·Review·

Recent developments in multivariate pattern analysis for functional MRI

Zhi Yang[1], Fang Fang[2,3], Xuchu Weng[2]

[1]*Key Laboratory of Behavioral Sciences, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China*
[2]*Center for Cognition and Brain Disorders, Hangzhou Normal University, Hangzhou 310015, China*
[3]*Department of Psychology, Peking University, Beijing 100817, China*

© Shanghai Institutes for Biological Sciences, CAS and Springer-Verlag Berlin Heidelberg 2012

Abstract: Multivariate pattern analysis (MVPA) is a recently-developed approach for functional magnetic resonance imaging (fMRI) data analyses. Compared with the traditional univariate methods, MVPA is more sensitive to subtle changes in multivariate patterns in fMRI data. In this review, we introduce several significant advances in MVPA applications and summarize various combinations of algorithms and parameters in different problem settings. The limitations of MVPA and some critical questions that need to be addressed in future research are also discussed.

Keywords: multivariate analysis; fMRI; pattern recognition; computational biology

1   Introduction

Unlike the traditional fMRI studies that mainly focus on mapping the magnitude of changes in the blood-oxygen-dependent level (BOLD) signal in various brain areas and during different task paradigms, a novel analytical approach to fMRI data, multivariate (or multi-voxel) pattern analysis (MVPA), aims to extract information from subtle voxel-wise activation patterns, and has attracted increasing attention[1-7]. Usually, it is also referred to as classification or decoding, because the activation of multiple voxels from fMRI data is treated as a pattern and this method aims to reveal information encoded in this pattern[8,9].

Specifically, MVPA considers each pattern (referred to as a sample) consisting of *n* voxels as a vector in an *n*-dimensional space. Classifiers are first trained to classify a set of samples (training-set) from different experimental conditions or brain states and then they are tested with another independent set of samples (test-set) to predict which conditions they are associated with. The accuracy of the prediction is referred to as the generalization performance, which indicates how well the classifier performs in extracting the differences between the samples from different conditions. Different algorithms can be used in the classification. Given the fact that the number of samples is usually less than the number of dimensions in fMRI data, certain algorithms featuring this case, such as supporting vector machine (SVM), are widely used.

Due to its multivariate nature, the MVPA approach is sensitive to differences in the voxel activation patterns between different brain states or representational contents[9]. It takes account of the relative activation changes across the voxels in patterns. In this sense, traditional univariate analysis based on the general linear model (GLM) can be viewed as a special case that aims to detect voxels significantly changing (passing a statistical threshold)

in the same direction[8]. Given this constraint, the usual practice in univariate methods is to smooth and average the activation of the voxels, which may potentially cancel out the information encoded among the voxels. MVPA preserves more information in fMRI data by extracting information from patterns of activity across multiple voxels. One striking example is from Kamitani & Tong[10], in which the authors demonstrated the capacity of MVPA to decode eight perceived orientations in early visual areas. Furthermore, they showed that attention strongly biased the orientation preference of the voxels. Such findings are difficult to obtain with univariate voxel-wise analysis, because the size of the voxel in fMRI data is usually much larger than the scale of the orientation-selective columns in early visual cortex, and thus the effects from neuronal populations with different orientation preferences are mixed together and hard to detect in a single voxel.

Unlike conventional GLM-based brain mapping, MVPA has been widely used to accurately predict (decode) behavioral variables encoded in a neuronal system[2]. The term 'MVPA' is therefore interchangeable with 'MVPA decoding' in many circumstances. However, the concept of treating fMRI data as multivariate patterns has the same purpose as GLM, as pointed out by Friston[11]: "showing that one decode activity in the visual cortex to classify (above-chance) a subject's percept is exactly the same as demonstrating significant visual cortex responses to perceptual changes". Indeed, recent MVPA studies primarily aimed to recover multivariate patterns carrying information on behavioral variables, and several sophisticated methods (e.g., cross-condition classification and representational similarity analysis) have been proposed to achieve this end.

Applications of the MVPA have rapidly developed, covering studies of neuronal mechanisms in various domains, such as perception[12-17], learning and memory[18-21], language[22], intention[23], decision-making[5,24-26], emotion[27-30], and mental disorders[31-33]. Instead of serving as a tutorial on applying multivariate classifiers to fMRI data (see Pereira *et al.*[34], an excellent reference for this purpose), this review attempts to give a general introduction to recent promising developments branching off from the original idea of treating fMRI data as multivariate patterns. We therefore classify the methods into two main categories according to their primary aims: methods that aim to accurately decode behavioral variables or identify/reconstruct stimuli (Sections 2 and 3), and those for identifying the brain activity patterns that encode the behavioral variables (Sections 4–6). Their features are illustrated with a few recent representative studies. Furthermore, with the rapid development of MVPA approaches, various combinations of algorithms and parameters have been used. These combinations in different problem settings are also summarized, to provide a clearer view from a technical perspective. Finally, some interesting questions for future research are included.

2   Predicting behavioral variables from fMRI data

Besides its ability to detect spatially distributed effects, MVPA provides a quantitative way to infer the representation of a specific mental process from brain activity, allowing examination of the predictive power of the spatial patterns of brain activity in depth. Recent studies on decoding spatial[35] and episodic[18] memory in the human hippocampus have clearly demonstrated the applications of MVPA and provided promising results. In an interactive virtual-reality navigation task (like a first-person video game) that requires a subject to navigate between four arbitrarily chosen target positions, MVPA was applied to the voxels within the hippocampus and the wider medial temporal lobe (MTL) in a searchlight manner[3] (see 'Section 5 Searchlight MVPA' for details) to search for local voxel activation patterns that carried predictive information about where the subjects were. The significance of the searchlight maps was assessed by nonparametric statistical tests based on permutations of the condition labels of the samples. As a result, the position of a subject within this virtual-reality environment was accurately decoded solely from the pattern of activity in the hippocampus, even when the visual input and the task were held constant. The authors also reported a dissociation of the responses in the

hippocampus from those in the parahippocampal gyrus, suggesting their different roles in navigation[3].

The same group further extended that study by predicting which clip of episodic memory a person is recalling[18]. Compared to the simple virtual-reality environment, human episodic memory is much richer in content. Intriguingly, different episodic memory clips were distinguishable merely from the activity patterns in some hippocampal regions, suggesting that the neuronal traces of episodic memories are stable (and thus predictable) even over many re-activations. These serial studies have important implications for understanding how information is represented within neuronal populations in the human hippocampus. Such findings were not observed using univariate approaches with fMRI data. Technically, the latter study is an excellent example of applying MVPA-searchlight feature selection to an fMRI dataset. Importantly, a three-split regime of the entire dataset is necessary in such applications. That is, one needs to first split the dataset into a training-set and a test-set, and perform the MVPA-searchlight feature selection only with the training-set to avoid circular reasoning issues[36]. To do this, one needs to further split the training-set into a sublevel training-set and test-set.

Besides decoding the stimuli from the fMRI signal, MVPA has also been used to make accurate predictions about the behavioral characteristics of subjects. One typical example is by Dosenbach *et al.*, who predicted individual maturity from the connectivity pattern across multiple brain regions[37]. In this study, supporting vector regression (SVR) was used to predict a continuous maturity index derived from the functional connectivity across 160 brain regions. The "importance" of each connection in the prediction was also reported, reflecting the key features related to brain maturity. Moreover, MVPA has been successfully used to distinguish patients with mental disorders from healthy control subjects[38-42]. One typical study by Zeng *et al.*[41] used the functional connectivity across 116 anatomical regions covering the entire brain to classify major depressive disorder (MDD) patients from healthy control subjects. With sophisticated techniques, the accuracy of identifying MDD patients reached 100%,

suggesting the clinical value of combining resting-state fMRI with MVPA approaches.

3 Identifying/reconstructing new stimuli from brain activity ("mind-reading")

In 2006, Thirion *et al.*[43] originated an attractive approach that developed MVPA towards "mind-reading". In their study, the retinotopy of the visual cortex was used to infer the visual content of real or imagined scenes from the brain activation patterns elicited. They successfully reconstructed a pattern imagined by the subjects. Later, Kay *et al.*[44] developed a more generalized decoding method based on quantitative receptive-field (RF) models that characterize the relationship between visual stimuli and fMRI activity in the early visual areas. Specifically, gray-scale natural images were modeled with linear combinations of a set of elementary images (basis images) containing Gabor wavelet pyramids[45,46]. By recording the activity in the early visual cortex with fMRI when the subject was viewing a large set of natural images, relationships (RF model) between the activation of each voxel and the basis images were computed (by solving linear equations). Combining the estimated RF models from all the voxels and their activation evoked by viewing natural images, the authors identified what the subject saw from a much larger image set. The clever connection of the RF model to the multi-voxel activation pattern allows this approach to "read out" images that have never been viewed by the subject. From a methodological perspective, this approach differs from multivariate pattern classification in that it requires an *a priori* explicit model to connect the spatial patterns with the stimuli, which enables the identification of novel stimuli that are never used in the tuning parameters of the model.

With a similar strategy, Miyawaki *et al.*[47] showed precisely reconstructed images from the early visual cortex, and suggested that if a perceptual state can be expressed by a combination of elemental features, a modular decoder can be trained for each feature with a small number of samples, but their combination can predict numerous states including those that have never been presented.

Meanwhile, Mitchell *et al.*[48] presented an approach to "read out" the words the subject was thinking of. They modeled the meaning of words with a set of intermediate semantic features (basis words), and determined the relationships from a large corpus of texts. After connecting this semantic model to the activation patterns of fMRI data, they were able to predict the words in the subject's mind by linearly combining the fMRI signatures with each of the basis words. Following these studies, more sophisticated models have been proposed to better reconstruct the image[49] or generalize the strategy to decode more complicated contents of the mind[14,50].

4   Cross-condition pattern classification

Stepping forward from the classic MVPA approach that addresses whether the brain representations of different mental processes are separable, cross-condition pattern classification aims to verify whether the information encoded in spatial patterns is consistent across different conditions or tasks. In such analyses, classifiers are trained to distinguish two or more mental states with spatial patterns obtained in one condition (or task), while their generalization performance is tested on spatial patterns observed in another condition (or task). An excellent example demonstrating the power of cross-condition pattern classification is from Meyer *et al.*[51], who demonstrated that the subjective experience of sound, in the absence of auditory stimulation, is associated with content-specific activity in early auditory cortices in humans. In this study, subjects viewed nine sound-implying, but silent, visual clips belonging to three conceptual categories, animals, musical instruments, and objects. With cross-condition pattern classification, the authors showed that the classifiers trained using the spatial patterns for two videos per category could accurately classify a third video into the correct category and that the classifiers trained with the spatial patterns for the silent video clips were able to classify the corresponding sound into animal or object categories. These results provide solid evidence for subjective content-specificity in the primary auditory cortex. More strikingly, using this method, Knops *et al.*[52] demonstrated a successful transfer of mul-

tivariate patterns across completely different experimental conditions by showing that the MVPA classifiers trained to distinguish eye movement direction (left *versus* right) could accurately classify arithmetic operations (subtraction *versus* addition).

In the same vein, we used the cross-condition pattern classification approach to determine whether the spatial patterns in the visual word form area (VWFA) are consistent across different attention levels[17] and different types of character (Chinese and Korean[53]). The spatial patterns in category-selective visual areas such as the fusiform face area (FFA) and the para-hippocampal place area (PPA) were shown to be less consistent at a lower than a higher attention level[54]. To further determine whether the attention level modulates the spatial representation of stimuli in these category-selective regions (FFA and VWFA), we trained linear classifiers to distinguish faces and characters with the spatial patterns in the low-attention condition and tested their performance with the spatial patterns in the high-attention condition. As a result, we found that the classification accuracy was higher than chance-level. More intriguingly, when increasing the contrast-to-noise ratio (CNR) of the spatial patterns in the low-attention condition, we found that the accuracy of the cross-condition classification kept increasing until it achieved the accuracy level when training and testing the classifier both with the spatial patterns under the high-attention condition[17]. This provided innovative evidence for stable category-selective spatial representations in the VWFA. With a slightly different implementation of the cross-condition pattern classification, we also demonstrated that the similarity of the spatial patterns in the VWFA representing Korean and Chinese characters was as high as that between Chinese characters[53]. This means that Chinese and Korean characters may be processed as a single visual category in Chinese-Korean bilinguals.

The above examples demonstrate that cross-condition pattern classification provides a tool for neuroscientists to determine whether the neuronal representations of different stimuli are highly similar. To answer this question in a conventional way, the only method is to compare the over-

all magnitude and the spatial location/extent of the activation clusters with different stimuli. In addition, the CNR manipulation approach we proposed may help to further disentangle the impact of CNR from variation of the multivariate patterns.

5 Searchlight MVPA

To detect brain regions that carry information about stimuli, researchers have developed a searchlight MVPA approach based on spatial filtering of the fMRI data[3,55-57]. The basic idea is to apply MVPA to classify stimuli within the local neighborhood around each voxel on the image, and assign the performance metric of the MVPA to the center voxel. Exploiting its information-mapping nature, a number of studies have applied this approach to a vast range of research topics[15,16,22,25,26,30,58-64]. On the other hand, searchlight MVPA can be considered as a feature-selection method that reflects which voxels (features) are most informative in distinguishing stimuli (see 'Section 7 Summary of MVPA algorithms and parameters'). A recent novel application of searchlight MVPA to classify obsessive-compulsive disorder (OCD) patients and healthy control subjects was demonstrated by Weygandt *et al.*[64]. In this study, the authors first applied whole-brain MVPA at the individual level to classify brain responses evoked by fear-eliciting and neutral pictures. Thus, the classifiers assigned a discrimination value to each voxel. The authors then applied searchlight MVPA to local patterns of discrimination values to classify OCD patients and healthy control subjects, yielding 100% accuracy in the orbitofrontal cortex and the caudate nucleus. Despite concerns about the impact of head motion[65] and errors in image registration, this approach is straightforward in detecting biomarkers between populations.

6 Representational similarity analysis

To give MVPA a boarder scope, Kriegeskorte *et al.*[66] proposed an approach aiming to bridge the three branches of systems neuroscience: behavioral experimentation, brain-activity experimentation, and computational modeling, which is termed representation

similarity analysis (RSA). The core idea of this approach is generalizing the concept of voxel activation pattern to a more abstract similarity pattern that can be measured among different conditions and stimuli. Specifically, there are two levels of pattern analysis in this approach. At the first level, a representational dissimilarity matrix (RDM) is constructed to reveal the similarity (or dissimilarity) of the activity patterns evoked by different stimuli or predicted by computational models. At the second level, the relationships between the RDMs from multi-channel measures are tested with nonparametric statistics. The authors showed the usefulness of RSA by quantitatively relating representations of visual objects in the early visual cortex and the FFA as measured with fMRI to computational models spanning a wide range of complexity.

7 Summary of MVPA algorithms and parameters

With the rapid development of MVPA, various techniques in the machine-learning field have been introduced into the analysis of fMRI data, yielding various combinations of algorithms and parameters used at stages of different analyses. In Table 1, we list different combinations of 'features', 'feature selection methods', 'classifiers', 'performance metrics', 'validation schemes', and 'significance tests' reported in the most recent studies using MVPA (most published in 2011 and 2012). The term 'features' stands for the metrics that compose the patterns in the analysis. The most common features used in MVPA are either raw data volumes or regression coefficients obtained from canonical GLM analysis. The former is usually from studies using block or slow event-related studies[17,60,64,67], while the latter choice can be found in fast event-related studies[21,62,68,69]. Another alternative feature is coefficients from GLM using impulse response (IRF) basis functions[58,59]. In MVPA studies aiming to classify subjects into different groups, the choice of features is much wider, including but not limited to raw fMRI data[40], grey matter density[70], volumetric brain morphometry[71], and functional connectivity measures[37,41]. The 'feature selection methods' are used to reduce the number of variables (dimensionality)

Table 1. Summary of algorithms and parameters used in recent MVPA applications

	Examples	Features	Feature selection methods	Classifiers	Performance metrics	Validation schemes	Significance tests
Classification of stimuli	[64]	Raw data volumes in block	N/A*	L-SVC	Classification accuracy	Split-half	Permutation test
	[67]	Averaged raw data	Functional localizer	L-SVC	Classification accuracy	Leave-one-out cross-validation	Group-level *t*-test
	[25, 58]	IRF coefficients	Searchlight	L-SVC	Classification accuracy	Leave-one-out cross-validation	Group-level *t*-test
	[60]	Averaged raw data	Searchlight	L-SVC	Classification accuracy	Leave-one-out cross-validation	Permutation test
	[15, 16, 22, 30, 59, 61, 72]	Canonical GLM coefficients	Searchlight	L-SVC	Classification accuracy	Leave-one-out cross-validation	Group-level *t*-test
	[68]	Canonical GLM coefficients	Functional localizer	L-SVC	Classification accuracy	Leave-one-run-out cross-validation	Group-level *t*-test
	[62]	Canonical GLM coefficients	Searchlight	NL-SVR	Classification accuracy	Leave-one-run-out cross-validation	Group-level *t*-test
	[21]	Canonical GLM coefficients	Searchlight	L-SVR	Correlation between predicted and actual behavior metrics	Leave-one-run-out cross-validation	Group-level *t*-test
Classification of subjects	[64]	Discrimination values from stimuli classifications	Searchlight	L-SVC	Classification accuracy	Leave-two (a pair)-out cross-validation	Binomial distribution; permutation test
	[73]	Volumetric brain morphometry	Nonlinear multivariate RGS; PCA; MRMR	NL-SVR	Correlation coefficients	Nested cross-validation	Permutation test
	[41]	Functional connectivity across 116 brain regions	Kendall rank correlation	L-SVC	Classification accuracy	Leave-one-out cross-validation	Permutation test
	[37]	Functional connectivity across 160 brain regions	Leave-one-out SVC	L-SVR	Correlation coefficients	Nested cross-validation	Permutation test
	[40]	Averaged raw data; canonical GLM coefficients	N/A*	L-SVC	Classification accuracy	Leave-one-block-out cross-validation	Binomial distribution
	[70]	Gray matter density	N/A*	SMLR	Classification accuracy	Leave-one-out cross-validation	Permutation test
Cross-task classification	[17]	Raw data volumes in block	Functional localizer	L-BP	Classification accuracy	Leave-one-block-out cross-validation	Group-level *t*-test
	[53]	Canonical GLM coefficients	Functional localizer	Pearson's correlation	Correlation coefficients	N/A	Group-level *t*-test
	[63]	Canonical GLM coefficients	Searchlight	L-SVC	Classification accuracy	Leave-one-run-out cross-validation	Group-level *t*-test
	[15]	Canonical GLM coefficients	Searchlight	Pearson's correlation	Correlation coefficients	N/A	Group-level *t*-test
	[69]	Canonical GLM coefficients	Functional localizer	L-SVC	Classification accuracy	Leave-one-run-out cross-validation	Group-level *t*-test

Note: *Anatomical brain masks were used to exclude non-brain voxels. GLM, general linear model; L-BP, linear back-propagation neural network; L-SVC, linear supporting vector classifier; L-SVR, linear supporting vector regression; MRMR, maximal relevance-minimal redundancy feature selection; NL-SVR, nonlinear supporting vector regression; RGS, nearest-neighbor-based filter algorithm; SMLR, sparse multinomial logistic regression classifier.

to avoid over-fitting the data with limited training samples. Most studies either used separate functional localizers to define features or directly applied searchlight MVPA to reveal the features. Much more sophisticated techniques from machine-learning theories have been applied in studies aiming to classify subjects (see Table 1 for details). For 'classifiers', the mainstream is the SVM, since it fits well to the data with high dimensionality and a small number of samples. Besides the supporting vector classifier (SVC), some studies used supporting vector regression (SVR) to predict continuous scales in behavioral performance and demographic measures. Instead of the SVC that decides among a finite number of classes, SVR can be a "predictor" of daily behavior that may contain up-to-infinite possibilities. A few studies simply used correlation coefficients to measure the similarity between patterns[15,53]. The advantage of this method is that the relationships between patterns from each trial are clearly revealed. Therefore, this is more suitable for investigating the fine-grained functional organization of very localized brain regions. The 'validation schemes' are used to assess the performance of the trained classifier. To train the classifiers with maximal samples, most studies chose to apply the 'leave-one-out cross-validation' scheme, which uses $n-1$ of the total n samples to train the classifier and test it with the one sample left. Depending on the features used in the studies, 'leave-one-run-out' and 'leave-one-block-out' schemes have also been used. To evaluate the significance of the MVPA accuracy at the group-level, both the *t*-test and non-parametrical permutation test were applied in previous studies, but there has not been a clear conclusion as to which test is optimal for this purpose.

8   Limitations and questions for future research

MVPA has been deemed sensitive in detecting information encoded in multivariate patterns of brain activity. However, some recent studies have urged caution when interpreting MVPA results. For example, a study by Kamitani and Tong[72] showed that MVPA is able to decode orientation performance in the early visual cortex even from neuronal structures smaller than the voxel size in fMRI images. Conversely, a recent study showed evidence for a topographic map of orientation-preference in human V1 at a much coarser scale, suggesting that the successful MVPA decoding of orientation still depends on the information revealed in fMRI voxels[73]. Another common flaw is the "reverse inference" that reasons backward from patterns of activation to infer the engagement of specific mental processes[74,75]. For example, highly sensitive multivariate classifiers may be able to distinguish face and house stimuli based on activation patterns in the primary visual cortex (V1), even though this region is not known to have category-preference. With *a priori* knowledge about V1, researchers can attribute the results to the subtle differences in the basic physical properties in the stimuli.

However, this inference is risky when *a priori* knowledge is unavailable[76]. From a more substantial perspective, Logothetis[77] stated that the presence of "voxels selective to two different stimuli attributes could be potentially detected by modern classifiers, yet the existence of two types of patterns does not necessarily imply the existence of two different types of neural populations". Thereby, an important question for future methodology research is how to identify patterns that not only are informative but also are really used to process the stimuli.

Finally, since different classifiers use different means of extracting information from data, an interesting question is how different types of classifiers affect MVPA performance[8]. For example, different classifiers encode the feature covariance in different ways. Linear discrimination analysis and quadratic discriminant analysis explicitly encode the covariance of the features, while the logistic regression, SVM, and random forest classifiers implicitly encode the covariance, and the Gaussian naïve Bayesian classifier does not encode the covariance. It would be interesting to see how this difference affects their performance in different problem settings. This is not merely a technical question, because performance differences between classifiers with different learning strategies may have implications for our understanding of the organization of neuronal encoding.

In conclusion, we have summarized both the applica-

tion and technical aspects of recent advances in MVPA methodologies. Despite some unsolved problems and controversies, the development of MVPA has allowed us to better exploit the information in fMRI signals. The methodology is still developing rapidly, and we believe it will provide further insights into the understanding of the human brain.

References:

[1]   Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 2001, 293: 2425–2430.

[2]   Haynes JD, Rees G. Decoding mental states from brain activity in humans. Nat Rev Neurosci 2006, 7: 523–534.

[3]   Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci U S A 2006, 103: 3863–3868.

[4]   O'Toole AJ, Jiang F, Abdi H, Penard N, Dunlop JP, Parent MA. Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. J Cogn Neurosci 2007, 19: 1735–1752.

[5]   Haynes JD. Decoding and predicting intentions. Ann N Y Acad Sci 2011, 1224: 9–21.

[6]   Kriegeskorte N. Pattern-information analysis: from stimulus decoding to computational-model testing. Neuroimage 2011, 56: 411–421.

[7]   Cox DD, Savoy RL. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 2003, 19: 261–270.

[8]   Bandettini PA. What's new in neuroimaging methods? Ann N Y Acad Sci 2009, 1156: 260–293.

[9]   Mur M, Bandettini PA, Kriegeskorte N. Revealing representational content with pattern-information fMRI--an introductory guide. Soc Cogn Affect Neurosci 2009, 4: 101–109.

[10] Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. Nat Neurosci 2005, 8: 679–685.

[11] Friston KJ. Modalities, modes, and models in functional neuroimaging. Science 2009, 326: 399–403.

[12] Kravitz DJ, Kriegeskorte N, Baker CI. High-level visual object representations are constrained by position. Cereb Cortex 2010, 20: 2916–2925.

[13] Mur M, Ruff DA, Bodurka J, Bandettini PA, Kriegeskorte N. Face-identity change activation outside the face system: "release from adaptation" may not always indicate neuronal selectivity. Cereb Cortex 2010, 20: 2027–2042.

[14] Chang KM, Mitchell T, Just MA. Quantitative modeling of the neural representation of objects: how semantic feature norms can account for fMRI activation. Neuroimage 2011, 56: 716–727.

[15] Corradi-Dell'Acqua C, Hofstetter C, Vuilleumier P. Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. J Neurosci 2011, 31: 17996–18006.

[16] Alink A, Euler F, Kriegeskorte N, Singer W, Kohler A. Auditory motion direction encoding in auditory cortex and high-level visual cortex. Hum Brain Mapp 2012, 33: 969–978.

[17] Xu G, Jiang Y, Ma L, Yang Z, Weng X. Similar spatial patterns of neural coding of category selectivity in FFA and VWFA under different attention conditions. Neuropsychologia 2012, 50: 862–868.

[18] Chadwick MJ, Hassabis D, Weiskopf N, Maguire EA. Decoding individual episodic memory traces in the human hippocampus. Curr Biol 2010, 20: 544–547.

[19] Mayhew SD, Li S, Storrar JK, Tsvetanov KA, Kourtzi Z. Learning shapes the representation of visual categories in the aging human brain. J Cogn Neurosci 2010, 22: 2899–2912.

[20] Schultz J. Brain imaging: decoding your memories. Curr Biol 2010, 20: R269–271.

[21] Kahnt T, Heinzle J, Park SQ, Haynes JD. Decoding the formation of reward predictions across learning. J Neurosci 2011, 31: 14624–14630.

[22] Herrmann B, Obleser J, Kalberlah C, Haynes JD, Friederici AD. Dissociable neural imprints of perception and grammar in auditory functional imaging. Hum Brain Mapp 2012, 33: 584–595.

[23] Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE. Reading hidden intentions in the human brain. Curr Biol 2007, 17: 323–328.

[24] Soon CS, Brass M, Heinze HJ, Haynes JD. Unconscious determinants of free decisions in the human brain. Nat Neurosci 2008, 11: 543–545.

[25] Bode S, He AH, Soon CS, Trampel R, Turner R, Haynes JD. Tracking the unconscious generation of free decisions using ultra-high field fMRI. PLoS One 2011, 6: e21612.

[26] Kahnt T, Heinzle J, Park SQ, Haynes JD. Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. Neuroimage 2011, 56: 709–715.

[27] Ethofer T, Van De Ville D, Scherer K, Vuilleumier P. Decoding of emotional information in voice-sensitive cortices. Curr Biol 2009, 19: 1028–1033.

[28] Lin D, Boyle MP, Dollar P, Lee H, Lein ES, Perona P, *et al.* Functional identification of an aggression locus in the mouse hypothalamus. Nature 2011, 470: 221–226.

[29] Gomez A, Rothkirch M, Kaul C, Weygandt M, Haynes JD, Rees G, *et al.* Emotion modulates the effects of endogenous attention on retinotopic visual processing. Neuroimage 2011, 57: 1542–1551.

[30] Kotz SA, Kalberlah C, Bahlmann J, Friederici AD, Haynes JD. Predicting vocal emotion expressions from the human brain. Hum Brain Mapp 2012. doi: 10.1002/hbm.22041.

[31] Engels AS, Heller W, Spielberg JM, Warren SL, Sutton BP, Banich MT, *et al.* Co-occurring anxiety influences patterns of brain activity in depression. Cogn Affect Behav Neurosci 2010, 10: 141–156.

[32] Hamilton JP, Chen G, Thomason ME, Schwartz ME, Gotlib IH. Investigating neural primacy in major depressive disorder: multivariate Granger causality analysis of resting-state fMRI time-series data. Mol Psychiatry 2011, 16: 763–772.

[33] Hoeft F, McCandliss BD, Black JM, Gantman A, Zakerani N, Hulme C, *et al.* Neural systems predicting long-term outcome in dyslexia. Proc Natl Acad Sci U S A 2011, 108: 361–366.

[34] Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 2009, 45: S199–209.

[35] Hassabis D, Chu C, Rees G, Weiskopf N, Molyneux PD, Maguire EA. Decoding neuronal ensembles in the human hippocampus. Curr Biol 2009, 19: 546–554.

[36] Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 2009, 12: 535–540.

[37] Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, *et al.* Prediction of individual brain maturity using fMRI. Science 2010, 329: 1358–1361.

[38] Craddock RC, Holtzheimer PE 3rd, Hu XP, Mayberg HS. Disease state prediction from resting state functional connectivity. Magn Reson Med 2009, 62: 1619–1628.

[39] Shen H, Wang L, Liu Y, Hu D. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. Neuroimage 2010, 49: 3110–3121.

[40] Fu CH, Mourao-Miranda J, Costafreda SG, Khanna A, Marquand AF, Williams SC, *et al.* Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. Biol Psychiatry 2008, 63: 656–662.

[41] Zeng LL, Shen H, Liu L, Wang LB, Li BJ, Fang P, *et al.* Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. Brain 2012, 135: 1498–1507.

[42] Oquendo MA, Baca-Garcia E, Artes-Rodriguez A, Perez-Cruz F, Galfalvy HC, Blasco-Fontecilla H, *et al.* Machine learning and data mining: strategies for hypothesis generation. Mol Psychiatry 2012 doi: 10.1038/mp.2011.173.

[43] Thirion B, Duchesnay E, Hubbard E, Dubois J, Poline JB, Lebihan D, *et al.* Inverse retinotopy: inferring the visual content of images from brain activation patterns. Neuroimage 2006, 33: 1104–1116.

[44] Kay KN, Naselaris T, Prenger RJ, Gallant JL. Identifying natural images from human brain activity. Nature 2008, 452: 352–355.

[45] Daugman JG. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am A 1985, 2: 1160–1169.

[46] Jones JP, Palmer LA. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J Neurophysiol 1987, 58: 1233–1258.

[47] Miyawaki Y, Uchida H, Yamashita O, Sato MA, Morito Y, Tanabe HC, *et al.* Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron 2008, 60: 915–929.

[48] Mitchell TM, Shinkareva SV, Carlson A, Chang KM, Malave VL, Mason RA, *et al.* Predicting human brain activity associated with the meanings of nouns. Science 2008, 320: 1191–1195.

[49] Naselaris T, Prenger RJ, Kay KN, Oliver M, Gallant JL. Bayesian reconstruction of natural images from human brain activity. Neuron 2009, 63: 902–915.

[50] Pereira F, Detre G, Botvinick M. Generating text from functional brain images. Front Hum Neurosci 2011, 5: 72.

[51] Meyer K, Kaplan JT, Essex R, Webber C, Damasio H, Damasio A. Predicting visual stimuli on the basis of activity in auditory cortices. Nat Neurosci 2010, 13: 667–668.

[52] Knops A, Thirion B, Hubbard EM, Michel V, Dehaene S. Recruitment of an area involved in eye movements during mental arithmetic. Science 2009, 324: 1583–1585.

[53] Bai J, Shi J, Jiang Y, He S, Weng X. Chinese and Korean characters engage the same visual word form area in proficient early Chinese-Korean bilinguals. PLoS one 2011, 6: e22765.

[54] Reddy L, Kanwisher N. Category selectivity in the ventral visual pathway confers robustness to clutter and diverted attention. Curr Biol 2007, 17: 2067–2072.

[55] Kriegeskorte N, Bandettini P. Analyzing for information, not activation, to exploit high-resolution fMRI. Neuroimage 2007, 38: 649–662.

[56] Kriegeskorte N, Bandettini P. Combining the tools: activation- and information-based fMRI analysis. Neuroimage 2007, 38: 666–668.

[57] Friman O, Borga M, Lundberg P, Knutsson H. Detection of neural activity in fMRI using maximum correlation modeling. Neuroimage 2002, 15: 386–395.

[58] Momennejad I, Haynes JD. Human anterior prefrontal cortex encodes the 'what' and 'when' of future intentions. Neuroimage 2012, 61: 139–148.

[59] Bode S, Bogler C, Soon CS, Haynes JD. The neural encoding of guesses in the human brain. Neuroimage 2012, 59: 1924–1931.

[60] Carlin JD, Rowe JB, Kriegeskorte N, Thompson R, Calder AJ. Di-

rection-sensitive codes for observed head turns in human superior temporal sulcus. Cereb Cortex 2012, 22: 735-744.

[61] Gilbert SJ. Decoding the content of delayed intentions. J Neurosci 2011, 31: 2888–2894.

[62] Bogler C, Bode S, Haynes JD. Decoding successive computational stages of saliency processing. Curr Biol 2011, 21: 1667–1671.

[63] Reverberi C, Gorgen K, Haynes JD. Compositionality of rule representations in human prefrontal cortex. Cereb Cortex 2012, 22: 1237–1246.

[64] Weygandt M, Blecker CR, Schafer A, Hackmack K, Haynes JD, Vaitl D, *et al.* fMRI pattern recognition in obsessive-compulsive disorder. Neuroimage 2012, 60: 1186–1193.

[65] Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 2012, 59: 2142–2154.

[66] Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci 2008, 2: 4.

[67] Liu T, Hospadaruk L, Zhu DC, Gardner JL. Feature-specific attentional priority signals in human cortex. J Neurosci 2011, 31: 4484–4495.

[68] Cichy RM, Heinzle J, Haynes JD. Imagery and perception share cortical representations of content and location. Cereb Cortex 2012, 22: 372–380.

[69] Cichy RM, Sterzer P, Heinzle J, Elliott LT, Ramirez F, Haynes JD. Probing principles of large-scale object representation: Category

preference and location encoding. Hum Brain Mapp 2012. doi: 10.1002/hbm.22020.

[70] Sun D, van Erp TG, Thompson PM, Bearden CE, Daley M, Kushan L, *et al.* Elucidating a magnetic resonance imaging-based neuroanatomic biomarker for psychosis: classification analysis using probabilistic brain atlas and machine learning algorithms. Biol Psychiatry 2009, 66: 1055–1060.

[71] Koutsouleris N, Meisenzahl EM, Davatzikos C, Bottlender R, Frodl T, Scheuerecker J, *et al.* Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch Gen Psychiatry 2009, 66: 700–712.

[72] Kamitani Y, Tong F. Decoding seen and attended motion directions from activity in the human visual cortex. Curr Biol 2006, 16: 1096–1102.

[73] Freeman J, Brouwer GJ, Heeger DJ, Merriam EP. Orientation decoding depends on maps, not columns. J Neurosci 2011, 31: 4792–4804.

[74] Poldrack RA. Can cognitive processes be inferred from neuroimaging data? Trends in cognitive sciences 2006, 10: 59–63.

[75] Poldrack RA. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. Neuron 2011, 72: 692–697.

[76] Bartels A, Logothetis NK, Moutoussis K. fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci 2008, 31: 444–453.

[77] Logothetis NK. What we can do and what we cannot do with fMRI. Nature 2008, 453: 869–878.