



Computational Neuroscience
Research article

Bayesian reconstruction of multiscale local contrast images from brain activity



Sutao Song^{a,b,c}, Xinyue Ma^b, Yu Zhan^b, Zhichao Zhan^b, Li Yao^{b,c}, Jiakai Zhang^{b,*}

^a School of Education and Psychology, Jinan University, Shandong, China

^b School of Information Science and Technology, Beijing Normal University, Beijing, China

^c State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

HIGHLIGHTS

- A new contrast decoding method based on ICA and Naive Bayesian methods was proposed.
- The modified method can decode contrast efficiently.
- The reconstruction accuracy of visual stimuli is far above the level of chance.

ARTICLE INFO

Article history:

Received 3 May 2013

Received in revised form 22 August 2013

Accepted 23 August 2013

Keywords:

Multi-scale local image decoder

Image reconstruction

Naive Bayesian

ICA

fMRI

ABSTRACT

Background: Recent advances in functional magnetic resonance imaging (fMRI) techniques make it possible to reconstruct contrast-defined visual images from brain activity. In this manner, the stimulus images are represented as the weighted sum of a set of element images with different scales. The contrast weight of local images were decoded using fMRI activity recorded when the subject was viewing the stimulus images. Multivariate methods, such as the sparse multinomial logistic regression model (SMLR), have been proven effective for learning the mapping between fMRI patterns of primary visual cortex voxels and contrast of stimulus images. However, the SMLR method is highly time-consuming in practical application.

New method: The Naive Bayesian classifier based on independent component analysis (NB-ICA) is proposed to efficiently decode the contrast of multi-scale local images. First, temporal independent components of fMRI data which were treated as new features for NB classifier were acquired by ICA decomposition. Second, the contrast for each local element image was computed based on NB estimation theory.

Results: NB-ICA method can be used to reconstruct novel visual images. The average spatial correlation between the represented and reconstructed images was 0.41 ± 0.13 ($p < 0.001$).

Comparison with existing method(s): At the expense of reconstruction accuracy, NB-ICA is more efficient than SMLR which reduces the computation time from hours to seconds.

Conclusions: A new method, termed NB-ICA, is proposed and can efficiently reconstruct contrast-defined visual images from fMRI data. This study provides theoretical support for brain-computer interface research and also provides ideas for the study of real-time fMRI data.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Functional magnetic resonance imaging (fMRI) provides a convenient tool for measuring blood-oxygen-level-dependent (BOLD) signals in the brain, a reflection of brain activity (Ogawa et al., 1992). Recently, some multivariate computational methods, as opposed to

univariate analysis methods, have been successfully applied to fMRI data to decipher visual experiences (Kay et al., 2008; Mitchell et al., 2008; Miyawaki et al., 2008; Naselaris et al., 2009; Nishimoto et al., 2011; Thirion et al., 2006). In general, all visual decoding studies establish systematic mapping between the visual stimuli and brain activity. However, these studies differ in the output information gathered from the fMRI data, that is, whether they classify, identify or reconstruct the stimulus images (Kay and Gallant, 2009).

In classification, multivariate pattern recognition methods, such as Fisher Linear Discriminant Analysis (Fisher LDA) and support vector machines (SVM), utilize the entire pattern of activity

* Corresponding author at: Xin Jie Kou Wai Da Jie19#, Beijing 100875, China. Tel.: +86 10 58807856; fax: +86 10 58800056.

E-mail address: jiakai.zhang@bnu.edu.cn (J. Zhang).

observed across multiple voxels to predict which stimulus or task condition the subject was experiencing (Carlson et al., 2003; Cox and Savoy, 2003; Davatzikos et al., 2005; Haxby et al., 2001; Haynes and Rees, 2005a,b; Kamitani and Tong, 2006, 2005; O'toole et al., 2005; Polyn et al., 2005; Spiridon and Kanwisher, 2002). In identification, a computation model that captures the exact correspondence between visual stimuli and brain activity is used to identify which image the subject was viewing out of a set of candidate images, based on measurements of brain activity (Kay et al., 2008; Mitchell et al., 2008). In an early study conducted by Thirion et al. (2006), a receptive-field model was fit separately to each voxel in the early visual areas to predict visual stimuli-elicited voxel activation. This receptive-field model is an example of a typical encoding model, which utilizes a Gabor wavelet pyramid to capture the relationship between fMRI activity and the characteristics of visual stimuli, such as space, orientation and spatial frequency. Then, the brain activity in early visual areas for each potential image was predicted using the receptive-field model. The stimulus decoding process (image identification) was accomplished by selecting closely matched images whose predicted activity was most similar to the measured brain activity recorded during image viewing (Kay et al., 2008).

Recently, researchers have been able to reconstruct visual images composed of flickering checkerboard patterns, more complicated actual natural images, or even dynamic movies, rather than simply choosing the stimulus materials from a known set (Naselaris et al., 2009; Nishimoto et al., 2011). Unlike classification, reconstruction relies heavily on the decoding model, which reproduces the visual stimulus from brain activity (Kay et al., 2008; Thirion et al., 2006). Kay et al. further advanced their fMRI decoding work beyond identification to reconstruction of images and movies by combining encoding models with prior stimuli. Given a fMRI response pattern to a stimulus, the reconstruction model first lists the candidate stimuli by sorting the posteriori in descending order, then the single image can be reconstructed using the maximum a posteriori (MAP) algorithm. Alternatively, the generation of a post-normalized movie clip is yielded by averaging clips near the peak of the posteriori, a method termed averaged high posterior (AHP) reconstruction (Naselaris et al., 2009; Nishimoto et al., 2011).

Another image reconstruction study was conducted by Miyawaki et al., which established a multi-scale local image decoder to directly model the relationship between a contrast-defined image stimulus and fMRI activity at the specific time when the image was presented to the subject (Miyawaki et al., 2008) (Fig. 1). In this reconstruction study, contrast-defined images were modeled by a linear combination of local image elements, as follows:

$$\hat{I}(z|\mathbf{r}) = \sum_{i=1}^M \lambda_i \psi_i(\mathbf{r}) \phi_i(z), \quad i = 1, \dots, M \quad (1)$$

Here, $\phi_i(z)$ is the local image basis of four scales, 1×1 , 1×2 , 2×1 , and 2×2 patch areas, and the coordinate of the image position is represented by z , $\phi_i(z) = 1$ if location z is contained in the area of the local image element, otherwise $\phi_i(z) = 0$. The number of the image basis is represented as M . Local image bases placed at every location with overlap can be seen as elements of the reconstructed image $\hat{I}(z|\mathbf{r})$, given the fMRI activity $\mathbf{r}$ recorded when the subject was looking at the stimulus image. The combination coefficient λ_i for each local image basis is employed to minimize reconstruction errors. The local decoder, $\psi_i(\mathbf{r})$, predicts the contrast level for the local image element $\phi_i(z)$ using a linearly weighted sum of activation of the fMRI voxels.

The weights in local decoder, w_k^d , denoting the weights of voxel d and contrast label k for the local image basis $\phi_i(x)$, are estimated using the multi-scale sparse multinomial logistic

regression (SMLR) model in Miyawaki's study (2008). Here, we use one local image base as an example to introduce the weight parameter estimation algorithm that states: given a training data set, maximum-likelihood (ML) estimation determines the weight by finding values that maximize the likelihood function,

$$p_w(\mathbf{S}|\mathbf{w}_1, \dots, \mathbf{w}_K) = \prod_{n=1}^N \prod_{k=1}^K p_w(k|\mathbf{r}_n)^{s_{nk}} \quad (2)$$

$$p_w(k|\mathbf{r}) = \frac{\exp[y_{w_k}(\mathbf{r})]}{\sum_{j=1}^K \exp[y_{w_j}(\mathbf{r})]} \quad (3)$$

$$y_{w_k}(\mathbf{r}) = \sum_{d=1}^D w_k^d r^d + w_k^0$$

The Eqs. (2) and (3) are used to estimate the weight parameters for local decoder in Eq. (1), where K is the number of the contrast classes, N is the number of trials, D is the number of voxels. The probability of fMRI pattern evoked by the stimulus image with contrast class k for the n th trial is represented as $p_w(k|\mathbf{r}_n)$, and class label matrix is represented as $\mathbf{S}$, whose element s_{nk} is 1 if the n th trial falls into contrast class k , otherwise 0. The weight vector for class k is represented as $\mathbf{w}_k$, which contains $D+1$ elements (with one bias term). In the SMLR model, the weight parameters w_k^d were estimated using a full-Bayesian approach (Yamashita et al., 2008), where the posterior distribution for each weight parameter was represented by multiplying the above likelihood (Eq. (2)) with a zero-mean normal prior distribution as Eq. (4) (Miyawaki et al., 2008),

$$p(w_k^d|\alpha_k^d) = N\left(0, \frac{1}{\alpha_k^d}\right) \quad (4)$$

where α_k^d is the hyper-parameter denoting the inverse of the variance, or precision, of the weight value for voxel d and contrast class k . The hyper-parameter is also treated as a random variable, whose distribution is defined by $p(\alpha_k^d) = 1/\alpha_k^d$.

One of the most important innovations in Miyawaki's study is the application of a multi-voxel pattern analysis method in element image contrast estimation. Compared to a univariate method, such as the contrast reconstruction model based on retinotopy in the visual cortex (Thirion et al., 2006), the multivariate method not only utilized the distinctiveness of a response to a given contrast due to the regions that responded maximally to that contrast level but also considered population encoding from the pattern in a wide expanse of the cortex (Haxby et al., 2001; Norman et al., 2006). The multivariate method takes into account the correlation of voxels, i.e. the activity of voxels were treated as a whole pattern, which can be used to characterize the correlation/covariance of activation across a set of voxels (such as brain regions), and has been shown to be important in stimulus decoding studies (Averbeck et al., 2006; Chen et al., 2006). Due to the tremendous amount of voxels, in contrast decoders the weights of voxels are assumed to be sparsely distributed to reduce computation consumption. The SMLR model used here simultaneously accomplishes feature selection and prediction of contrast. Feature selection was achieved by assuming sparse distribution of the weight parameters. For SMLR, there are two ways to introduce a priori constraints for sparsity: (1) implicit a priori constraints, based on automatic relevance determination (ARD) as shown in Eq. (4) (Wipf and Nagarajan, 2008; Yamashita et al., 2008), and (2) explicit Laplacian prior constraints, $p(\mathbf{w}) \propto \exp(-\lambda \|\mathbf{w}\|_1)$, where $\|\mathbf{w}\|_1 = \sum_i w_i$ denotes the l_1 norm (Krishnapuram et al., 2005). Mathematically, the computation time for ARD method is proportional to the feature number D^3 , while the latter algorithm only requires computation time proportional to D (Yamashita et al., 2008). In the study of Miyawaki et al. (2008),

Download English Version:

<https://daneshyari.com/en/article/6268938>

Download Persian Version:

<https://daneshyari.com/article/6268938>

[Daneshyari.com](https://daneshyari.com)