Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

The neuroscientific exploitation of high-resolution functional magnetic resonance imaging

Nikolaus Kriegeskorte and Peter Bandettini

Section on Functional Imaging Methods, Lab of Brain and Cognition, National Institute of Mental Health

Abstract

High-resolution functional magnetic resonance imaging (hi-res fMRI) promises to bridge the gap between the macro- and the microview of brain function afforded by conventional neuroimaging and invasive cell recording, respectively. Hi-res fMRI (nominal voxel sizes $\leq (2\ mm)^3$) is robustly achievable in human studies today using widely available clinical 3-Tesla scanners. However, the neuroscientific exploitation of the greater spatial detail poses three challenges: (1) Fine-scale neuronal activity patterns are inaccurately depicted in the hemodynamic images obtained. (2) Single small voxels yield very noisy measurements. (3) For groups of subjects, the interindividual correspondency mapping is unknown at the fine scale of millimeters. Here we argue that these challenges can be met by abstracting from the regional fine-scale activity patterns themselves and instead asking how well they distinguish the experimental conditions.

Introduction

Neuroimaging has helped define the big picture of primate brain function at the systems level. Its macroview of functional regions activated in a given cognitive process has complemented the microview offered by invasive cell recording. However, with each blob of a neuroimaging map reflecting the averaged activity of hundreds of thousands of neurons over extended periods of time, macro- and microview are separated by many orders of magnitude. The gap between them represents one of the obstacles to forming a coherent overall account of the functional dynamics of the primate brain.

Hi-res fMRI (i.e. nominal voxel sizes $\leq (2\ mm)^3$) promises to help bridge the gap between the macro- and the microview of brain function. A voxel size of $(2\ mm)^3$ is robustly achievable in human studies today on widely available clinical scanners operating at 3 Tesla (T). At high field (>3T) and using parallel imaging (Uğurbil et al. 2003, Prüssmann 2004), hi-res fMRI is invading the submillimeter range.

The neuroscientific interpretation of the greater spatial detail now available poses three challenges: (1) Fine-scale neuronal activity patterns are inaccurately depicted in the fMRI images because of hemodynamic blurring, distortion and displacement. (2) Single small voxels yield noisy measurements. Conventional functional maps therefore show salt-and-pepper patterns that are hard to distinguish from noise. (3) For groups of subjects, the interindividual spatial correspondency mapping is unknown. The Talairach standard space (Talairach & Tournoux 1988) conventionally used for group analysis is too coarse to define correspondency between hi-res fMRI voxels.

Although attempts are underway to improve fMRI measurements in order to alleviate challenges (1) and (2), a complete solution seems unlikely. As for (3), the definition of a high-precision functional correspondency mapping (or, equivalently, a high-precision standard brain space) is not merely a technical challenge. It concerns a fundamental neuroscientific question: To what level of spatial detail does a functional correspondency even exist between different human brains?

If hi-res fMRI is to fulfill its promise, neuroscientists will need to learn how to extract the rich information about neuronal activity, which is present in the data despite the drawbacks of the measurement and the lack of a high-precision standard brain space. Here we argue that the challenges can be met by focusing on the *information* that regional fine-scale fMRI activity patterns convey, rather than on finding activated regions.

Information-based fMRI analysis

The idea of analyzing fMRI data for the information in local activity patterns has recently gained momentum as neuroscientists began to see spatial fMRI patterns as widely distributed codes or macropopulation codes (Haxby et al. 2001, Cox & Savoy 2003, Carlson et al. 2003, Sidtis et al. 2003, Mitchell et al. 2004, Polyn et al. 2005, Kamitani & Tong 2005, Haynes & Rees 2005, LaConte et al. 2005, Mourao-Miranda et al. 2005, O'Toole et al. 2005, Davatzikos et al. 2005, Kriegeskorte & Goebel 2002, Kriegeskorte et al. 2006, for a review see Haynes & Rees 2006).

Most of these studies used multivariate pattern-recognition techniques for *decoding*, i.e. they aimed to identify a perceptual representation or cognitive state on the basis of multivoxel regional fMRI signals. When a perceptual representation can be decoded from the multivoxel response, clearly the brain region studied contains information about the stimulus. In general, information-based analysis requires multivariate techniques, though not necessarily decoding.

Multivariate techniques have been explored early on in the neuroimaging analysis literature (e.g. Worsley et al. 1997). However, the current momentum of the information-based approach derives from a specific neuroscientific idea: the idea of reading out a local macropopulation code. This idea motivates the joint multivariate analysis of contiguous functional regions in single subjects without smoothing of the data.

If neuroscientists are to utilize the spatial fine-scale information present in hi-res fMRI data, they must not smooth the data. Unsmoothed data, however, yield salt-and-pepper statistical maps that are hard to distinguish from noise. In fact they may not reach significance in any single voxel (challenge (2)). Furthermore, they may neither accurately reflect the pattern of neuronal activity (challenge (1)), nor match up between subjects (challenge (3)). In addition, their salt-and-pepper structure makes them impossible to report verbally or

relate to systems-level brain theory. How then can those patterns form the basis of solid empirical neuroscience?

Despite the three challenges, an information-based locally multivariate analysis can yield stable results of neuroscientific interest. Since such an analysis locally combines the signals without averaging them, it increases statistical power. If the local activity patterns differ across conditions, this indicates the presence of information about the experimental condition, even if the spatial structure of the patterns is greatly blurred, distorted and displaced (Figure 2). If such information is present in the same macroscopic region in each subject, a group analysis can be performed at the level of local information, even if the fine-grained activity patterns within the region are – like fingerprints – unique to each subject.

We refer to an analysis as "information-based" if it meets the following two loose criteria: (1) An analysis is information-based if it is aimed at determining whether there is a statistical dependency (i.e. mutual information) between experimental condition and multivoxel spatiotemporal fMRI signal. In many cases a direct estimation of mutual information (e.g. Kraskov et al. 2004) may not be optimal, because the number of time points available is small in relation to the large number of voxels (i.e. multivariate dimensions). Functional MRI typically does not allow experimental events to be repeated more than about 100 times per run and provides no more than a few hundred time points. Restrictive assumptions can help by focusing sensitivity on a neuroscientifically motivated subclass of dependencies. For example, the cited studies all restrict the analysis to information in the spatial structure of the activity patterns, averaging over their temporal structure. Moreover, demonstrating a dependency does not require information-theoretic quantification. A significant test result on any effect statistic indicates a dependency between the experimental condition and the multivariate response. (2) The assumptions should not be too restrictive. For example, the classical approach to brain mapping analysis (e.g. Worsley et al. 1992) restricts the focus to changes in local activation, i.e. it reduces a high-dimensional space of possible effects to a single dimension: the local spatial average. This approach is optimal when activation of whole regions is the sole target of the analysis. We refer to it as activation-based, in contrast to the information-based approaches we are concerned with here, which are sensitive to more general effect classes.

Information undetected by activation-based mapping is often present in neuroimaging data (e.g. Sidtis et al. 2003). If the information resides in the fine-scale pattern of the activity, the spatial average may be similar between conditions, so no effect may be found by conventional methods with the data spatially averaged for region-of-interest analysis or smoothed for statistical mapping (e.g. Haynes et al. 2005).

Challenge 1: ***Neuronal activity patterns are inaccurately depicted***

Current fMRI does not directly measure neuronal activity. It measures the local hemodynamic response that delivers more energy to active neurons. Different fMRI techniques target different aspects of the hemodynamic response (e.g. blood flow, blood volume, blood oxygenation level). The spatial accuracy of the depiction of neuronal activity patterns may thus vary somewhat. However, all techniques share the basic limitation that the spatiotemporal patterns measured are transformed by the hemodynamics. First, the hemodynamic changes measured may be more extended than the site of neuronal activity. In other words, the hemodynamic activation unit may be larger than a single neuron. Second, effects originating in the capillaries, where vessels interface with neurons, are washed out into venules and draining veins. This hemodynamic process distorts, blurs and displaces local effect patterns. At the macroscopic level of low-resolution fMRI these effects may be negligible. In hi-res fMRI they are substantial and cannot be ignored.

Consider the study of Cheng et al. (2001), who used of hi-res fMRI to reveal the pattern of ocular dominance columns in human primary visual cortex (Figure 1). Ocular-dominance domains are regions of V1 responding primarily to stimulation from one eye. The two ocular domains form a zebra-like pattern of stripes, whose width ranges between 0.5 and 1 mm. Cheng et al. performed BOLD fMRI with $0.47 \times 0.47 \times 3$ mm^3 voxels using a gradient-echo multi-shot echoplanar imaging pulse sequence at 4T. The activity patterns elicited by prolonged left- and right-eye stimulation (as compared to no visual stimulation) do not reveal the cortical ocular-dominance pattern (Figure 1a, b). This attests to the inaccurate depiction of the neuronal activity pattern by the hemodynamic signal. (In particular, the two conditions activate largely overlapping sets of voxels, suggesting that voxels do not respond specifically to neuronal activity occurring within their boundaries.)

When the fMRI patterns for left- and right-eye stimulation were subtracted, however, significant differences were found between them. The V1 voxels clearly contained a lot of information about which eye was stimulated. This would not be expected in higher visual regions that do not have separate ocular domains.

Since ocular-dominance domains are nonoverlapping by definition, the authors managed to reveal them by mapping the difference between the left- and the right-eye pattern. The difference map (Figure 1, right panel) might be slightly distorted, but it does reveal the ocular domains, correctly depicting their shape and width.

Differential mapping can accurately depict two neuronal activity patterns if (1) the contrasted conditions are associated with complementary neuronal activity patterns and (2) the artefacts corrupting the hemodynamic single-condition maps cancel out in differential mapping (i.e. they are additive and equal for both conditions).

In general, we cannot make either of these assumptions. Differential mapping will not reveal what two noncomplementary neuronal activity patterns look like. In most studies, therefore, hi-res fMRI patterns will somewhat inaccurately depict neuronal patterns – just like the single-condition patterns obtained by Cheng and colleagues.

Information-based analysis does not require accurate depiction of neuronal activity patterns

Even if the neuronal patterns cannot be depicted accurately, a difference between the distorted fMRI patterns indicates that the two underlying neuronal patterns differ (Figure 2). Information-based analysis targets such differences between activity patterns and reveals whether a given brain region distinguishes the experimental conditions.

Of course, the hemodynamic blurring and distortion could in principle obliterate the information about the experimental

condition contained in a region's neuronal activity. This might be expected if the information is known to be confined to the high-spatial-frequency band of the neuronal activity patterns. However, studies like that of Cheng et al. clearly indicate that neuronal pattern information at a scale as small as cortical columns is present in hi-res fMRI data.

Challenge 2: ***Small voxels yield noisy responses***

Smaller voxels give noisier time series. If hi-res is not needed to resolve the pattern of activity (consider a large homogeneously activated region), it is advisable to image at a lower resolution: Larger voxels will yield lower noise than smaller voxels locally averaged to give the same lower resolution. In studies that target fine-scale spatial patterns of activity, however, small voxels give more detailed information – although they may not singly yield significant effects using conventional amounts of data.

Information-based analysis reduces noise by locally combining signals without averaging

Two recent studies (Kamitani & Tong 2005, Haynes & Rees 2005) beautifully demonstrate how information-based analysis can reveal neuronal pattern information even when single voxels do not show significant effects.

These authors imaged activity patterns in early visual areas while subjects viewed oriented gratings. Under these conditions, single voxels do not necessarily show significant effects of grating orientation.

In both studies the information-based analysis involved using one portion of the data (the training set) to determine a weight for each voxel. The weight is a positive or negative real number and reflects how well the voxel distinguishes the viewed grating orientations. The weighted sum of the voxels then can be shown (using an independent portion of the data: the test set) to reflect the orientation viewed, demonstrating that the region in question (e.g. V1) contains orientation information.

These studies did not use hi-res fMRI, but a more conventional resolution of $3 \times 3 \times 3$ mm^3. Nevertheless they demonstrate how information-based locally multivariate analysis can reveal effects invisible to conventional univariate analysis. This superiority of the information-based approach holds regardless of whether the univariate analysis is applied to each voxel separately or to the overall spatial average of a region.

Challenge 3: ***The fine-scale interindividual correspondency mapping is unknown***

Neuroimaging group analysis is usually performed by projecting data into the Talairach standard space. Each subject's data is first spatially smoothed (by convolution of each volume with a Gaussian of 4-8 mm full width at half maximum), because it is known that matching regions can be off by many millimeters between subjects in Talairach space. This approach is clearly unsuited for hi-res fMRI as it would obliterate the targeted information in fine-scale activity patterns.

Cortex-based intersubject alignment can provide a more precise correspondency mapping than Talairach space (Fischl et al. 1999, Argall et al. 2006). However, the problem at hand concerns a fundamental neuroscientific question: To what level of spatial detail does a functional correspondency even exist between different human brains? This is an empirical question. Correspondency cannot exist at the single-neuron level, because the number of neurons varies across brains. The precision of intersubject correspondency is likely to depend on the brain region studied. A high-precision standard brain space may never be available for group analysis.

Information-based group analysis does not require fine-scale interindividual correspondency

When activity patterns cannot be matched between subjects, fine-scale effects need to be investigated at the single-subject level first. Once multivariate effect measures have been obtained for corresponding regions in all subjects, group analysis can proceed at the level of whole regions characterized by these effect measures. This two-scale approach can be extended to continuous mapping of local information (Kriegeskorte et al. 2006). At the fine spatial scale of millimeters, activity patterns are assumed to be unique to each individual and therefore analyzed separately for each subject. At the coarse spatial scale of centimeters, single-subject statistical maps reflecting local information are combined (using Talairach or a cortex-based common space) to increase statistical power and obtain a group-statistical summary of the individual results.

Conclusion

We argued that hi-res fMRI patterns can form the basis of solid empirical neuroscience even if they do not accurately depict neuronal activity patterns, do not reach significance in single voxels and do not match up between subjects. The three challenges can be met by abstracting from the regional fine-scale activity patterns themselves and instead asking how well they distinguish the experimental conditions. Techniques of information-based, locally multivariate analysis, which have recently gained momentum in the analysis of low-resolution fMRI data, address this question and will be crucial if hi-res fMRI is to fulfill its promise.

References

Argall, B. D., Saad, Z. S. & Beauchamp, M. S. (2006) Simplified intersubject averaging on the cortical surface using SUMA. *Human Brain Mapping* **27**(1), 14-27

Carlson, T. A., Schrater, P. & He, S. (2003) Patterns of activity in the categorical representation of objects. J. Cogn. Neurosci. 15, 704–717.

Cheng, K., Waggoner, R. A. & Tanaka, K. (2001) Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron* **32**, 359–374.

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**, 261-70.

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

Fischl, B., Sereno, M. I., Tootell, R. B. H. & Dale, A. M. (1999) High-resolution inter-subject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* **8**:272–284.

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**, 2425-30.

Haynes, J. D. & Rees G. (2006) Neuroimaging: Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience* **7**, 523-534.

Haynes, J. D. & Rees, G. (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* **8**(5), 686-91.

Kamitani, Y. & Tong, F. (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* **8**(5), 679-85.

Kraskov, A., Stögbauer, H. & Grassberger, P. (2004) Estimating mutual information, *Phys Rev E* **69**, 066138.

Kriegeskorte, N. & Goebel, R. (2002) Information-based multivariate functional brain mapping. In *Proc. of the 8th Int Conf on Func Map of the Human Brain*, June 2-6, 2002, Sendai, Japan. *NeuroImage* 16, No 2, Part 2.

Kriegeskorte, N., Goebel R. & Bandettini, P.A. (2006) Information-based functional brain mapping. Proc Natl Acad Sci U S A. **103**(10):3863-8.

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

Mitchell, T. M., Hutchinson, R., Niculescu, R. S., Pereira, F. & Wang, X. (2004) Learning to decode cognitive states from brain images. Machine Learning 57, 145–175.

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.

O'Toole, A., Jiang, F., Abdi, H. & Haxby, J. V. (2005) Partially distributed representation of objects and faces in ventral temporal cortex. J. Cogn. Neurosci. 17, 580–590.

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

Prüssmann, K. P. (2004) Parallel imaging at high field strength: synergies and joint potential. *Top Magn Reson Imaging* **15**(4), 237-44.

Sidtis, J. J., Strother, S. C. & Rottenberg, D. A. (2003) Predicting performance from functional imaging data: methods matter. Neuroimage 20, 615–624.

Talairach, J. & Tournoux, P. (1988) *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publishers, New York.

Uğurbil, K., Toth, L. & Kim, D. S. (2003) How accurate is magnetic resonance imaging of brain function? Trends Neurosci **26**, 108–14.

Worsley, K. J., Evans, A. C., Marrett, S. & Neelin P. (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab.* **12**(6), 900-18.

Worsley, K. J., Poline, J.-B., Friston., K. J. & Evans, A. C. (1997) Characterizing the response of PET and fMRI data using multivariate linear models. *Neuroimage* **6,** 305–319.

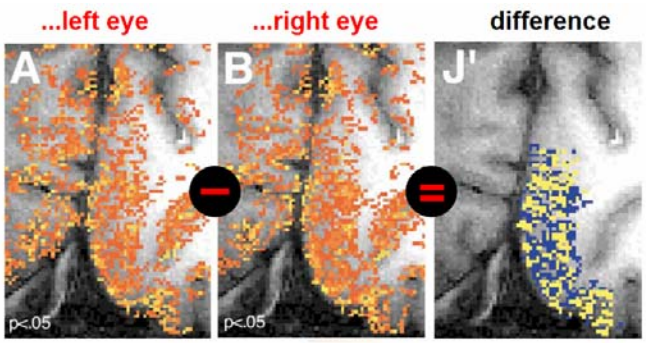

Figure 1: **Hi-res fMRI patterns inaccurately depict neuronal activity (challenge 1), but their difference is informative.** Results of Cheng et al. (2001) show that fMRI patterns elicited in V1 by left- and right-eye stimulation are corrupted by vascular artefacts obscuring the ocular dominance pattern present in the neuronal responses. When the fMRI patterns were subtracted, however, a replicable representation of the ocular-dominance pattern could be obtained. Cheng et al. performed gradient-echo echoplanar imaging at 4T using $(0.47\text{mm})^2 \times$ 3mm voxels.

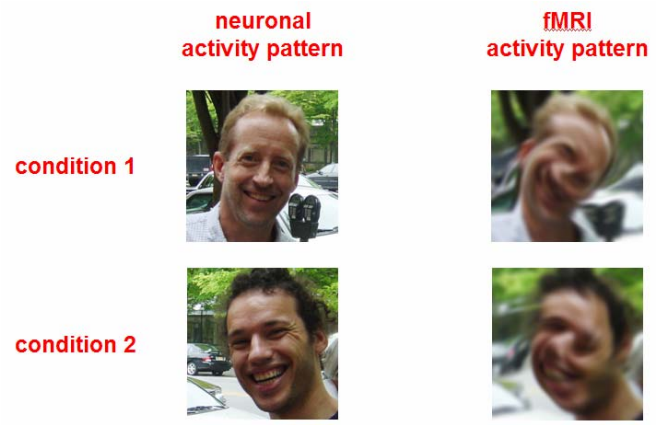

Figure 2: **Despite blurring, distortion and displacement, fMRI activity patterns may distinguish experimental conditions.** In this figure photographs serve as a stand-in for activity patterns. Even if the neuronal activity is corrupted beyond recognition in the fMRI patterns, information distinguishing the experimental conditions may still be present. This would not be the case in a brain region displaying the same neuronal (and thus the same fMRI) activity pattern in both conditions.

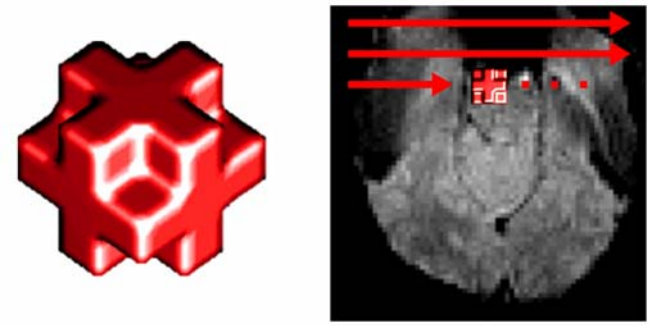

Figure 3: **Spherical multivariate searchlight for information-based functional brain mapping.** A spherical multivariate searchlight, here consisting of 19 voxels, can be used to map a neuroimaging volume for information in local fine-scale activity patterns, rather than for blobs of activation, which are detected by conventional univariate mapping based on smoothed data.