Classification Analysis of the Auditory Cognitive States for Sighted and Blind People Based on a Whole-brain "Searchlight" SVM

Jianlong Zheng and Junhai Xu*

School of Computer Science and technology, Tianjin Key laboratory of Conitive Computing and Application, Tianjin University, Tianjin 300072, P.R. China
*Corresponding author

Abstract—When talking about auditory perception, it remains unclear how people could achieve sound identification and localization efficiently even in the middle of chaos. Though the dual-pathway model is widely held, there are still controversies in specific brain regions. Moreover, some machine learning methods are successfully applied to study neural pattern information in our brain. In this study, both sighted and blind subjects were recruited and instructed to perform a sound localization and identification task. Functional magnetic resonance imaging (fMRI) data were collected, then we used a whole brain "searchlight" support vector machine (SVM) to decode the pattern information successfully. We identified the distinct cortical networks encoding sound pattern and spatial information, and we also found a large-scale cortical reorganization in blind people from a multi-voxel perspective.

Keywords-sound identification; sound localization; blind; functional magnetic resonance imaging; support vector machine

I. INTRODUCTION

Human beings can easily perceive outside world, such as sound categories and other spatial information, despite in a complex environment. Previous studies have showed that the distinct neural pathways were involved in sound identification and sound localization[1], which also suggested a dual-pathway processing mechanism for the auditory cortex, wherein sound identification and localization were processed in the so-called "what" (ventral) and "where" (dorsal) streams respectively. A further invasive study found that transcranial magnetic stimulation (TMS) to posterior auditory cortex (AC) showed a more significant influence in reaction times (RTs) during the localization task than sound identification task, while TMS to anterior areas delayed RTs significantly more during sound identification task[2]. But there were still many controversies for this clear anatomo-functional dissociation between two independent neural pathways, while another study suggested an interaction between the ventral and dorsal streams during auditory cognitive processing[3].

Blind people were found to have a more prominent performance in sound localization task than sighted. So exploring their cognitive mechanism during the auditory task may give us some novel hint. Furthermore, several studies found the brain regions activated by specific sound in blind subjects was significantly different with those in sighted ones, which was from a univariate perspective[4].

Traditional researches on auditory cognition were based on a univariate analysis, which studied voxel by voxel by using statistical methods to test if it was activated by specific experimental conditions. But this led to a loss of fine-grained pattern information across voxels though it had already achieved good success. Recently, the multi-variate pattern analysis (MVPA) became more and more popular in human neuroscience. MVPA is a data-driven method, and it can extract feature information from multi-dimensional space and separate data samples into different classes. In a specific region of interest, if the voxel pattern can be sorted correctly, it indicates that this brain region contains the stimulus information in the experiment. We aimed at exploring which voxels were involved in processing of sound pattern and spatial information using a support vector machine (SVM) classification algorithm, and making a comparison between the blind and sighted group.

II. EXPERIMENTAL PROCEDURE AND METHOD

A. Experimental Procedure

Our study included twenty healthy participants aged between 20 and 22 years old (10 males, 10 females; age = 21.4 ± 0.6 years old), together with sixteen blind subjects (10 males, 6 females; age = 21 ± 1.8 years). Five participants in the blind group have no light perception (NLP), and the remaining only have faint light perceptions, and all of blind matched to sighted ones on age and even on level of education.

The sound stimuli consisted of four types of conceptual categories: human, animal, environmental and mechanical sound, and all sounds were convolved with head related transfer functions (HRTFs) to simulate different spatial information. In addition, we also added white Gaussian noise to generate two different signal-to-noise ratios (SNRs). So our experiment included sixteen sound conditions: 4 categories × 2 directions (left 45° and right 45°) × 2 SNRs (clear and noise).

Our experiment adopted a block design including four task runs. Each run consisted of sixteen blocks corresponding to sound conditions mentioned above, and the fixation between two successive blocks lasted 10 seconds. At the beginning of each run, a 10-s prescan was also conducted to stabilize the MRI equipment. In each block, eight different exemplars that belonged to the same categories were presented in a pseudo-randomly order. Both sighted and blind participants were instructed to perform sound localization and identification task
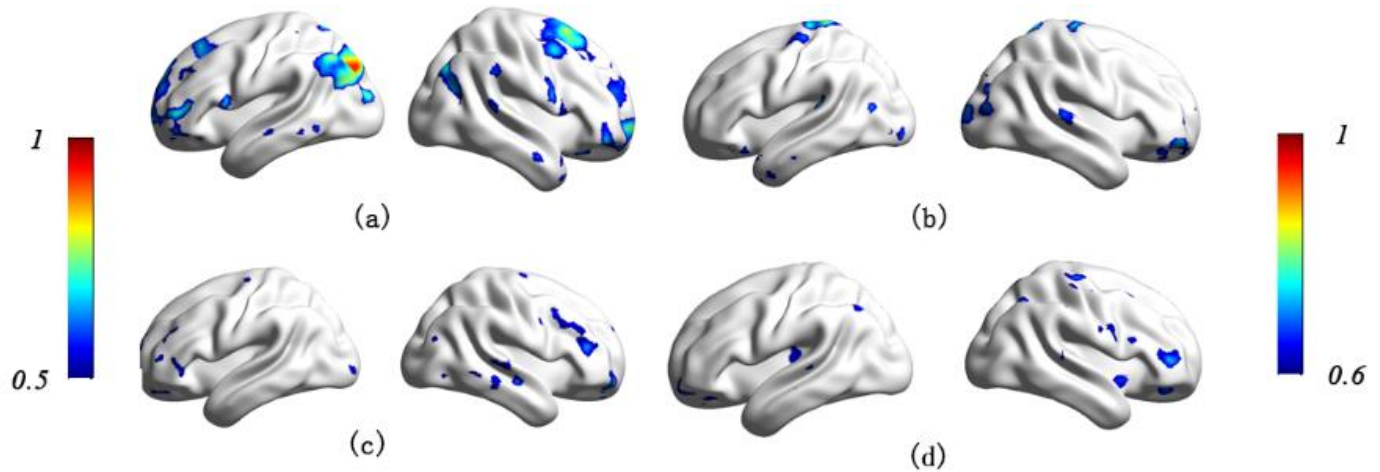
FIGURE I. THE SEARCHLIGHT RESULTS OF SIGHTED (ABOVE) AND BLIND GROUP (BELOW), THE LEFT IS 4-CATEGORY CLASSIFICATION RESULT, AND RIGHT IS 2-CATEGORY CLASSIFICATION RESULT.

simultaneously. MRI scanning was conducted on a 3.0 T Siemens scanner in Yantai Hospital Affiliated to Binzhou Medical University with a 32-channel head coil. The functional MRI data were acquired using an echo-planar image (EPI) sequence, and high-resolution magnetization-prepared rapid acquisition gradient echo structural image was also acquired for each subject at the end of task runs.

B. Data Processing

We used Statistical Parametric Mapping (SPM) toolbox to process functional data, including preprocess and first level statistical analysis. During preprocessing, functional images were slice-time corrected to the middle slice, and then a realigned procedure was done to correct for head movements. Finally, the parameters from a high-resolution structural image to Montreal Neurological Institute (MNI) space were used to normalize the functional images.

After preprocessing, follow-up general linear model (GLM) and statistical inference based on univoxel were done. For each voxel, the time sequence was extracted firstly. If task lasted T seconds and the TR of MRI equipment was 2 seconds, the number of samples of time sequence would be T/2. Then the data was submitted to GLM framework, with a high-pass filter of a standard cut-off frequency of 1/128 Hz, and the regressor in each condition was constructed by convolving the onset of stimulus with the canonical hemodynamic response function (HRF). The regression variables included sixteen experimental conditions and other conditions of no interest, such as six parameters of head motion (X, Y, Z translation and rotation). We assumed that there was a linear relation between the neural response and sound stimuli,

$$Y = \beta_1 * X_1 + \beta_2 * X_2 + \beta_{16} * X_{16} \cdots + \beta_n * X_n + \varepsilon_i.$$

For each voxel, the weights of all regression conditions would be estimated. In other words, we computed beta images corresponding to experimental conditions. Finally, the statistical inference was performed to test which voxels were sensitive to specific sound stimuli.

C. Algorithm

As one of the most popular multi-variate methods, SVM was usually used as pattern-classification algorithm to study neural representation. SVM mapped input features into a high dimensional feature space, and then classified the input data by constructing the optimum separating hyperplane. In the current study, we also chose SVM as classification methods because of our high-dimension and small-sample data set. A whole-brain searchlight analysis for each subject and second level group analysis were done in sequence, then we treated the voxels selected by searchlight analysis and GLM as regions of interest (ROI) and did a validation procedure.

In the searchlight analysis, a searchlight sphere with 6-mm radius was constructed for each voxel and its surrounding neighbors first, and then all samples were extracted from time series using a MVPA toolbox and labeled by corresponding conditions. Then SVM was used to classify these predefined labels. In order to avoid overfitting bias, a sixteen-folds cross-validation procedure was applied. We divided the data set into sixteen part randomly, and trained the model on fifteen parts, then tested our model on the remaining one. This procedure was repeated for sixteen times until all samples were be tested and the average of sixteen accuracies was treated as classification performance for a specific voxel. A searchlight map was generated for each subject, and then one-side Wilcoxon sign-rank test was conducted to examine which voxel could classify these sound conditions significantly.

Then we tested the effectiveness of the selected voxels, for sighted and blind group, we built the mask based on these significant voxels and activate pattern estimated by GLM respectively. Finally, we treated these voxels as ROI and ran SVM procedure to test them from a pattern perspective.

III. EXPERIMENTAL RESULTS

The searchlight results are shown in FIGURE I (a ~ d). In the second level group analysis, we did the statistical inference and computed which voxels could correctly classify the experimental conditions. We tested the accuracies of which regions were more than chance level significantly (4-categories

classification: 25%, 2-class: 50%). In the sighted group, the voxels that could classify four categories (human, animal, mechanical and environmental) were mainly located at bilateral precuneus, left inferior parietal lobe (IPL), right superior frontal gyrus (SFG), middle frontal gyrus (MFG), as shown in FIGURE I (a). Differently, the voxels that could classify direction conditions (left and right) were mainly in bilateral superior parietal lobe (SPL), right precuneus and right postcentral gyrus FIGURE I (b). For the blind group, the cortical network that contained conceptual category information included the lingual gyrus (occipital lobe), right MFG, right middle temporal gyrus (MTG) and bilateral middle frontal gyrus FIGURE I (c). In contrast, the voxels in left middle occipital gyrus (MOG), bilateral IPL, and left posterior cingulate contained sound spatial information FIGURE I (d). Only the voxels in which the classification accuracy reached above chance level were rendered on PALS template and displayed.
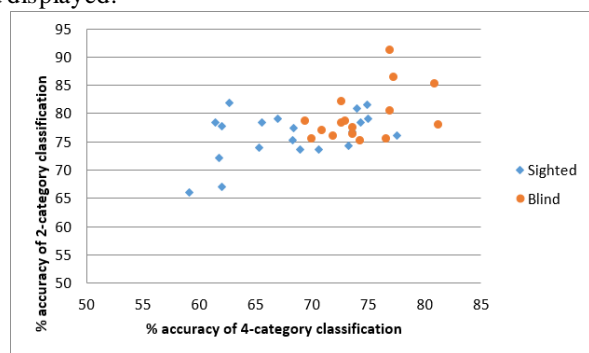

FIGURE II. ROI-BASED CLASSIFICATION ACCURACY IN SIGHTED AND BLIND GROUP

The validation result is shown in FIGURE II. This step is based on previous searchlight analyses, and we included the voxels which are tested significantly classify the experimental conditions and also activated in GLM group level in our ROIs. Then the response pattern in this ROI was submitted to SVM classification framework, 4-category (4 conceptual categories) and 2-category classification (2 sound directions) were done respectively. FIGURE II showed the average of cross-validation accuracy, plotted by 4-category classification accuracy on the horizontal axis and 2-category classification on the vertical. Blue and yellow points indicated sighted and blind subjects respectively.

IV. DISCUSSIONS

Our results showed that the distinct cortical networks were involved in sound identification and localization. Instead of a large part of ventral regions containing sound pattern information, the voxels those had good performances on sorting sound spatial condition were mainly located at the dorsal stream. Bind group also showed separated cortical networks for different auditory cognitive tasks to some degree, suggesting that visual experiences may have no influence in this anatomo-functional dissociation. This was consistent with one previous study[5]. However, there were many voxels that could both significantly classify sound category and direction, which suggested the two neural pathway may have some interaction there. This hypothesis have been proposed by a previous study [6]. Besides,

we further identified the cortical networks that contained sound pattern and spatial information in the whole brain from a multivariate perspective.

The blind subjects showed some differences with sighted group. The second level result of searchlight analysis included a large regions of occipital lobe, which was the traditional visual cortex, and this cross-modal processing for auditory task may explain the highly developed auditory abilities of blind people. Moreover, the ROI-based validation procedure demonstrated the voxels those selected during a whole brain searchlight analysis, as a pattern, could classify our experimental conditions effectively. The scatter plot showed blue points located on the top right corner of yellow points, which also indicated that blind subjects had better classification performance than sighted ones.

The SVM was applied to classify sound conditions successfully. Its good performance may be caused by the high dimension of fMRI data set, and one brain volume image included hundreds of thousand voxels. SVM algorithm computed support vectors (SV), and only these SV had influences in the classifying performances. Moreover, we can further test the different performances of linear and non-linear kernels.

REFERENCES

[1] Rauschecker, J.P., Processing of Complex Sounds in the Auditory Cortex of Cat, Monkey, and Man. Acta Oto-Laryngologica, 1997. 532(Supp 532): p. 34-38.

[2] Ahveninen, J., et al., Evidence for distinct human auditory cortex regions for sound location versus identity processing. Nature Communications, 2013. 4(10): p. 2585.

[3] Cloutman, L.L., Interaction between dorsal and ventral processing streams: where, when and how? Brain Lang, 2013. 127(2): p. 251-63.

[4] Occelli, V., C. Spence, and M. Zampini, Auditory, tactile, and audiotactile information processing following visual deprivation. Psychological Bulletin, 2013. 139(1): p. 189-212.

[5] Zundorf, I.C., J. Lewald, and H.O. Karnath, Testing the dual-pathway model for auditory processing in human cortex. Neuroimage, 2016. 124(Pt A): p. 672-81.

[6] Jörg, L. and G. Stephan, When and Where of Auditory Spatial Processing in Cortex: A Novel Approach Using Electrotomography. Plos One, 2011. 6(9): p. e25146.