RESEARCH ARTICLE

Decoding disparity categories in 3-dimensional images from fMRI data using functional connectivity patterns

Chunyu Liu[1] · Yuan Li[2] · Sutao Song[3] · Jiacai Zhang[1]

Abstract

Humans use binocular disparity to extract depth information from two-dimensional retinal images in a process called stereopsis. Previous studies usually introduce the standard univariate analysis to describe the correlation between disparity level and brain activity within a given brain region based on functional magnetic resonance imaging (fMRI) data. Recently, multivariate pattern analysis has been developed to extract activity patterns across multiple voxels for deciphering categories of binocular disparity. However, the functional connectivity (FC) of patterns based on regions of interest or voxels and their mapping onto disparity category perception remain unknown. The present study extracted functional connectivity patterns for three disparity conditions (crossed disparity, uncrossed disparity, and zero disparity) at distinct spatial scales to decode the binocular disparity. Results of 27 subjects' fMRI data demonstrate that FC features are more discriminatory than traditional voxel activity features in binocular disparity classification. The average binary classification of the whole brain and visual areas are respectively 87% and 79% at single subject level, and thus above the chance level (50%). Our research highlights the importance of exploring functional connectivity patterns to achieve a novel understanding of 3D image processing.

Introduction

One of fundamental visual abilities is the perception of three-dimensional (3D) word. This ability is closely connected with the binocular disparity. Namely: the horizontal separation of the eyes allows both of eyes to have a slightly different viewpoint of the world (Bridge et al. 2007). And the visual system uses this difference to extract depth information from the two-dimensional retinal image, this process is called stereopsis. Even though scientists have spent nearly 50 years studying binocular depth perception in animals and humans and have gained substantial achievements, it is still worth being further researched.

Early researches reveal that the cortical neurons can convey signals related to binocular depth perception. Recordings from the cortices of cats and macaques show that the disparity populations are structured in V1 (Barlow et al. 1967), V2 (Nienborg et al. 2006), V3/V3A (Hubel et al. 2015; Anzai et al. 2011), and MT (DeAngelis et al. 1998; Krug et al. 2011). Functional magnetic resonance imaging (fMRI) studies of the human visual cortex report that V3A, V7, MT, and LO produced disparity-evoked responses (Bridge et al. 2007; Backus et al. 2001; Neri et al. 2004).

As initial studies of the visual field were dominated by univariate analyses, i.e. general linear model (GLM) proposed by Friston et al. (1995), research on the interaction between neurons or voxels was relatively scarce. Recently, however, researchers have introduced multivariate pattern analysis (MVPA) (Norman et al. 2006) to characterize multivoxel pattern selective perception of binocular disparity (Preston et al. 2008; Li et al. 2017), decode the depth

✉ Jiacai Zhang
jiacai.zhang@bnu.edu.cn

[1] College of Information Science and Technology, Beijing Normal University, Beijing, China

[2] School of Electrical and Information Engineering, Tianjin University, Tianjin, China

[3] School of Education and Psychology, University of Jinan, Jinan, China

position (Finlayson et al. 2017) and discriminate the binocular disparity categories (Li et al. 2017) across the human visual cortex. Compared with traditional univariate method, MVPA methods are more sensitive which capture the correlative information among voxels. The most commonly used MVPA methods were the machine learning methods, such as 'one-nearest neighbor' classifier (Haxby et al. 2001), support vector machine (SVM) (Norman et al. 2006), logistic regression model (Yamashita et al. 2008) and random forest (Langs et al. 2011). However, as Naselaris pointed out, most of these MVPA methods use patterns of activity across an array of voxels to discriminate between different levels of stimulus, experimental or task variables (Naselaris et al. 2011). Therefore, they do not fully utilize the discriminative information across brain regions, and the explorations of the hidden connections (such as functional connectivity) between voxels or regions of interest (ROIs) are especially scarce.

Functional connectivity (FC) patterns in fMRI can exhibit dynamic behavior in the range of seconds (Deli et al. 2017) and with a rich spatiotemporal structure, thereby providing a novel representation of mental states. Functional connectivity refers to the functionally integrated relationship between spatially separated brain regions that represent similar patterns of activation (Smith et al. 2011; Tozzi et al. 2017). FC can be extracted at different spatial scales (Betzel et al. 2017), such as at the level of the whole brain, across a set of related brain regions, or within a specific ROI. Establishing FC patterns first requires the definition of functional nodes based on spatial ROIs, after which a connectivity analysis between each pair of nodes based on their fMRI time series may be performed (Pantazatos et al. 2012).

Functional connectivity (FC) patterns have been successfully used to decode subject-driven cognitive states, such as memory retrieval, silent-sing, mental arithmetic, rest (Shirer et al. 2012), emotion categories (Pantazatos et al. 2012; Dasdemir et al. 2017), object categories (Hutchison et al. 2014; Wang et al. 2016; Fields et al. 2017; Liu et al. 2018), semantic representation (Fang et al. 2018; Mizraji et al. 2017), visual attention (Parhizi et al. 2018), and tracking ongoing cognition (Gonzalez-Castillo et al. 2015). However, whether the disparity information in 3D perception can be represented by FC patterns and whether functional connectivity across brain areas facilitates the discrimination of disparity categories still remain unknown. Previous studies using traditional MVPA methods have revealed that many single brain regions are involved in disparity processing (Preston et al. 2008; Li et al. 2017). However, few studies have investigated perceptual binocular disparity at the whole-brain level, as it may require the amalgamation of calculations conducted at the voxel or neuronal level. Calculations based on FC

patterns are much smaller than those based on voxel or neuronal patterns, providing the approach suitable to study the perception of binocular disparity at larger scales level. In addition, the potential contributions of FC patterns to characterize binocular disparity recognition at large scales remain unexplored.

To estimate FC pattern of implicit disparity processing, we introduced traditional fMRI block design to obtain blood oxygen level dependent (BOLD) activity. To decode disparity categories, the present study constructed FC patterns at three different spatial scales: whole brain, visual cortex, and single ROI. The discrimination of FC patterns is evaluated by Random Forest classifier. To verify the superiority of FC patterns, we also compared accuracies achieved when using interactivity (pair-wise correlations) versus independent voxel (voxels time series) within the same visual-cortex ROIs.

The remainder of the paper is organized as follows. "Materials and methods" section describes our datasets and the details of the proposed data-processing method. In "Results" section, the results of accuracies of decoding the three disparity level accuracies under the three spatial scale FC patterns were presented. Finally, in "Discussion" section, the details of results are discussed, including the potential implications of this study and indication of future directions of related researches.

Materials and methods

Materials

Participants

Twenty-seven participants (13 men; right-handed; median age, 25 years) volunteered for the experiment and were compensated for their time. All participants had normal or corrected-to-normal visual acuity; provided written, informed consent; and were debriefed on the purpose of the study immediately following completion. Participant recruitment and experiments were conducted with the approval of the Beijing Normal University (BNU) Imaging Center for Brain Research, National Key Laboratory of Cognitive Neuroscience and Learning as well as the Human Research Ethics Committee. All participants were screened for stereo deficits using four stereo tests (Yan 1985). The stereo tests guaranteed that the participants were able to distinguish between crossed and uncrossed disparities in the experiment. Meanwhile, after the experiment, we asked each participant if they could perceive the disparity in the crossed-disparity and uncrossed-disparity tasks. All participants responded that they could perceive the disparity.

fMRI data acquisition

A 3T Siemens scanner equipped for echo planar imaging (EPI) at the Brain Imaging Center of BNU was used for image acquisition. For each participant, functional images were collected with the following parameters: repetition time (TR), 2000 ms; echo time (TE), 30 ms; flip angle (FA), 90°; field of view (FOV), 200×200 mm^2; in-plane resolution, 3.13×3.13 mm^2; 33 slices; slice thickness, 3.5 mm; gap thickness, 0.7 mm. The localizer scans were obtained using the EPI with the following parameters: TE, 30 ms; TR, 2000 ms; flip angle, 90°; FOV; 192×192 mm^2; in-plane resolution, 3×3 mm^2; 33 slices; slice thickness, 3 mm; gap thickness, 0 mm.

Stimuli and experimental design

The present study used forty stimulus shapes (Fig. 1) with a line width of 1.3°. Each shape covered a maximum visual angle of $12° \times 12°$ and was presented against a mid-gray background ($26° \times 14°$). A CIE 1931 xyY color space was used to describe the luminance and chromaticity values of the shape (Y = 1130, X = 1/3, y = 1/3) and background (Y = 672, x = 1/3, y = 1/3). For each shape, we generated stimuli with 3 different levels of disparity ($\pm$ 30 arcmin and 0 arcmin). The positive disparity levels (+ 30 arcmin) correspond to uncrossed-disparity and the negative disparity levels (− 30 arcmin) correspond to crossed-disparity. The stimuli presented in our experiments totaled to 120. Each stimulus was displayed using a 3D LCD with LED Backlight (LGD 2343p, 1920×1080 pixels)

positioned in the bore of a magnet at a distance of 110 cm from the point of observation. Participants wore polarized glasses and viewed the stimuli through a mirror attached to the head coil and tilted at 45°.

For the selection of crossed and uncrossed disparity levels, the disparity levels varied from − 36 to 36 arcmin in studies of Preston et al. (2008) and Goncalves et al. (2015). In our study, the crossed and uncrossed disparity levels ($\pm$ 30 arcmin) is within the disparity range in those previous studies. Moreover, Minini's study pointed out that large disparities can induce stronger neural activities in visual cortex than small disparities (Minini et al. 2010). Meanwhile, previous fMRI studies indicated that larger disparity can easily cause the visual fatigue (Lambooij et al. 2009) and the decrease in cortical activity (Backus et al. 2001). Considering the cortical activity and the visual fatigue, very large or small disparity levels are not fit for our experiment. So the selected disparity levels in our study ($\pm$ 30 arcmin) should be reasonable.

The fMRI experiment featured a block-design paradigm that consisted of two runs. Crossed-disparity (C), uncrossed-disparity (U), and zero-disparity (Z) tasks were presented in a pseudorandom fashion. The stimuli sequences CUZUZCZUCZCU and UCZCZUZCUZUC indicate two sequential runs, where C, U, and Z refer to crossed-disparity, uncrossed-disparity and zero-disparity blocks, respectively. Each run consisted of twelve 24-s task blocks. Between successive blocks, we displayed a 12-s fixation block. An individual trial consisted of a stimulus shape presentation period of 1.5 s and an inter-stimuli interval of 0.5 s, thus 2 s in total. Twelve trials with the same disparity
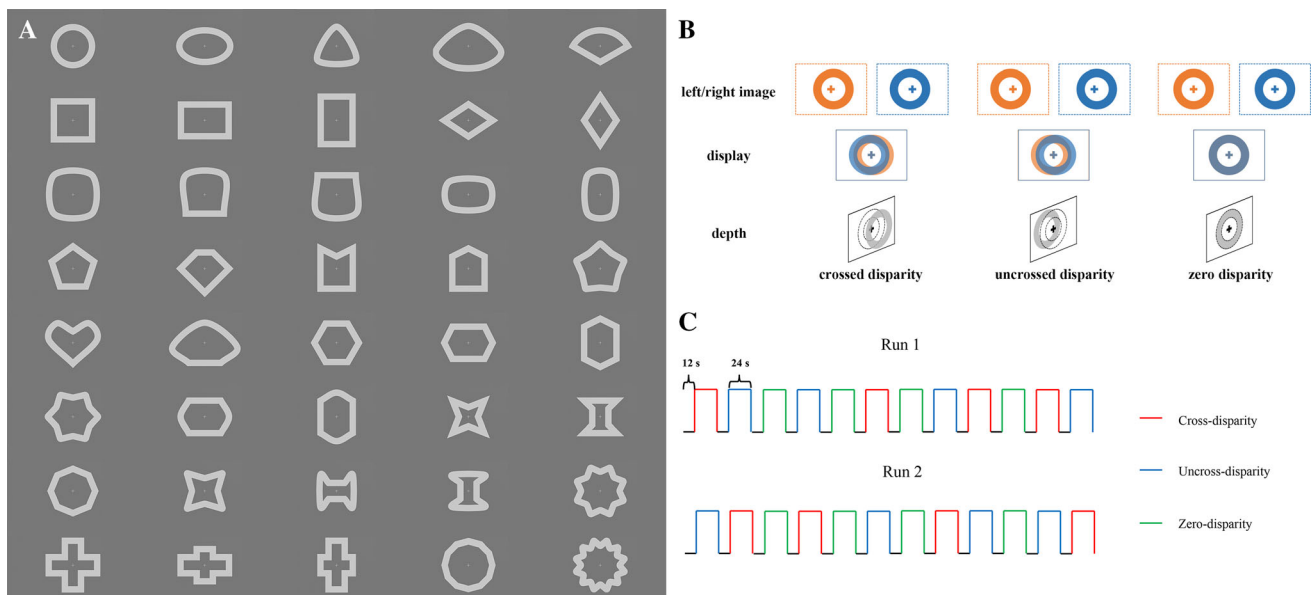

Fig. 1 Schematic illustration of the stimuli-viewing task (Li et al. 2017). **a** Forty shapes that were used to generate stimuli with different disparities. **b** Diagram of the depth arrangement in the stimuli. **c** The experiment featured a block-design paradigm that consisted of two runs

were presented in a given task block. Twelve stimulus shapes included in each block were randomly selected from forty stimulus shapes (Fig. 1). During the task block, each participant was required to press a button with his or her right index finger if the two successive stimuli were different and with his or her left index finger if the two successive stimuli were the same. During the fixation blocks, participants were asked to fixate on a cross at the center of the screen.

Data preprocessing

The preprocessing of functional images was performed using SPM8 (Welcome Department of Cognitive Neurology, University College of London, London, UK; https://www.fil.ion.ucl.ac.uk/spm/) and the DPABI toolbox (Chao-Gan et al. 2016). The processing included the deletion of the first five volumes due to artifacts related to signal stabilization in the initial images of each session, realignment, normalization, and smoothing. The fMRI images were realigned to the first image of the scan run and normalized to the Montreal Neurological Institute (MNI) template. The voxel size of the normalized image was set to $3 \times 3 \times 4$ mm^3. Images were smoothed with an 8-mm full-width at half maximum (FWHM) Gaussian kernel.

ROI

For whole-brain FC patterns, we used the Harvard–Oxford Atlas to obtain 112 ROIs across the entire in every subject. For FC patterns across visual cortical areas, we mapped visual regions to select ROIs for each participant, using standard mapping techniques. Retinotopic areas V1, V2d, V3d, V7, V2v, V3v, and V4v were defined using rotating wedges and expanding stimuli comprised of concentric rings. V4v was defined as the region of retinotopic activation in the ventral visual cortex adjacent to V3v. V7 was defined as the region anterior and dorsal to V3A. The human motion complex (hMT +/V5) and localizing image (LOC) were defined using independent localizer scans. The hMT +/V5 area was defined as the set of voxels in the lateral temporal cortex that responded with a significantly high rate ($p < 10^{-4}$) to a coherently moving array of dots. The LOC was defined as the set of voxels in the lateral occipito-temporal cortex that produced significantly stronger responses ($p < 10^{-3}$) to intact object images than to scrambled images (Kourtzi et al. 2000). The scrambled images were created by dividing the intact images in a square grid and by scrambling the positions of each of the resulting squares.

For each visual ROI in the Talairach space, gray matter voxels were selected from both hemispheres and sorted according to their T-statistics calculated by contrasting all stimulus conditions (C, U, and Z) with the fixation baseline across the two runs. For each hemisphere of each ROI, 100 voxels were selected from among those with a T of > 0 as calculated via the aforementioned T-statistical analysis. When 200 voxels were not available in a cortical area, we used the maximum number of voxels that featured a T-statistic of > 0. For example, only 194 voxels were selected from the V3v area in participant 1.

Methods

Sliding window

The classification problem in fMRI data is described using the following mathematical formula:

$$f : \quad fMRI\_sequence(t_1, t_2) \rightarrow k \tag{1}$$

where $t_1, t_2 (t_1 < t_2)$ are two different time points, and k is the category of binocular disparity of the 3D images viewed by the subject. We thus mapped the FC patterns constructed in the time interval (t_1, t_2) to the respective binocular disparity categories.

Previously, a limited set of mental tasks (Shirer et al. 2012; Richiardi et al. 2011), different categories of object (Wang et al. 2016), and arousal levels (Tagliazucchi et al. 2012) were discriminated in task fMRI studies on the basis of localized changes in FC in the range of 45 s to 2 min. In our experiment, one block only consists of 12 time points, and we were concerned that fMRI time series in one block are not long enough to estimate a stable FC pattern. In addition, Pantazatos et al. (2012) have adopted fMRI time series of two blocks to estimate FC patterns to decode unattended fearful face states successfully. Based on these findings, we adopted two different window lengths (48 s and 72 s) to estimate the FC pattern of each disparity task across multiple blocks.

We concatenated all TRs under the same disparity categories and obtained corresponding concatenated fMRI time series for C, U, and Z. There was a total of 96 TRs for each concatenated category time series. Segmentation of the fMRI time series without any overlap would have yielded small sample sizes and resulted in overfitting. Instead, a sliding-window method allowed for the analysis of temporal cross-correlations between windowed segments of the BOLD fMRI signal over time. Here, the sliding window had a step of one TR along the fMRI time series. FC patterns for the nodes i and j in the sliding window were estimated based on the correlations among ROIs as follows:

$$\rho_{i,j,t,w} = \frac{\sum_{s=t-w-1}^{t}(y_{i,s} - \hat{\mu}_{i,w})(y_{j,s} - \hat{\mu}_{j,w})}{\sqrt{\left(\sum_{s=t-w-1}^{t}(y_{i,s} - \hat{\mu}_{i,w})^2\right)\left(\sum_{s=t-w-1}^{t}(y_{j,s} - \hat{\mu}_{j,w})^2\right)}}$$

(2)

where w is the length of the sliding window, t represents the time point, w indicates $t_2 - t_1$, and i, j denote the ROI index (or voxels), $\rho_{i,j,t,w}$ represents the functional connectivity values of two brain regions (ROIs or voxels) under time-series length w, and $\mu_{i,w}$ represents the time series mean value along the sliding window w.

We obtained the FC pattern matrices between each pair of nodes for each fMRI segment by:

$$X_t^k = (\rho_{i,j,t,w})_{i,j}$$

(3)

where $i, j = 1, \ldots, m$, m represents the number of nodes (ROIs or voxels), X_t^k represents the FC pattern matrices of k category of binocular disparity at time point t. Each FC pattern was a $m \times m$ connectivity symmetry matrix. Given the symmetry of FC pattern matrices, the lower-left triangle matrix was used to reflect the connectivity strengths among all m ROIs.

FC pattern space

For each binocular disparity category, the brain-state space was denoted as Ω^k, while k indicated the category of binocular disparity level of a given stimulus. In previous studies on binocular disparity, the brain-state space was represented by patterns of voxel activation, assuming that voxel patterns are independent of one another. This notion has dominated many methods via an essential spatial coding mechanism across signal independent-neuronal populations, ignoring the relationships among regions. In fact, the FC patterns can identify these correlations among brain regions. In our study, a brain-state space of binocular disparity was projected onto the space represented by FC features. We denoted the FC pattern samples in each category space as

$$
\begin{aligned}
x_t^k &= linearly\{lowhalf(X_t^k)\} \\
&= (\rho_{1,1,t}, \rho_{2,1,t}, \ldots, \rho_{i,j(i \geq j),t}, \ldots, \rho_{m,m,t}) \\
&= (x_{1,t}, x_{2,t}, \ldots, x_{M,t}), \quad M = \frac{m(m-1)}{2}.
\end{aligned}
$$

(4)

where x_t^k represents FC pattern samples in k category space. All FC patterns x_t^k for each disparity category were pooled together to yield the FC pattern spaces of each category. Furthermore, we calculated the lower-left triangle matrix

$$X^k = \{x_t^k | t = 1, \ldots, n\}, X^k \subseteq \Omega^k$$

(5)

where n is the number of samples (FC patterns) for a given disparity category, and k denotes the disparity category.

The construction of FC-pattern spaces can be summarized into two steps: (1) all the time series for a given category were segmented with a sliding window, and (2) windowed fMRI time series of all m ROIs were used to compute the FC matrix. Within each window, we averaged all voxel values within a ROI to compute ROI the time series. We then calculated the ROI-pairwise correlations. The lower-left triangle matrix X^k was used as the brain representation of disparity stimuli in the classification model. Figure 2 shows the flow of the decoding disparity categories using MVPA based on FC patterns.

MVPA

MVPA refers to a diverse set of methods that analyze neural response as patterns of activity that reflect the varying brain states that a cortical field or system can produce (Haxby 2012). MVPA analyzes brain activity with methods such as pattern classification, representational similarity analysis, hyperalignment, and stimulus-model-based encoding and decoding. MVPA treats the fMRI signal as a set of pattern vectors stored in an $N \times M$ matrix with N observations (e.g., stimulus conditions, time points) and M features (e.g., voxels, cortical surface nodes, functional connectivity) defining an M-dimensional vector space (Haxby et al. 2014). The goal of MVPA is to analyze the structure of these high-dimensional representational spaces. The random forest (RF) is an ensemble classifier that uses decision trees as base learners (Breiman 2001). In visual fMRI task, Langs et al. (2011) found that RF could detect stable distributed patterns of brain activation and capture the complex relationships between the experimental conditions and the fMRI signals. Therefore, RF was used to classify the functional connectivity patterns of different tasks.

Classifier

We denoted the training data and test data as follows:

$$
\begin{aligned}
T_{train} &= \{(x_i, k) | i = 1, \ldots, n_1\}, \quad x_i \in X^k \\
T_{test} &= \{(x_j, k) | j = 1, \ldots, n_2\}, \quad x_j \in X^k
\end{aligned}
$$

(6)

where x_i, x_j represents the FC patterns, k denotes the disparity category, n_1, n_2 represent the sample number. The four fMRI blocks from the first runs were used to train the classifier and the four fMRI blocks from the second runs were used compile the test dataset.

Decision trees are discriminative classifiers performing recursive partitioning of a feature space to yield a potentially non-linear decision boundary. The objective of
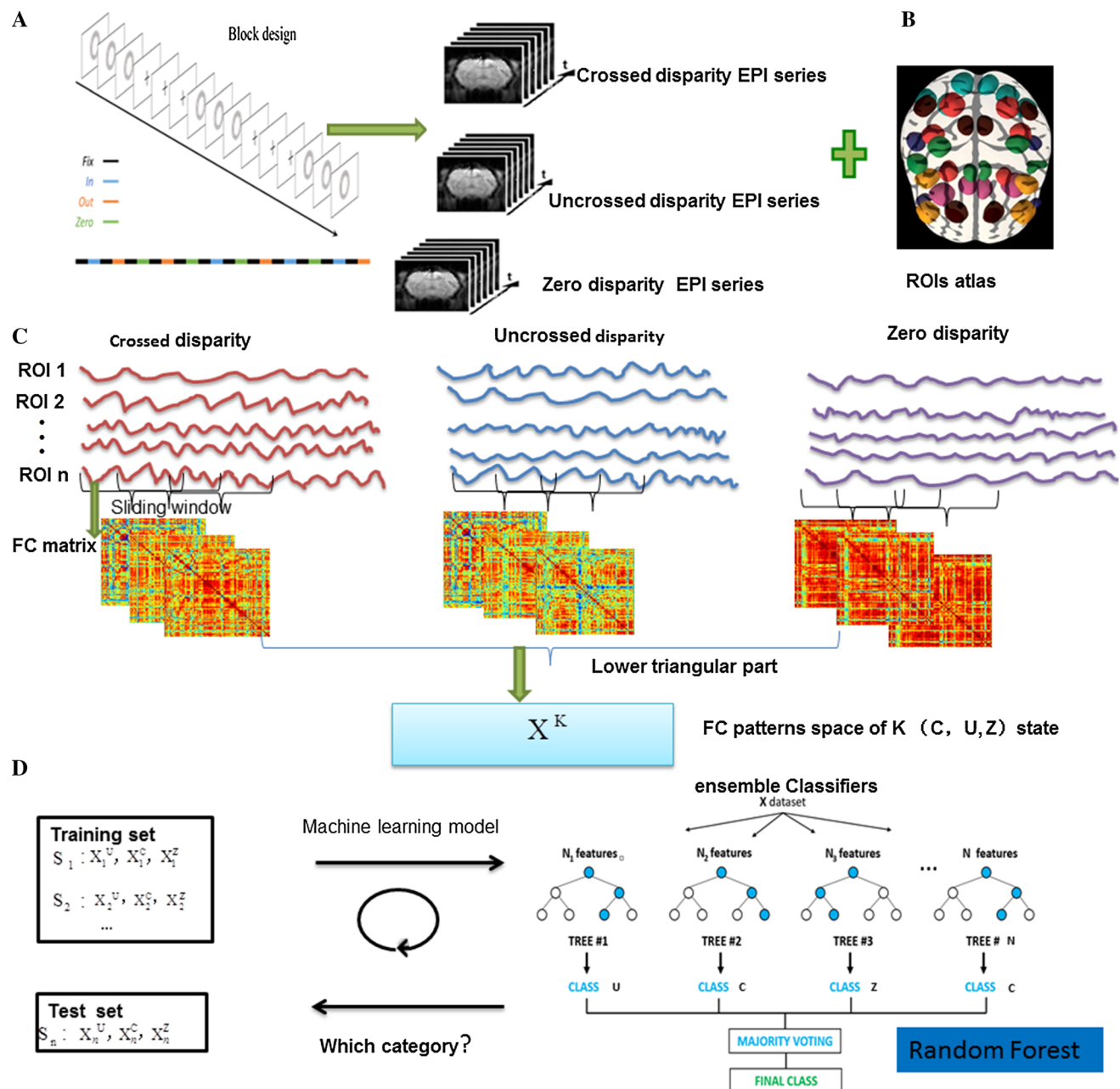
A

Block design

Crossed disparity EPI series

Uncrossed disparity EPI series

Zero disparity EPI series

Fix —
In —
Out —
Zero —

B

ROIs atlas

C

Crossed disparity   Uncrossed disparity   Zero disparity

ROI 1
ROI 2
ROI n

Sliding window

FC matrix

Lower triangular part

X^K

FC patterns space of K （C，U, Z) state

D

Training set
$S_1 : X_1^U, X_1^C, X_1^Z$

$S_2 : X_2^U, X_2^C, X_2^Z$

...

Test set
$S_n : X_n^U, X_n^C, X_n^Z$

Machine learning model

Which category?

ensemble Classifiers
X dataset

N_1 features N_2 features N_3 features N features

TREE #1   TREE #2   TREE #3   TREE # N

CLASS U   CLASS C   CLASS Z   CLASS C

MAJORITY VOTING

FINAL CLASS

Random Forest

Fig. 2 The flowchart of the task and analysis procedures. **a** Participants were presented in various block (24 s in length) with depth information of three categories-crossed disparity (C), uncrossed disparity (U), and zero disparity (Z), putting together the same depth information EPI sequences by the cross block technique. **b** Choosing appropriate ROIs atlas. **c** Constructing FC patterns space. Firstly, time series were extracted from each ROI, and then using a length window with fixed step to sliding the time series, calculating the pairs of ROI time sequence correlation coefficient under the fixed window length relative to each category, last taking the lower half of the matrix and transform it into a vector. **d** Category-paired samples and Random Forest machine learning model were trained in training set and tested in test set. X_i^K represent i set of FC patterns space of K state (C, U, Z)

decision tree building is to obtain the set of decision rules that can be used to predict the outcome of a set of input variables. At each decision node of the decision tree, the CART algorithm (Breiman 2017) seeks cut-off points in the continuous variable function f that minimize the Gini value. The Gini value was computed as follows:

$$i(f') = \sum_{k \neq l} p(k|f')p(l|f') \tag{7}$$

where k, l denote the class of the disparity level, and $p(k|f')$ is the conditional probability of class k for the FC pattern feature f'.

The FC samples were randomly selected from the training set to yield a new training set. An arbitrary number of K attributes of the samples in the new training set were then randomly selected to train a Decision Tree classifier using the CART methodology (Breiman 2017) without pruning. Repeating the above process N times we obtained a Random Forest model containing N classifiers. The final disparity category was voted by the N classifiers. In the experimental process N was set to 40. The random forest was implemented with sklearn (Abraham et al. 2014) on Python 3.6 with the following parameters: bootstrap = True, class_weight = None, criterion = Gini value, max_depth = None, max_feature = sqrt, max_leaf_nodes = None, min_impurity_decrease = 0.0, min_impurity_split = None, min_samples_leaf = 1, min_samples_split = 2, min_weight_fraction_leaf = 0.0, n_trees = 40, random_state = None.

Results

Results of individual classifiers

Figures 3 and 4 show the decoding accuracies of the FC patterns under the three disparity conditions at the individual level. FC patterns were extracted at different sliding-window lengths (48 and 72 s) and brain areas (whole brain and visual areas) to train classifiers for each subject.

For the 48-s (24-TR) sliding window, we obtained 25 (48 − 24 + 1) training samples from the first run and 25
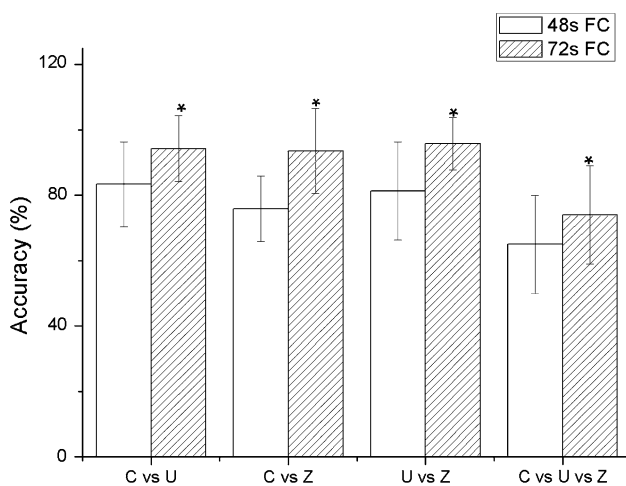

Fig. 3 The average decoding classification accuracies from two kind time scale FC pattern on individual subject. Figure showed the three binary classification, including crossed disparity(C) versus uncrossed disparity(U), crossed disparity(C) versus zero disparity(Z), uncrossed disparity(U) versus zero disparity(Z), and one three classification (C vs U vs Z). Error bar represent the variance between all subjects. Asterisk represent the significant difference ($p < 0.05$), which was obtained by paired t test
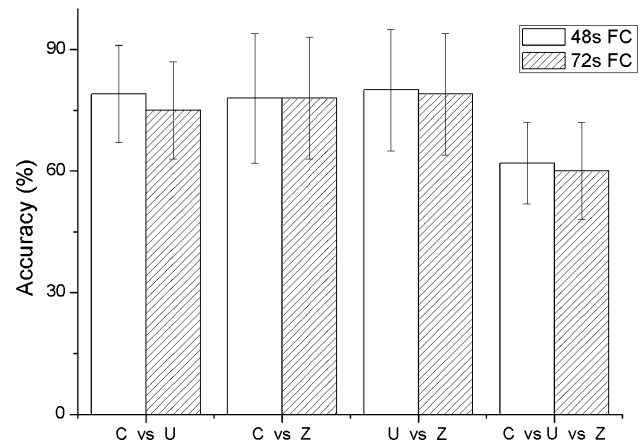

Fig. 4 The average decoding classification accuracies from two kinds of brain visual area FC pattern on individual subject. Figure show three binary classification, including crossed disparity(C) versus uncrossed disparity(U), cross disparity(C) versus zero disparity(Z), uncrossed disparity(U) versus zero disparity(Z), and one three classification (C vs U vs Z). Error bar represent the variance between all subjects. Asterisk represent the significant difference ($p < 0.05$), which was obtained by paired t test

testing samples from the second run per subject and category. For the 72-s (36-TR) sliding window, we obtained 13 (48 − 36 + 1) training samples from the first run and 13 testing samples from the second run. Figures 3 and 4 display the individual results from the FC patterns in the whole brain and visual areas, respectively.

The average binary accuracy was 80.2% for the 48-s FC patterns covering the whole brain and 94% when the sliding window increased to 72 s (Fig. 3). The classification accuracy of 72-s window was significantly better ($p < 0.05$) than that of the 48-s FC pattern for both the two-category and the three-category classification.

The average binary classification accuracies of the FC patterns restricted to visual areas were 79% for the 48-s FC pattern and 78% for the 72-s FC pattern (Fig. 4). There was no significant difference between the two values. Comparing these results to those obtained for the whole-brain FC pattern, we found that the classification effect was better for the whole brain than for the ROIs in a single visual area.

Results of common classifiers

Our analysis of individual results demonstrated that FC patterns computed from windowed fMRI time series can predict the viewed image category at the individual level. We further investigated the comparison among individual FC patterns with the performance of a universal classifier in all subjects. All samples from the 27 subjects were pooled together to yield an across-subject dataset. For the 48-s and 72-s sliding windows, the training datasets across

subjects contained 1350 (50 × 27) and 702 (26 × 27) samples per disparity category, respectively. We randomly divided the data set into five subsets and applied the Random Forest models to classify the FC patterns according to five-fold cross-validation. The average decoding accuracy results of the binary classification and the three-way classification of the FC patterns in the whole brain and the visual areas are shown in Figs. 5 and 6, respectively.

The average binary decoding accuracies for common classifiers were 58% (48-s FC) and 59% (72-s FC) at the whole-brain level (Fig. 5). The difference between these measurements was non-significant. The across-subject average binary decoding accuracies were 67% (48-s FC) and 65% (72-s FC) at the brain visual level (Fig. 6). The difference between the two values was also non-significant. Of note, both pairs of values were significantly above chance (50%). Comparing the results presented in Figs. 5 and 6, we found that the classification effect of the FC pattern in the visual area is better than the classification effect of the whole-brain FC pattern, regardless of the time window.

Comparison of voxel activation and FC patterns

The classification results using voxel activation and FC patterns of selected ROIs in visual areas are shown in Fig. 7. The time series course of each voxel was detrended and transformed into a z-score for each experimental run. The FC pattern in the signal ROI was obtained by
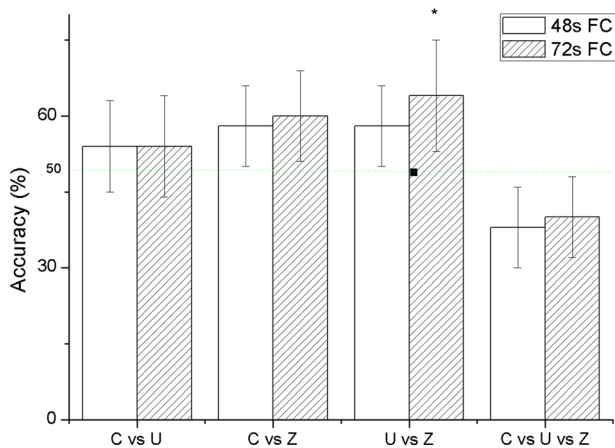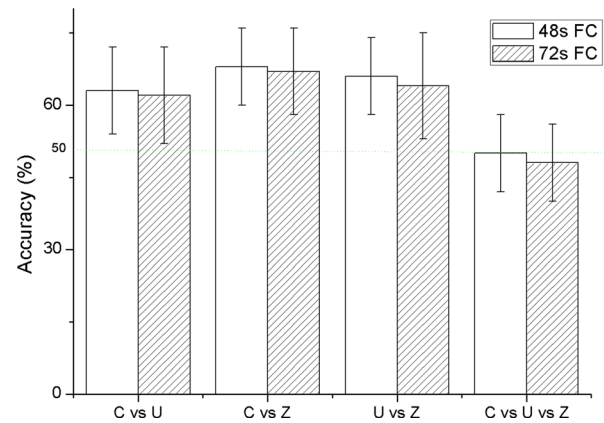

Fig. 6 The average decoding classification accuracies from two kinds of brain visual area FC pattern across subject. Figure show the three binary classification, including crossed disparity (C) versus uncrossed disparity (U), crossed disparity (C) versus zero disparity (Z), uncrossed disparity (U) versus zero disparity (Z), and one three classification (C vs U vs Z). Error bar represent the variance between all subjects

calculating the correlation matrix among voxels in the ROI with a 72-s sliding window.

We found that the classification accuracies of the FC pattern were significantly better than those of the voxel pattern for all ROIs in all subjects (Fig. 7). We also found that V7, V3d, V3A, and MT showed the highest accuracy during the decoding disparity level of the FC patterns.

Discussion

The present experiment represents that FC patterns extracted from fMRI data can be used to discriminate disparity categories in 3D images. As the extraction of the FC pattern requires enough long time series of fMRI data, we concatenated TRs in different blocks that presented stimuli of the same categories. We then used sliding-window technology to construct fMRI segments corresponding to FC samples. To obtain the nodes in the FC patterns, we considered different spatial scales, including whole-brain and visual-area FC patterns. We then verified the performance of binocular disparity decoding from the FC patterns by adopting a machine-learning method (random forest). We also evaluated the effect of decoding disparity level at the individual and across-subject levels. Our finding that decoding accuracies were generally above chance level demonstrates that FC patterns represent the information that is involved in the processing of disparity categories. Our comparison between the classification effects of the FC patterns and voxel values in visual-area ROIs further supports the superiority of the former.

To decode disparity categories at the individual level, we adopted two sliding windows (24 TR and 36 TR) at
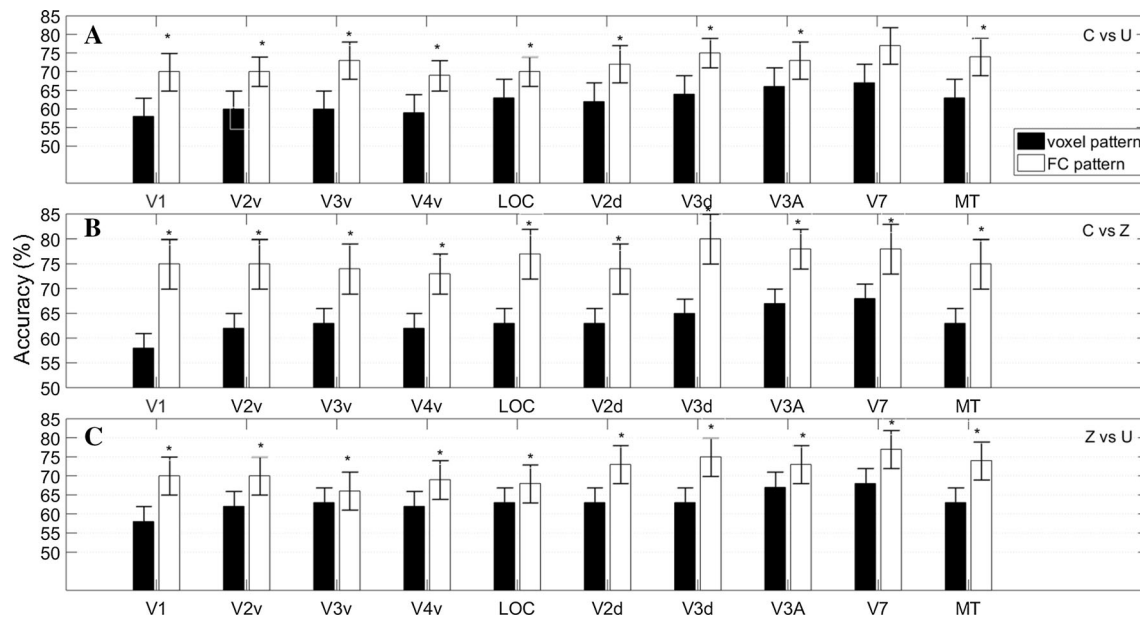

Fig. 5 The average decoding classification accuracies from two kinds of whole brain FC pattern across subject. Figure show the three binary classification, including crossed disparity (C) versus uncrossed disparity (U), crossed disparity (C) versus zero disparity (Z), uncrossed disparity (U) versus zero disparity (Z), and one three classification (C vs U vs Z). Error bar represent the variance between all subjects. Asterisk represent the significant difference ($p < 0.05$), which was obtained by paired t test

Fig. 7 The comparison results of decoding disparity categories using MVPA based on functional connection pattern and voxel pattern within 10 visual ROIs. The error bar represented the standard error mean. Asterisk represent the significant difference ($p < 0.05$), which was obtained by paired t test. **a** The mean prediction accuracies of crossed disparity(C) versus uncrossed disparity (U). **b** The mean prediction accuracies of crossed disparity (C) versus zero disparity (Z). **c** The mean prediction accuracies of zero disparity (Z) versus uncrossed disparity (U)

three spatial scales (ROI, visual area, whole brain). For the whole-brain and visual-area FC patterns, we found that both sliding windows could decode the disparity category with accuracy above chance level, indicating that FC patterns contain discrimination information for different stimulus tasks when the subject viewing different disparity categories at short temporal scales. These results agree with those of previous studies: FC can reflect discriminative information for decoding categorical (Wang et al. 2016; Stevens et al. 2015) or semantic information (Pantazatos et al. 2012; Fang et al. 2018). For the whole-brain FC patterns, we found that the classification effect of the 72-s FC pattern was significantly higher than that of the 48-s FC pattern but identified no significant differences among the local visual-area FC patterns. This finding suggests that additional time points may be important to estimate the of stable FC patterns at large-scale spatial FC patterns level. Furthermore, the classification results of the whole-brain FC pattern were better than those of the local visual-area FC pattern. This finding agree with those of previous studies reporting that: the whole-brain FC pattern contains more information to decode brain states (Shirer et al. 2012; Gonzalez-Castillo et al. 2015; Richiardi et al. 2011). For the whole-brain FC pattern, the mean accuracies of the 48-s and 72-s FC patterns were above 80%, providing further support for the feasibility of separating fMRI time series into time windows of 30–60 s to identify favorable whole-

brain connectivity (Shirer et al. 2012; Gonzalez-Castillo et al. 2015; Allen et al. 2014; Yang et al. 2014).

For the across-subject classification results, the effect accuracy of the whole-brain FC patterns was much lower than the individual-subject level, indicating variance between subjects was large, and the whole-brain FC patterns cannot overcome variance at a short temporal scale. However, the classification effects of local visual-area FC patterns were better than whole-brain FC patterns. This finding may be attributed to less variance between subjects at the level of local visual areas than at that of the whole brain.

As shown in Fig. 7, the classification accuracies using MVPA based on FC patterns were higher than those based on the voxel time series values. Our comparison demonstrates that the three binary classifiers utilized more discrimination information to decode disparity information than that conveyed in the voxel time series. This conclusion is also confirmed by previous research (Pantazatos et al. 2012; Wang et al. 2016). Pantazatos et al. (2012) found that functional connectivity features are more informative than beta estimates derived from the original, smaller voxel size when they decoded unattended fearful faces using MVPA based on voxel patterns and FC patterns. Wang et al. (2016) successfully decoded four object categories based on FC patterns in whole-brain areas, even when the contributions of regions showing classical category-selective activations (general linear model analysis

based on voxels) were excluded. In the present study, the decoding results from the ROIs in the areas V7, V3d, V3A, and MT were relatively better than those from other ROIs in visual areas, which support previous findings that ROIs in these visual areas better perceive disparity categories (Bridge et al. 2007; Anzai et al. 2011; DeAngelis et al. 1998; Neri et al. 2004). Of note, V7, V3d, and MT are situated in dorsal areas of the brain; our results therefore agree with previous reports of dorsal areas evincing more adaptation to disparity categories (Preston et al. 2008; Li et al. 2017). Previous research has further shown that areaV3A contributes the most to the perception of disparity levels; however, our results contradict this finding. As V3A is adjacent to V3d and V7, the ROIs of these areas may overlap, and we therefore speculate that small deviations that rise when the single ROIs are manually cut may lead to the conflicting data. Regardless of this inconsistency, prior research and our own results indicate that visual areas that respond strongly to binocular disparity are located in the dorsal visual cortex.

Several additional factors may contribute to the high accuracy of decoding disparity information. First, functional connectivity across ROIs contain more information than a single ROI or voxel. Second, the Random Forest model is a boost classifier that can improve the accuracy of category decoding. Third, we used sliding window technology to construct functional connectivity patterns. A sliding window analysis is typically marred by a relatively small number of data points, which may be overcome with recent technological advances in fMRI acquisition (Baczkowski et al. 2017). For example, multiband imaging allows for sub-second repetition times (Feinberg et al. 2010) and thereby increases the number of sample points. This feature further increases the robustness of correlation estimates and enables a better characterization of high-frequency components, which are mostly affected by non-neuronal noise, such as cardiac and respiratory rhythms. Finally, the face that we considered whole-brain connectivity patterns instead of a subset of connections may have increased accuracy for the shortest windows. Gonzalea-Castillo et al. (2015) also found that valuable information for tackling cognition is spatially distributed across the whole brain.

Even though accuracies of decoding binocular disparity categories by FC patterns were higher than voxel patterns, the former method still has some limitations. For example, estimating FC patterns requires much more time information than MVPA methods based on voxel patterns. Finally, our investigation did not consider the peak information of brain networks, such as vertex-position information or vertex angle information. Combining our results with such peak information may better characterize brain state patterns to decode 3D visual categories.

Conclusion

In this study, MVPA based on FC patterns was adopted to decode disparity categories. We constructed FC patterns at three different spatial scales by calculating the correlations of BOLD time series across different ROIs or voxels. Our results indicate that FC patterns could be used as brain-state patterns to decode 3D visual information. Comparing the effects of MVPA based on FC patterns and those based on voxel patterns, we found that FC patterns contained more discrimination information during the decoding of disparity categories. Our method is different from previously reported MVPA methods based on voxels or ROIs, in that, it identifies effective connectivity information patterns whose feature representation functions have yet to be investigated in 3D image. Our research therefore highlights the importance of exploring functional connectivity patterns to achieve a full understanding of 3D image processing.

References

Abraham A et al (2014) Machine learning for neuroimaging with scikit-learn. Front Neuroinform 8:14

Allen EA et al (2014) Tracking whole-brain connectivity dynamics in the resting state. Cereb Cortex 24:663–676

Anzai A et al (2011) Coding of stereoscopic depth information in visual areas V3 and V3A. J Neurosci 31:10270–10282

Backus BT et al (2001) Human cortical activity correlates with stereoscopic depth perception. J Neurophysiol 86:2054–2068

Baczkowski BM et al (2017) Sliding-window analysis tracks fluctuations in amygdala functional connectivity associated with physiological arousal and vigilance during fear conditioning. NeuroImage 153:168–178

Barlow HB, Blakemore C, Pettigrew JD (1967) The neural mechanism of binocular depth discrimination. J Physiol 193:327–342

Betzel RF et al (2017) Multi-scale brain networks. NeuroImage 160:73–83

Breiman L (2001) Random forests. Mach Learn 45:5–32

Breiman L (2017) Classification and regression trees. Routledge

Bridge H et al (2007) Topographical representation of binocular depth in the human visual cortex using fMRI. J Vis 7(15):1–14

Chao-Gan Y et al (2016) DPABI: data processing & analysis for (resting-state) brain imaging. Neuroinformatics 14:339–351

Dasdemir Y et al (2017) Analysis of functional brain connections for positive–negative emotions using phase locking value. Cogn Neurodyn 11:487–500

DeAngelis GC et al (1998) Cortical area MT and the perception of stereoscopic depth. Nature 394:677

Deli E et al (2017) Relationships between short and fast brain timescales. Cogn Neurodyn 11:539–552

Fang Y et al (2018) Semantic representation in the white matter pathway. PLoS Biol 16:e2003993

Feinberg DA et al (2010) Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. PLoS ONE 5:e15710

Fields C et al (2017) Disrupted development and imbalanced function in the global neuronal workspace: a positive-feedback mechanism for the emergence of ASD in early infancy. Cogn Neurodyn 11:1–21

Finlayson NJ et al (2017) Differential patterns of 2D location versus depth decoding along the visual hierarchy. NeuroImage 147:507–516

Friston KJ et al (1995) Analysis of fMRI time-series revisited. NeuroImage 2:45–53

Goncalves NR et al (2015) 7 Tesla FMRI reveals systematic functional organization for binocular disparity in dorsal visual cortex. J Neurosci 35:3056–3072

Gonzalez-Castillo J et al (2015) Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns. Proc Natl Acad Sci 112:8762–8767

Haxby JV (2012) Multivariate pattern analysis of fMRI: the early beginnings. NeuroImage 62:852–855

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

Haxby JV et al (2014) Decoding neural representational spaces using multivariate pattern analysis. Annu Rev Neurosci 37:435–456

Hubel DH et al (2015) Binocular stereoscopy in visual areas V-2, V-3, and V-3A of the macaque monkey. Cereb Cortex 25:959–971

Hutchison RM et al (2014) Distinct and distributed functional connectivity patterns across cortex reflect the domain-specific constraints of object, face, scene, body, and tool category-selective modules in the ventral visual pathway. NeuroImage 96:216–236

Kourtzi Z et al (2000) Cortical regions involved in perceiving object shape. J Neurosci 20:3310–3318

Krug K et al (2011) Neurons in dorsal visual area V5/MT signal relative disparity. J Neurosci 31:17892–17904

Lambooij M et al (2009) Visual discomfort and visual fatigue of stereoscopic displays: a review. J Imaging Sci Technol 53:30201-1–30201-14

Langs G et al (2011) Detecting stable distributed patterns of brain activation using gini contrast. NeuroImage 56:497–507

Li Y et al (2017) Stereoscopic processing of crossed and uncrossed disparities in the human visual cortex. BMC Neurosci 18:80

Liu C et al (2018) Image categorization from functional magnetic resonance imaging using functional connectivity. J Neurosci Meth 309:71–80

Minini L et al (2010) Neural modulation by binocular disparity greatest in human dorsal visual stream. J Neurophysiol 104:169–178

Mizraji E et al (2017) The feeling of understanding: an exploration with neural models. Cogn Neurodyn 11:135–146

Naselaris T et al (2011) Encoding and decoding in fMRI. NeuroImage 56:400–410

Neri P et al (2004) Stereoscopic processing of absolute and relative disparity in human visual cortex. J Neurophysiol 92:1880–1891

Nienborg H et al (2006) Macaque V2 neurons, but not V1 neurons, show choice-related activity. J Neurosci 26:9567–9578

Norman KA et al (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci 10:424–430

Pantazatos SP et al (2012) Decoding unattended fearful faces with whole-brain correlations: an approach to identify condition-dependent large-scale functional connectivity. PLoS Comput Biol 8:e1002441

Parhizi B et al (2018) Decoding the different states of visual attention using functional and effective connectivity features in fMRI data. Cogn Neurodyn 12:157–170

Preston TJ et al (2008) Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. J Neurosci 28:11315–11327

Richiardi J et al (2011) Decoding brain states from fMRI connectivity graphs. NeuroImage 56:616–626

Shirer WR et al (2012) Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex 22:158–165

Smith SM et al (2011) Network modelling methods for FMRI. NeuroImage 54:875–891

Stevens WD et al (2015) Functional connectivity constrains the category-related organization of human ventral occipitotemporal cortex. Hum Brain Mapp 36:2187–2206

Tagliazucchi E et al (2012) Automatic sleep staging using fMRI functional connectivity data. NeuroImage 63:63–72

Tozzi A et al (2017) From abstract topology to real thermodynamic brain activity. Cogn Neurodyn 11:283–292

Wang X et al (2016) Representing object categories by connections: evidence from a mutivariate connectivity pattern classification approach. Hum Brain Mapp 37:3685–3697

Yamashita O et al (2008) Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. NeuroImage 42:1414–1429

Yan SM (1985) Digital stereoscopic test charts. People's Medical Publishing House

Yang Z et al (2014) Common intrinsic connectivity states among posteromedial cortex subdivisions: insights from analysis of temporal dynamics. NeuroImage 93(Pt 1):124–137