presented (60 frames per second) at the center of a screen, using a Python script with the Python OpenGL® binding (http://pyopengl.sourceforge.net/). The objects were scaled to have the same length along their longest axis, and subtended a view angle of 7.68°. Each trial lasted for 4 s, followed by a fixation-only period of 4, 6 or 8 s (randomized). Objects rotated around a randomly changing axis at four different angular velocities (0, 0.5, 1.5, 4.5 cycles/s), where angular velocity 0 corresponded to a static object presented in a randomly chosen view (Fig. 1A). Each of the 12 experimental conditions (3 different objects in 4 rotating velocities) was presented 5 times per run. In order to minimize potential effects of sequencing that might affect the later analysis, the order of trials was randomized in every run independently for each subject. Throughout the whole experiment subjects were engaged in a Landolt C fixation task (Fig. 1B). The fixation stimulus subtended a view angle of 0.154° and randomly changed its opening side (left or right) every 2 s. Subjects were asked to indicate the sides of the opening by a button press with the corresponding hand. Example images of the three objects used in the experiment are shown in Fig. 1C.

Data acquisition and preprocessing

Functional MRI scans (gradient-echo EPI) were acquired on a Bruker 3 T (MedSpec 30/100) scanner with standard head coil. The fMRI volumes contained 30 axial slices to completely cover the occipital and temporal lobe for every subject (TR = 2000 ms, TE = 30 ms, in plane resolution 3×3 mm, slice thickness 2 mm with 1 mm gap). Five runs of 302 volumes were acquired. The first two images of each run were discarded to allow for magnetic saturation effects. In addition, for every subject, a sagittal T1-weighted anatomical scan was acquired in a separate session on a different scanner (3T Siemens Trio MR System, TR = 1300 ms, TE = 3.93 ms, FOV = 256×240 mm^2, slab thickness 192 mm). The anatomical images had a resolution of $1 \times 1 \times 1$ mm^3 and were used for cortical surface reconstruction.

All functional images were corrected for head-motion and realigned to the first functional image for each subject (SPM2, Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). After the coregistration, a standard hemodynamic response function model was fitted to the data in order to estimate the statistical parameters (beta values in SPM) for each of the 12 experimental conditions (3 objects × 4 velocities).

Construction of cortical surface-based searchlight

The construction of the 2-dimensional surface-based searchlight involved several steps which are described in detail below. First, the cortical surface was extracted. Then, we defined disks with a fixed radius around each point on the surface deformed to respect the surface curvature. Finally, all voxels of the functional images that contained some part of each disk were assigned to the surface-based searchlight corresponding to that disk's centre. Fig. 2 illustrates the searchlight extraction mechanism.

Circular searchlights on the cortical surface

High resolution ($1 \times 1 \times 1$ mm) T1-weighted whole brain images were acquired for every subject. Meshes of the white matter and pial surfaces of the cortex were reconstructed individually for each subject using the Freesurfer software package (Dale et al., 1999; Fischl et al., 1999a). A third surface in the middle of the gray matter was defined as the surface lying equidistant from both the pial and white–gray boundary surfaces. This 'graymid' surface was generated by inflating the white matter surface by 50% towards the pial surface for each subject, also using Freesurfer.

The surfaces generated by Freesurfer are defined as discrete meshes and can be represented as graphs with weights corresponding to the lengths of the edges between vertices. Based on this weighted representation, pair-wise geodesic distances between vertices of the mesh can be approximated with the length of the shortest paths on the graph by using a modified Dijkstra algorithm (Dijkstra, 1959; Fischl et al., 1999a). For a given radius r ($r = 9$ mm in this study), we constructed disks around each vertex v of the surface mesh by including all vertices u such that the geodesic distance between u and v was less than r (Fig. 2A). This procedure defined a searchlight structure for each vertex as the set of neighboring vertices within the radial distance. Note that neighborhood is defined with the intrinsic

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

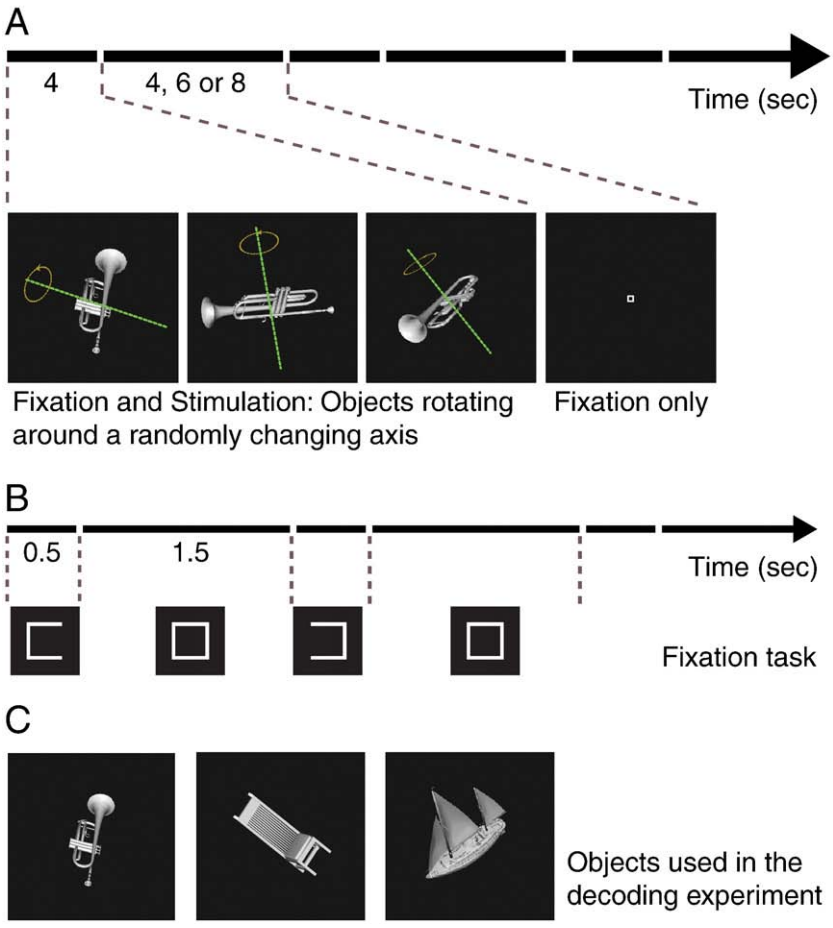

Fig. 1. Experimental design. (A) Rotating objects (trumpet, chair and sailing boat) were dynamically presented for 4 s (at 60 frames/s). The axis of rotation changed randomly during the trial. Between individual presentations of objects (trials) a fixation only period of random length (4, 6 or 8 s) was introduced. (B) During the whole experiment (trial duration and interleaved fixation periods) subjects were engaged in a fixation task. Small black squares illustrate the central part of the screen. Images are highly enlarged compared to (A). The size of one fixation square corresponds to 0.154° of visual angle. (C) Example images of the three objects: trumpet, chair and sailing boat.

metric of the surface, therefore respects the cortical anatomical structure.

Surface searchlight in functional image space and extraction of activity patterns

To use the surface-based searchlight structure for pattern extraction from functional images the sets of vertices were projected into the functional image space. We first calculated the transformation parameters by coregistering (SPM2) the functional images to the T1-weighted image which was used for surface reconstruction. The functional image voxels belonging to the surface searchlight were extracted by assigning to each vertex its nearest neighbor voxel within three-dimensional space (Euclidian distance). The list of voxels obtained in this way constitutes the set of features for a single surface-based searchlight (see Fig. 2B and C). By assigning nearest neighbor voxels to vertices, we avoided any smoothing or interpolation between neighboring voxels, which might have affected the patterns and thereby also might have affected classification accuracies.

Because the surface meshes created by Freesurfer have a much finer spatial resolution than the functional images, several vertices within the searchlight corresponded to the same voxel. To avoid redundant features in the classification analysis, we disallowed multiple occurrences of the same voxel within a single searchlight.

We constructed the searchlight structure on three different surfaces, the white–gray boundary (W), graymid (G) and pial (P), all of which are inflations of the original white–gray boundary surface. Since tiny structures on the original white matter surface often induce errors in the estimation of geodesic distance, a smoothed version of the white–gray boundary computed with Freesurfer (Fischl et al., 1999a) was used instead. For each of the surfaces we included all voxels within a certain distance from the center vertex in the corresponding searchlight. Note that now distance was calculated on the respective surface and the 3-D coordinates of that surface were used to find the corresponding voxels. Moreover as the original surface was inflated to form these three surfaces, there is a one-to-one correspondence between the vertices of the three surfaces. We thus were able to combine the searchlights of one vertex of all or any combination of the three extracted surfaces and created searchlights of contiguous surfaces (the set union of voxel indices) to obtain the following combinations of 2D-surface searchlights: W–G, G–P, and W–G–P. These searchlights from combined surfaces can be viewed as 3-D cylindrical searchlights.

Multivariate pattern analysis

Once the searchlight structure for every vertex was constructed and mapped to functional image space, each vertex v was associated with a list of all voxels that contained at least one vertex that belonged to the searchlight centered at v. From this list the set of features associated with a given vertex was extracted from the parameter estimates for the three different objects. For every center vertex we

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

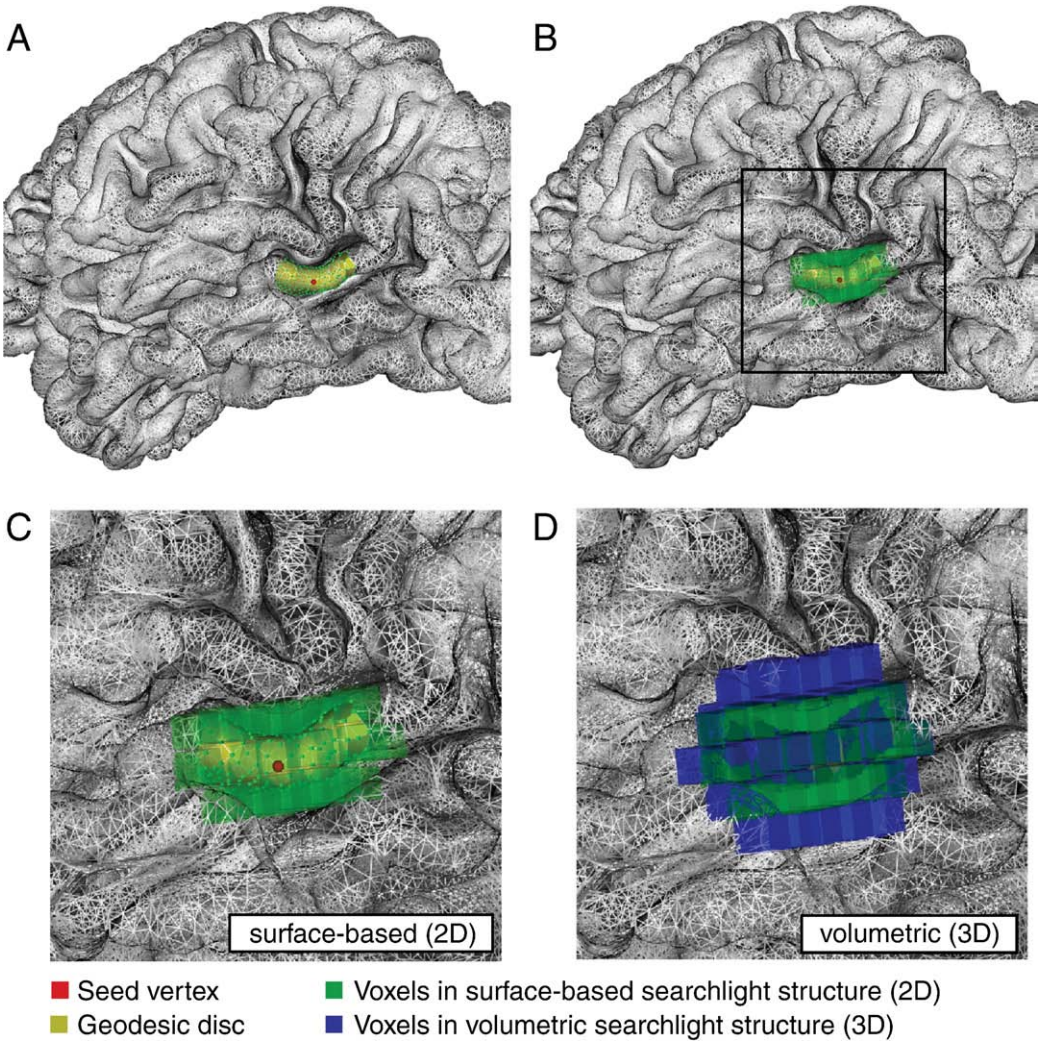

Fig. 2. Construction of surface-based searchlights. (A) Example seed vertex (red dot) and corresponding geodesic disk (faces in yellow and vertices in green) around it (radius<9 mm). (B) Voxels (translucent green cubes) in functional image space that are included in the searchlight structure defined by the geodesic disc. (C) Magnified view of the surface-based searchlight structure and the voxels defined by it. Note that the voxels preserve the surface structure, even though the voxels of the functional images have a coarser resolution ($3\times3\times3$ mm^3) than the average distance between vertices (<1 mm). (D) The voxel cluster (translucent blue cubes) defined by the volumetric searchlight structure at the same location and with the same radius is overlaid on the surface-based searchlight structure. Note that the surface-based searchlight voxels (green) constitute a subset of the volumetric searchlight (blue).

used a linear support vector machine (SVM, LIBSVM implementation, http://www.csie.ntu.edu.tw/~cjlin/libsvm) to perform a 5 fold cross-validated classification over the 5 functional runs (Pereira et al., 2009). In each of the cross-validating folds, 12 samples (4 from each category) were used as training set and 3 samples (one from each category) served as the testing set. The cross-validation procedure yielded an average accuracy map for each individual subject, reflecting the distribution of decodable information about visual object category on the cortical surface. Note that to avoid potential confounds from sequencing effects, individual trials were randomized independently for each run and subject, and hence the parameter estimates should be unbiased.

For the purpose of this paper, we restricted the analysis to a single angular velocity (0.5 cycles/s). This angular velocity was slow enough to give the subjective impression of a smoothly rotating object. Faster speeds were omitted because they result in flickering effects, which may have influenced visual processing. The static condition is prone to leading to significant decoding in early visual areas, which are not related to objects but to the spatial arrangement of contrast edges. A detailed account of the influence of different rotation speeds goes beyond the scope of this paper and will be published separately.

Volumetric searchlight method

In order to assess the differences between a volumetric and a surface-based searchlight approach, we also performed a standard volumetric searchlight analysis (see Bode and Haynes, 2009, for a detailed description). We extracted around each voxel all voxels within the same radius as for the surface-based method (9 mm). Fig. 2D shows a comparison between a surface-based and a volumetric searchlight. We then proceeded with the identical classification procedure as for the surface-based decoding. Again, a 5 fold cross-validation with linear SVM was used to estimate the decoding accuracy for each voxel. The resulting accuracies were then assigned to the nearest-neighbor vertices on the surface mesh providing an accuracy map of volumetric searchlight decoding that we could directly compare to the surface-based decoding.

Group analysis

To assess the statistical validity of the decoding among visual categories on the group level we used Freesurfer to construct the average white matter and pial surfaces together with an average anatomical volume for the whole group (Fischl et al., 1999b). Based on

the cortical curvature, we then aligned the accuracy map for every individual subject onto the average surface and smoothed it with a Gaussian kernel (FWHM = 9 mm). Finally, a *t*-test was used to compare the accuracy maps against chance level for every vertex. We report the resultant significance as *p*-value maps.

In addition to the *t*-test, permutation tests for classification analysis (Golland and Fischl, 2003) were conducted for all subjects. One thousand random permutations of class labeling, including the true one from the experiment, were used for each subject at every location, where locations are either vertices (2D) or voxels (3D). For each permuted labeling, we computed the accuracy with the same 5-fold cross-validation procedure as described above. After mapping the accuracies from the individual subjects to the average subject cortical surface, we obtained 1000 randomly sampled accuracies representing the unknown accuracy distribution for each subject at each vertex of the average cortical surface.

Based on these random samples of single subjects we estimated the empirical significance for the group mean accuracy of the 12 subjects. The group mean accuracy is a statistic based on a random sample of size *N* taken independently from the *N* subjects ($N = 12$). The significance of the true mean resulting from the correct labeling in all subjects is thus the probability that a random sample has a mean value greater than or equal to the true mean accuracy. To estimate this probability we drew 10^5 random samples of combinations of *N* accuracies (one from each of the *N* subjects) in a uniformly randomized way, and calculated the percentile of the samples that had a mean greater or equal to the true mean. This percentile is an approximation to the statistical significance with a minimal *p*-value of 10^{-5}.

Results

In this part we report the results of decoding object categories from visual cortex using the surface-based searchlight. We report the main results using surface-based searchlights of radius 9 mm defined on the graymid surface in the middle of the gray matter. Differences between the surface-based and volumetric methods are illustrated by looking at the spatial specificity of the results. We present the decoding results from surfaces of different cortical depths and their combinations and, finally, compare the results for various sizes of surface-based and volumetric searchlights.

Object decoding in surface-based and volumetric local patterns

For both the surface-based on the graymid and the volumetric searchlight method object category could be decoded from regions in the ventral occipito-temporal cortex (Fig. 3). Consistent with previous findings (Carlson et al., 2003; Levy et al., 2001; Malach et al., 2002), the category-specific information was located within the collateral sulcus (CoS). While the surface-based decoding did not yield any other significant region, the volumetric approach resulted in two additional regions of significant decoding in earlier visual areas (presumably V2 or V3). However, contrary to most previous studies we did not find any object-specific information in patterns decoded from the lateral surface of the occipital cortex (e.g. in the Lateral Occipital Complex, LOC). This is presumably due to differences in the stimulation paradigm (i.e. our study used objects that rotated within a single trial).

Although the distribution of object information across the cortical surface was similar in both volumetric and surface-based decoding, there are important differences between the results derived from the two searchlight techniques. Fig. 3 illustrates differences in the spatial spread of the informative regions. On the one hand, the volumetric searchlight showed larger significant regions. Some information was found in regions where the surface-based searchlight did not reveal any significant accuracy, as in early visual cortex. On the other hand, closer examination of the region around the CoS reveals some

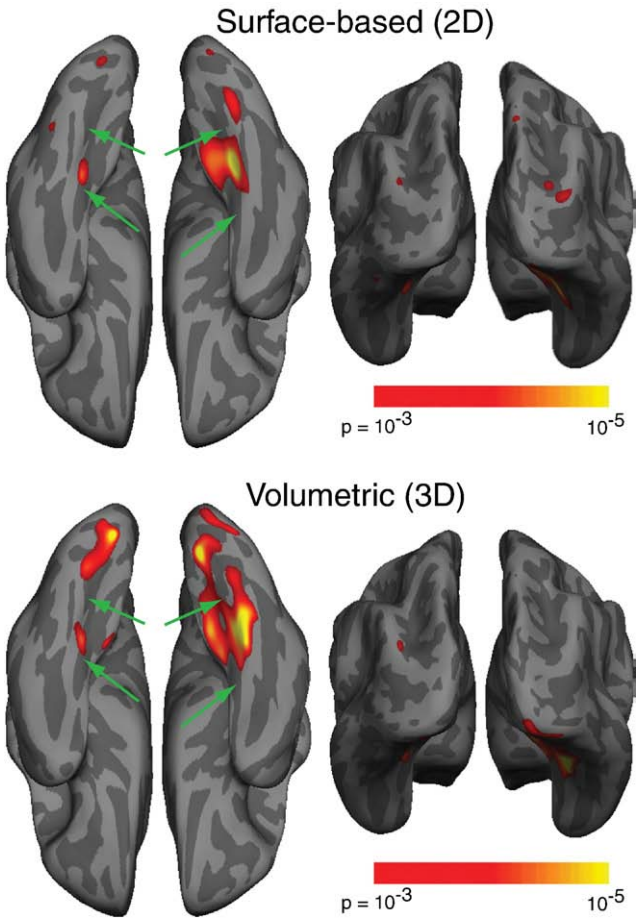

Fig. 3. Decoding object categories. Distributions of *p*-values from a group statistic comparing the decoding accuracy against chance level are shown on the averaged cortical surface. Colored regions indicate cortical areas where the category of the rotating objects can be decoded from local patterns significantly across the group of subjects. Top: Results obtained with the surface-based searchlight method defined on the graymid surface. Bottom: Results obtained with the standard volumetric method projected onto the surface. Note that the scaling of *p*-values is identical between the two figures. Thus same colors indicate the same statistical significance in both subfigures. Green arrows indicate the location of the Collateral Sulcus (CoS). Supplementary Fig. S1 depicts the results on the flattened cortical surface.

important differences. While the results from the surface-based decoding were located in a single region within the CoS, the significant regions of the volumetric decoding spread much further. See also Supplementary Fig. S1, where the results are shown on the flattened cortical surface. Fig. 4 shows the difference in spatial specificity between two techniques. The significance maps are shown on the folded surface (projected to the smoothed white–gray boundary). In addition, the same maps are also illustrated on two orthogonal slices of the anatomical volume (a coronal and a sagittal cut through the CoS). The volumetric sampling caused the activation to spread into the fusiform gyrus giving the impression of two independent peaks. In contrast, the surface-based method located the effects within the collateral sulcus. Note that although the region that reached a significance value below $p = 0.001$ (red area) is considerably smaller in the surface-based searchlight, the most significant parts showed a comparable effect ($p = 10^{-6}$).

Varying cortical depth of the searchlight

It seems intuitive to sample cortex using a surface reconstruction that lies in the middle of the gray matter (graymid), which is between 1 and 4.5 mm thick (Fischl and Dale, 2000). As this thickness is of the same size as the voxels of the functional scans ($3 \times 3 \times 3$ mm) the

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

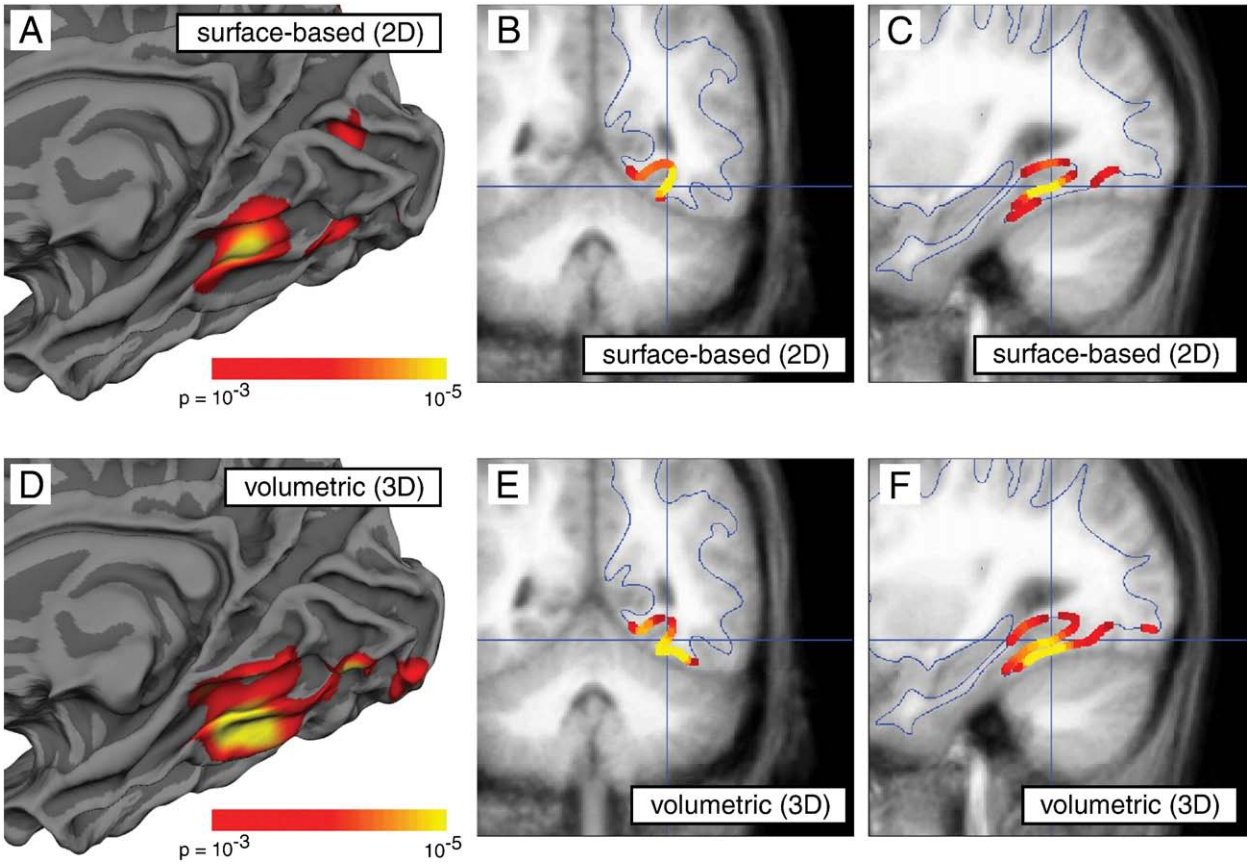

Fig. 4. Comparison of spatial specificity between surface-based and volumetric decoding. Excerpt of the same decoding results as in Fig. 3 rendered on the white matter surface (A, D) and overlaid on coronal (B, E) and sagittal (C, F) anatomical slices through the CoS. While the results obtained with the volumetric method (D, E, F) clearly spread through the fusiform gyrus, the surface-based method (A, B, C) locates the effects only within the collateral sulcus. Note that the surface-based results are based on searchlights defined on the graymid surface. This discrepancy in localization between the two methods is due to the spatial extent of the volumetric searchlight, which reaches across the white matter and samples from within the collateral sulcus even if the seed voxel is not within the sulcus. Note that the peak statistic values are indistinguishable between the two methods.

voxels within a surface searchlight will not completely cover the gray matter. To investigate the effect of sampling the cortical surface at varying depth, we ran a comparison analysis where we created three different types of surface searchlight based on the white–gray (W) boundary, the graymid (G) and the pial (P) surface respectively. The extraction of the surface-based searchlight was identical to the initial analysis (see Methods and materials) but using different surfaces. Finally, in order to sample from different cortical depths simultaneously, combinations of these three surfaces (W–G, G–P and W–G–P) were considered. Fig. 5 illustrates the results and compares them to volumetric decoding. The *p*-maps clearly show that sampling at the white matter (peak significance $p_{peak} = 8.9^{*}10^{-6}$) boundary or at the pial surface ($p_{peak} = 4.8^{*}10^{-6}$) produced results inferior to the graymid ($p_{peak} = 6.6^{*}10^{-7}$). It is also worth noting that including two or three surfaces produced results more similar to the volumetric searchlight. The minimal *p*-values for the combinations are the following (W–G: $p_{peak} = 2.3^{*}10^{-5}$; G–P: $p_{peak} = 1.7^{*}10^{-6}$; W–G–P: $p_{peak} = 3.7^{*}10^{-6}$) and for the volumetric searchlight (VOL: $p_{peak} = 6.4^{*}10^{-7}$). Note that the searchlights using different cortical depth surfaces have considerable overlap (compare Fig. 5, bottom left).

Influence of searchlight size on decoding

In order to investigate the influence of the searchlight size on the decoding results, we analyzed the data for both the surface-based and the volumetric searchlight for four different radii (6, 9, 12 and 15 mm). Fig. 6 summarizes these results. The area of significant decoding increased as the searchlight radius increased in both cases. However, this gradual increase in significant region was much more restricted for the surface-based searchlight. Importantly, although the size of the significant area increased gradually, the shape of the significant area was highly conserved for the surface-based method. For the volumetric searchlight the increase of the radius from 9 to 12 mm dramatically increased the size of the significant area. When the searchlights became large (in particular at $r = 15$ mm) the results obtained by the two methods became more similar. The peak accuracies for the two methods were again comparable and increased as the searchlight radius increased. However, this increase was not strictly monotonic for the surface-based searchlight (see Table 1).

Non-parametric statistical analysis on the group level

We also assessed the statistical significance of decoding accuracies using a non-parametric method based on permutation tests. The results from this analysis are given in Supplementary Fig. S3. The significant regions obtained by this non-parametric method were larger than the ones obtained by a *t*-test with the same threshold. Importantly, the surface-based searchlight resulted in a higher spatial specificity than the volumetric searchlight irrespective of which of the two statistical analyses was used (see Supplementary Fig. S3).

Discussion

The surface-based searchlight technique is a new alternative to 3D-searchlight techniques (Haynes et al., 2007; Kriegeskorte et al., 2006) that extract spherical voxel clusters as a basis for decoding at each position of the brain. We tested the surface-based searchlight

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

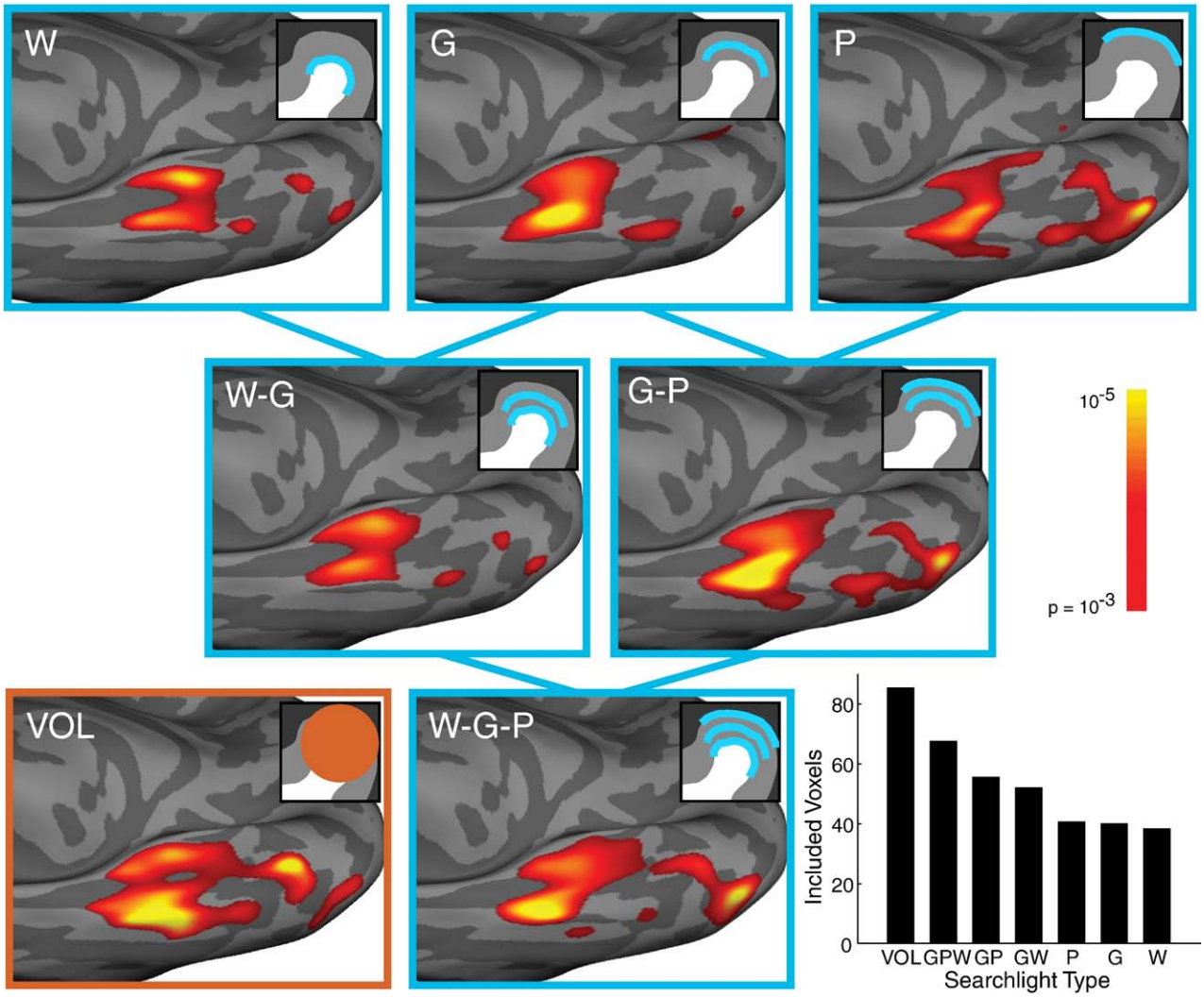

Fig. 5. Surface searchlights at different cortical depths. The graphs illustrate the same region of the cortex as in Fig. 4. Uncorrected *p*-maps of classification accuracy are shown on the inflated cortical surface for surface-based decoding (blue frames) and the volumetric searchlight (red frame). Top row: Results for searchlights extracted from single surfaces: the white–gray boundary (W), the middle between the white–gray boundary and the pial surface (graymid: G) and the pial surface (P). Middle row: Combinations of surface-based searchlight from two different cortical depths (white and graymid: W–G, graymid and pial: G–P). Bottom row: Combination of searchlights from all three cortical surfaces (W–G–P) and volumetric searchlight (VOL, red frame). Insets sketch the surface or combination of surfaces used for the extraction of searchlights (white: white matter, gray: gray matter, black: CSF, blue lines: surface-based searchlights, red disk: volumetric searchlight). All searchlights were first defined on the corresponding surfaces using the same radius. Then searchlights from different cortical depths were combined to yield 3D cylindrical searchlights. All *p*-maps are equally scaled and colored. Only *p*-values $p<0.001$ are shown. The minimum values of *p* (highest significance) for the individual maps are as follows (W: $8.9*10^{-6}$, G: $6.6*10^{-7}$, P: $4.8*10^{-6}$, W–G: $2.3*10^{-5}$, G–P: $1.7*10^{-6}$, W–G–P: $3.7*10^{-6}$, VOL: $6.4*10^{-7}$). The graph on the bottom right illustrates the average number of voxels that was included in the specific combination of searchlights.

using SVM classifiers to decode categories of rotating objects. The results from surface-based decoding were compared to the standard volumetric searchlight on decoding sensitivity and spatial specificity. Although both techniques gave similar results in terms of peak statistical significance, the results from the surface-based methods showed more spatial specificity.

Feature extraction for pattern classification

Broadly speaking, there are three major paradigms in fMRI decoding which differ in their methods of feature extraction.

- First, there are approaches that try to extract the maximum amount of information from the whole brain, treating all of the tens of thousands of voxels as a single pattern vector in a high-dimensional space (Mitchell et al., 2008; Mourao-Miranda et al., 2005). These approaches are particularly efficient for decoding mental states ("brain reading"). Even though the weight distribution of a whole-brain classifier can provide information about the contribution of different brain regions to the classification, the detailed neurophysiological origins of the fMRI-patterning are typically of minor interest.
- Second, some methods look at the information contained in specific regions of interest (Formisano et al., 2008; Haxby et al., 2001; Haynes et al., 2005; Haynes and Rees, 2005; Kamitani and Tong, 2005). This is particularly useful if a clear hypothesis is given about which cortical regions are informative (e.g. the early visual cortex is selected as a ROI for low level visual experiments, as in (Haynes and Rees, 2005; Kamitani and Tong, 2005).
- Third, searchlight methods evaluate local information in a spatially unbiased fashion across the whole brain (Haynes et al., 2007; Kriegeskorte et al., 2006).

The surface-based searchlight presented embodies a significant refinement in searchlight techniques. The assumption underlying the use of the searchlight method as a tool to extract patterns is that

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

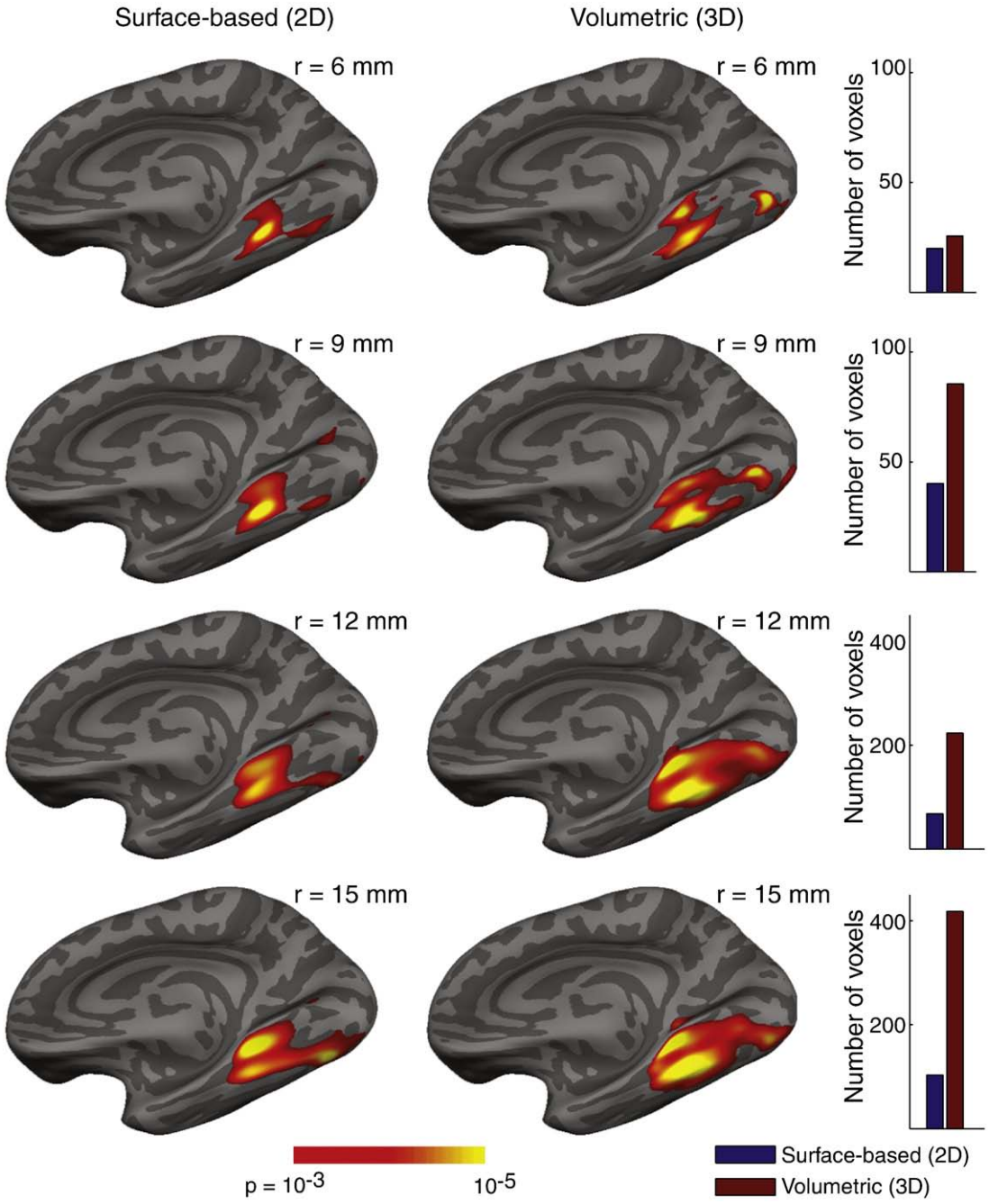

Fig. 6. Effect of searchlight size. Uncorrected *p*-maps of classification accuracy are shown on the inflated cortical surface for surface-based decoding (left) and the volumetric searchlight (right). All surface-based searchlights are extracted from the graymid surface. The radius of the searchlights increases from top to bottom (see insets). All *p*-maps are equally scaled and colored (see color bars at the bottom). Only *p*-values ($p<0.001$) are shown. Bar plots on the right show the average number of voxels included in the searchlights. Note that the scale of the bar plot is different for the two large radii (12 and 15 mm).

information is processed locally within the cortex. However, as we explain above, the surface-based searchlight method respects cortical surface topology and mitigates inappropriate combination of voxels within a single searchlight. Two points which are close in three-dimensional voxel space are not necessarily close in cortical space. By selecting the voxels that are included within a searchlight according to their distance along the cortical surface it is possible to look at patterns within a local cortical region. In fact, we have shown here that the surface-based searchlight method results in a spatial distribution of decoding accuracies that we think is spatially more accurate than the volumetric approach. The reasons why this might be so are discussed in the following.

Decoding of rotating objects in the collateral sulcus

Both surface-based and volumetric searchlight methods showed that the CoS is the region where object identity can best be decoded. However, most studies concerned with object representation report specific responses to objects in the lateral occipital region as well (Malach et al., 2002). Although LOC is often regarded as one of the most important regions for object representation (Malach et al., 1995), decoding the identity of rotating objects was not possible within the LOC, irrespective of the searchlight method used. One possible explanation for this discrepancy is the motion of the

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

Table 1
Peak values of significance from different sizes of searchlights (2D: surface-based searchlight; 3D: volumetric searchlight).

Radius (mm)	6	9	12	15
p-values (2D)	$1.9*10^{-6}$	$1.6*10^{-7}$	$1.5*10^{-5}$	$8.4*10^{-7}$
p-values (3D)	$1.3*10^{-6}$	$6.4*10^{-7}$	$2.3*10^{-7}$	$1.2*10^{-7}$

stimulus. Usually static objects are presented (see Grill-Spector and Malach, 2004, for a review), and therefore the dynamic nature of the stimulation might be the cause for the discrepancy. The area of significant decoding accuracies is less restricted when different rotation speeds are included (see Supplementary Fig. S2).

A note on statistical testing on the group level

We have tested the significance of the decoding accuracy using a *t*-test across the subjects. However, the distribution of decoding accuracies based on cross-validation may not be well approximated by a normal distribution and might vary across subjects. Hence, a *t*-test might not be valid for reporting the group level statistical significance. We therefore conducted a non-parametric statistical analysis (Nichols and Holmes, 2002) in addition to the *t*-tests described above. The main finding of this additional analysis was that significant regions from non-parametric analysis were larger than those obtained by a *t*-test with the same threshold. This suggested that for the current data the *t*-test was conservative. Nevertheless, the spatial specificity of the results from the surface-based method is higher than that of the volumetric method, irrespective of the statistical analysis being used (see Supplementary Fig. S3).

Comparing volumetric and surface-based methods

The comparison between the two searchlight methods has several aspects which speak in favor of one or the other. Volumetric searchlights contain many more voxels than surface-based searchlights with the same radius. In particular, the number of voxels in the volumetric searchlight grows with the cube of the radius, whereas the number of voxels in the surface-based searchlight grows quadratically. In the case of a radius of 9 mm and fMRI voxel side lengths of 3 mm, the former has more than twice as many (85.5 vs. 40.2, mean values across 12 subjects). More voxels might carry more information, but they might also contain more noise. In particular, if meaningful fMRI activations are restricted to the gray matter, then most of the additional voxels in a volumetric searchlight will introduce noise in the classification. This is especially critical in fMRI, where often only few training samples are available to train a classifier in a high-dimensional space. In order to characterize the influence of the searchlight size on decoding, we have compared several searchlight sizes (see Fig. 6). Generally, we observed that the size of the significant regions increased as expected for increased searchlight size. Although it might seem trivial that an increased searchlight increases the sensitivity because it includes more voxels and therefore more information, this is not necessarily the case. As explained above, an increased searchlight size might also include more noisy voxels and, due to the relatively few samples in the training set, it could lead to an overfitting which will effectively decrease the decoding accuracy. Importantly, over all tested sizes the surface-based searchlight conserves the significant regions, which increase only slightly at their border.

As illustrated in Fig. 4 (see also Fig. 2D), a volumetric searchlight can also include gray matter voxels that are far away in cortical distance. This can cause dislocation of patterns and spatial spreading of activation if projected to the surface. In the case of the CoS region, the volumetric searchlight carried information into a cortical region that was, measured in cortical distance, relatively far from the place where the information was actually represented. However, there are also some regions (especially the two occipital peaks), where the volumetric searchlight found information, whereas the surface-based searchlight did not. There are several possible reasons for this disparity:

- First, although we took great care in coregistering the functional to the anatomical images, it is possible that the surfaces extracted from the individual anatomical T1 images were not perfectly aligned with the functional images of individual subjects. The reasons for such a misalignment could be either different distortions in the functional and anatomical MRI images, caused by different sequences, or some errors in the segmentation and surface extraction process in Freesurfer.
- Second, at the top of gyri and in the bank of sulci, the volumetric searchlight covered a much larger region of gray matter than the surface-based searchlight, which could give better classification results.
- Third, although a normalization procedure is needed in both the surface- and volume-based approaches, there are differences between the two approaches that might affect the results. For the surface-based searchlight coregistration of the EPI images to the anatomical T1 is a necessary part of the analysis. The precise selection of voxels that are used for the classification is directly affected by this step of the normalization procedure. Spatial normalization on the cortical surface is only possible if the correspondence of functional voxels to the extracted surface is known. In the volume-based decoding the coregistration is not necessarily needed. In fact, it is common practice to directly normalize the accuracy maps of the subjects to the EPI MNI template for the standard volume-based decoding (Bode and Haynes, 2009; Soon et al., 2008), which is not possible for the surface-based searchlight. However, in order to minimize the differences in the analysis between the two methods, we used the same coregistration procedure for both searchlights here.
- Fourth, the surface mapping *across subjects* based on curvature might not align functional activities perfectly and such misalignments may compromise the statistical significance of group analyses. Such a mismatch between the gyrification of cortex and its functional architecture would affect the volumetric searchlight much less, because it does not rely on any anatomical structure. Compared to the surface-based searchlight, the volumetric searchlight might compensate for between-subject differences in the association between gyrification and function better because it is spatially more extended and potentially samples information from more distal parts of cortex. The volumetric searchlight thus might be able to compensate to some extent for these discrepancies between functional architecture and cortical folding. However, a surface-based searchlight will result in a better spatial sensitivity. This holds only if the alignment of functional and anatomical volumes and the extraction of the cortical surface are precise.

What is the best size of searchlights for decoding?

An important choice that has to be made when applying a searchlight technique is the choice of the size of the searchlight. It is not possible to give a general advice for the "best" searchlight size, because the question is inherently bound to the neural mechanism under investigation, or more specifically, the expected compactness and size of the underlying cortical representation. As shown in Fig. 6 and Table 1, the searchlight size has an effect on both the spatial compactness of significant regions and the peak values of the significance level. The significant area increased in both cases with the searchlight radius, and so too did the peak values of significance level, though not strictly monotonically. As in our study the major

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035

concern is to delineate the fine functional structure of the inferior temporal area in representing dynamic objects, we preferred the relatively small size of 9 mm searchlight, with which a satisfactory peak value of the significance level and a highly compacted informative region were simultaneously achieved. If, however, we were looking for any regions potentially involved in representing objects, then larger searchlight sizes of 15 mm or even more might be preferred, as such a size is less likely to miss weakly or widely distributed information encoded in the cortical regions (e.g. the occipital peaks shown by larger searchlight size and volumetric method).

Which cortical depth is best suited for the extraction of searchlights?

Although all neurons are located within the gray matter, it is not clear whether sampling in the middle between the white matter and pial surface captures all the informative cortical activations. Inaccurate segmentations, partial volume effects, unknown distributions of blood vessels and cortical thickness factors might result in informative voxels lying outside or at the border of the gray matter. Therefore, the selection of a particular surface depth for sampling might influence the classification results. Voxels containing a vertex equidistant from the pial surface and the boundary between white and gray matter contain a higher ratio of cortex than any other set of voxels. Thus, it seems intuitive that decoding from searchlights defined on the graymid surface reconstruction would provide strongest accuracy and localization. The thickness of the cortex varies between 1 and 4.5 mm, and the inter subject standard deviation of the average cortical thickness is 0.5 mm (Fischl and Dale, 2000).

High-field fMRI (with gradient echo EPI imaging) has shown that the strongest BOLD signals in visual cortex are found in the upper layers of cortex and around the pial surface (Jin and Kim, 2008). This finding leads to the prediction that the best signals are measured by sampling gray matter close to the pial surface. On the other hand, the anatomy of cortical vasculature suggests exactly the opposite. Large draining vessels are located at the pial surface (Duvernoy et al., 1981; Weber et al., 2008). These large vessels supply a large region of cortex, and therefore the BOLD signal around them is unlikely to incorporate specific fine-grained patterns. This suggests that the cortex should be sampled as far away as possible from the pial surface, i.e. at the white–gray matter boundary. However, there is evidence that the BOLD signal in large vessels might also contain highly specific information, e.g. about the stimulated eye in visual experiments (Shmuel et al., 2007). Here, we found that the surface decoding results are best for surface-based searchlights extended along the graymid surface. This is a compromise between the two extremes discussed above, although success of this depth may be related to the ratio of cortex covered rather than vascular or BOLD signal considerations. A recent paper has interpreted the BOLD response of a voxel as a complex spatio-temporally filtered version of the local neural firing (Kriegeskorte et al., 2010). Our results suggest that the information contained in this filtered signal is highest when the searchlights sample large parts of the gray matter, as is the case for the searchlights based on the graymid surface. Searchlights centered at the border of the gray matter provide less information. This might be because they sample large amounts of uninformative white matter or include large draining vessels that might carry a less specific signal. It is important to note here that the spatial resolution of 3 mm of the EPI volumes does not allow us to distinguish cortical layers, since the whole cortical sheet extends only over a few millimeters in depth. Combining multiple surfaces (the W–G, G–P and W–G–P conditions above) extends the searchlight into a third dimension forming a cylindrical searchlight. As voxels on the G and P surfaces straddle the grey matter boundary this increases the average number of voxels per searchlight that sample from regions across sulci or small white matter parts in gyri. This should lead to results more similar to the volume based decoding. Indeed, the decoding results for combinations of surface-searchlights were comparable to the volumetric searchlight results.

We have shown that although surface-based and volumetric searchlights produce superficially similar informative distribution on the cortical surface, the surface-based searchlight methods provide a more localized map. There seems to be a difference in spatial specificity between the two methods. However, there might be informative patterns which are missed by the surface-based searchlight method, if the functional activation pattern is not aligned with the cortical surface extraction. In these cases, the volumetric searchlight can still detect the information, because it samples across a larger area and overlaps more with the activation pattern. Some of such misalignments, like a poor EPI-T1 coregistration or EPI distortion, could possibly be reduced by an improved spatial coregistration or distortion correction. In summary, the surface-based searchlight method offers a spatially unbiased approach to sampling information along the cortical surface. It considers patterns that are local in cortical coordinates. It avoids the accumulation of noise sources that lie outside of the gray matter and samples information from voxels that represent 'true' cortical neighbors.

Acknowledgments

This work was funded by the Max Planck Society, the Bernstein Computational Neuroscience Program of the German Federal Ministry of Education and Research (BMBF Grant 01GQ0411), the Excellence Initiative of the German Federal Ministry of Education and Research (DFG Grant GSC86/1-2009), Defence Science and Technology Agency Singapore (DSTA, POD0713897), and National Research Foundation Singapore (NRF, STaR Award). We thank Stefan Bode for his helpful comments on the manuscript and his help in data acquisition and Carsten Allefeld for helpful discussion.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2010.07.035.

References

Bode, S., Haynes, J.D., 2009. Decoding sequential stages of task preparation in the human brain. Neuroimage 45, 606–613.
Carlson, T.A., Schrater, P., He, S., 2003. Patterns of activity in the categorical representations of objects. J. Cogn. Neurosci. 15, 704–717.
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, 261–270.
Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.
Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271.
Duvernoy, H.M., Delon, S., Vannson, J.L., 1981. Cortical blood vessels of the human brain. Brain Res. Bull. 7, 519–579.
Eger, E., Ashburner, J., Haynes, J.D., Dolan, R.J., Rees, G., 2008. fMRI activity patterns in human LOC carry information about object exemplars within category. J. Cogn. Neurosci. 20, 356–370.
Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. USA 97, 11050–11055.
Fischl, B., Sereno, M.I., Dale, A.M., 1999a. Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 9, 195–207.
Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999b. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.
Formisano, E., De Martino, F., Bonte, M., Goebel, R., 2008. "Who" is saying "what"? Brain-based decoding of human voice and speech. Science 322, 970–973.
Golland, P., Fischl, B., 2003. Permutation tests for classification: towards statistical significance in image-based studies. Inf. Process. Med. Imaging 18, 330–341.
Grill-Spector, K., Malach, R., 2004. The human visual cortex. Annu. Rev. Neurosci. 27, 649–677.
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, 2425–2430.

Haynes, J.D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. 8, 686–691.
Haynes, J.D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.
Haynes, J.D., Deichmann, R., Rees, G., 2005. Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. Nature 438, 496–499.
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, 323–328.
Jin, T., Kim, S.G., 2008. Cortical layer-dependent dynamic blood oxygenation, cerebral blood flow and cerebral blood volume responses during visual stimulation. Neuroimage 43, 1–9.
Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8, 679–685.
Knops, A., Thirion, B., Hubbard, E.M., Michel, V., Dehaene, S., 2009. Recruitment of an area involved in eye movements during mental arithmetic. Science 324, 1583–1585.
Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. Proc. Natl. Acad. Sci. USA 103, 3863–3868.
Kriegeskorte, N., Cusack, R., Bandettini, P., 2010. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? Neuroimage 49, 1965–1976.
Levy, I., Hasson, U., Avidan, G., Hendler, T., Malach, R., 2001. Center-periphery organization of human object areas. Nat. Neurosci. 4, 533–539.
Logothetis, N.K., 2008. What we can do and what we cannot do with fMRI. Nature 453, 869–878.
Logothetis, N.K., Wandell, B.A., 2004. Interpreting the BOLD signal. Annu. Rev. Physiol. 66, 735–769.
Macevoy, S.P., Epstein, R.A., 2009. Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. Curr. Biol. 19, 943–947.
Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. USA 92, 8135–8139.
Malach, R., Levy, I., Hasson, U., 2002. The topography of high-order human object areas. Trends Cogn. Sci. 6, 176–184.
Mitchell, T.M., Shinkareva, S.V., Carlson, A., Chang, K.M., Malave, V.L., Mason, R.A., Just, M.A., 2008. Predicting human brain activity associated with the meanings of nouns. Science 320, 1191–1195.
Mourao-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, 980–995.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
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, 424–430.
Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 45, S199–S209.
Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N.K., Ugurbil, K., 2007. Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. Neuroimage 35, 539–552.
Soon, C.S., Brass, M., Heinze, H.J., Haynes, J.D., 2008. Unconscious determinants of free decisions in the human brain. Nat. Neurosci. 11, 543–545.
Weber, B., Keller, A.L., Reichold, J., Logothetis, N.K., 2008. The microvascular system of the striate and extrastriate visual cortex of the macaque. Cereb. Cortex 18, 2318–2330.
Williams, M.A., Baker, C.I., Op de Beeck, H.P., Shim, W.M., Dang, S., Triantafyllou, C., Kanwisher, N., 2008. Feedback of visual object information to foveal retinotopic cortex. Nat. Neurosci. 11, 1439–1445.

Please cite this article as: Chen, Y., et al., Cortical surface-based searchlight decoding, NeuroImage (2010), doi:10.1016/j.neuroimage.2010.07.035