Decoding auditory frequencies and directions based on brain functional features

Mingxi Wang and Gaoyan Zhang*
College of Intelligence and Computing, Tianjin key Laboratory of Cognitive Computing and Application, Tianjin University, China
*E-mail: zhanggaoyan@tju.edu.cn

***Abstract*— Decoding auditory stimuli based on brain function data is of great significance to understand auditory functional mechanism. At present, there is still controversy about whether the brain processing mechanism of auditory stimuli is parallel hierarchical processing or distributed processing [1]. Different from previous studies that used univariate analysis to study auditory processing, this study intends to build a decoding model and study the auditory processing mechanism from the perspective of multivariate pattern analysis. In this paper, we analyzed functional MRI data from 27 subjects under perception of different auditory frequencies and directions, using brain activation and functional connectivity as features, and using support vector machine for decoding. The decoding accuracy of frequencies and directions was 70.7% and 71.6% with brain activation features. On the other hand, the accuracy rate reached 73.7% and 77.7% respectively with functional connectivity features. Then we analyzed the weights and found that the activation patterns in precuneus and the superior temporal gyrus (STG) contributed to sound frequency discrimination, and STG also represented differences in direction. The connectivity patterns between the bilateral precuneus showed obvious changes under different frequency conditions, while the bilateral middle occipital gyrus and STG showed significant changes under different directions of sound stimulation. The results support a distributed auditory processing model.**

I. Introduction

Hearing is an important way for human beings to know the world and obtain external information. The information acquired by the auditory pathway is very important for people to understand surrounding environment.

Similar to the dual-pathway model in visual information processing, researchers believe that auditory frequency information and auditory orientation information are processed through different pathways [2]. That is, the information processing related to the recognition of sound type is carried out in the ventral "What" pathway, while the information processing related to the recognition of sound orientation is carried out in the dorsal "where" pathway. The two pathways are parallel and respectively process the sound information. Although the auditory dual-pathway model has been supported and verified in many relevant literatures [3-6], there are still many studies questioning this model. On the one hand, many studies believe that the two pathways in this model are not completely independent of each other and deal with sound information separately. In some cases, the two pathways interact with each other [7]. On the other hand, some studies proposed that the brain regions corresponding to sound recognition and spatial location were widely distributed in the cortex [8-10].

As we all know, auditory information processing in the human brain can be studied though a forward encoding way and also by a backward decoding way. Most previous studies revealed the processing mechanism of stimulus information through univariate analysis from the perspective of encoding [11, 13]. For example, Okada et al. used univariate analysis on neural mechanism of speech stimuli and proposed a hierarchical organization of human auditory cortex in an encoding way [11]. By functional integration analysis, researchers found a distributive functional connectivity pattern for auditory direction processing during encoding [13]. In recent years, the multivariate pattern analysis (MVPA) method based on decoding model, has a more sensitive detection ability to reveal the spatial pattern of brain regions and their information interaction pattern during responding to stimuli [12, 13]. Some researchers used multivariate pattern analysis to decode auditory stimuli and found that besides auditory regions, visual regions can also represent auditory semantic information, suggesting a distributive processing mechanism underlying auditory processing [12]. Considering auditory frequencies and directions are basic components in auditory stimuli and they separately occupy the two auditory pathways, in this study, we aimed to used MVPA based decoding method to study responding pattern of auditory frequencies and directions in the human brain more sensitively from the way of decoding and further investigate the auditory processing mechanism.

II. EXPERIMENT AND ANALYSIS

A. *Experiment*

The experimental materials include functional imaging data and T1 structural image data. The block experiment paradigm was adopted in this experiment. The whole experiment includes direction task and frequency task, and each task includes 3 function runs. Each run lasts 544s, which starts with a pre-scan of 8s, followed by 18 blocks, and ends with 8s of rest. Each block lasts for 18s and presents sounds in the same direction or frequency. Blocks are spaced 12s apart. A block contains nine trials, each consisting of a 1s stimulus and a 1s rest. In this experiment, the stimulus materials included sounds of high, medium and low frequencies and sounds of left, middle and right directions. The experimental paradigm is shown in Fig. 1.

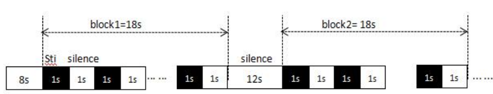

Fig.1 Experimental paradigm design

B. *Data collection*

All the imaging data is collected in a 3.0 T Siements Tim Trio MRI scanner. The experiment recruited 28 healthy College students. (Data from one of the subjects was not available and was removed from subsequent experiments) There were 14 male and 14 female, with an average age of 22.3 years (SD=1.1). All subjects were born right-handed, had normal hearing and had no mental or neurological problems. Foam pads and earplugs were used for all participants; besides that, eyeshade was worn to prevent vision effect. T1-weighted anatomical images were acquired with a three-dimensional magnetization-prepared rapid acquisition gradient echo (3D MPRAGE) sequence with the following parameters: TR = 1900 ms, TE = 2.52 ms, TI = 1100 ms, voxel size = 1 × 1 × 1 mm^3, matrix size = 256 × 256. T2*-weighted images were acquired using a gradient echoplanar imaging (EPI) sequence with the follow parameters: TR = 2000 ms, TE = 30 ms, FOV = 192 × 192mm^2, matrix size = 64 × 64, slices = 33, slices thickness = 4 mm, slice gap = 0.6 mm.

C. *Data preprocessing*

Unified processing of fMRI data is required prior to analysis. First, data format was converted into NII format. Then, the first 8 seconds of fMRI data in each run were removed because they were dummy scan. SPM8 toolbox (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) was employed to preprocess the structural and functional images in the following steps: (1) time slice correction, (2) headmotion correction, (3) coregistration of functional images with structural images, (4) structural images segmentation, (5) spatial normalization to make sure each subject in the same MNI space, (6) spatial smoothing to improve signal-to-noise ratio. Through these steps of data preprocessing, the errors caused by the physiological characteristics of the subjects as well as the errors generated in the data collection process can be reduced as far as possible.

D. *Decoding based on brain activation feature*

Firstly, the activity intensity of each voxel was used as feature when building the decoding model on sound frequencies and directions. The preprocessed 4-D data were converted into a two-dimensional matrix with each row represents the stimuli samples and each column is the spatial voxels in the brain. As there is a certain delay between the generation of stimulus signals and the hemodynamic response, the relative positions of all samples and labels should be moved back 4s as a whole. The excess data from the sample, such as the rest time between stimuli, were removed.

In this experiment, F-score was used to select features. F-score is a method to measure the distinguishing ability of different spatial voxels between two categories, which can effectively realize feature selection. The formula is as follows:

$$F(\mathrm{i}) \equiv \frac{(\overline{\mathrm{x}}_{\mathrm{i}}^{(+)} - \overline{\mathrm{x}}_{\mathrm{i}})^2 + (\overline{\mathrm{x}}_{\mathrm{i}}^{(-)} - \overline{\mathrm{x}}_{\mathrm{i}})^2}{\frac{1}{n_+ - 1}\sum_{k=1}^{n_+}(x_{k,i}^{(+)} - \overline{x}_i^{(+)})^2 + \frac{1}{n_- - 1}\sum_{k=1}^{n_-}(x_{k,i}^{(-)} - \overline{x}_i^{(-)})^2}$$

The higher the F-score, the better the discrimination. After calculating the F value of all features, the top 200 features with the highest F value are selected for the final training and testing. The support vector machine (SVM) was used as classifier. The k-fold cross-validation strategy is adopted. The data of 27 subjects were divided into 9 groups on average. For each training, 8 groups of data were used as the training set, and the remaining 1 group was used as the test set. This process was repeated for 9 times.

E. Decoding based on functional connectivity features

A brain template (AAL) with 116 nodes were used and time series of each node was extracted from the preprocessing fMRI data. The whole brain functional connection matrix for each subject was constructed by Pearson correlation of time series in each brain region node, which represents the functional connection mode of a subject under a category condition. Before classification, the functional connection matrix needs to go through a series of processes. The first step is to eliminate redundant information from the functional connection matrix. In this experiment, the data of lower triangular matrix is retained and represented by one-dimensional vector. The vectors of all subjects are formed into two matrices corresponding to the direction and frequency. Each row of the matrix represents a category. In addition, according to previous research results, the internal mechanism of negative functional connections is not clear so far [14, 15]. So we removed the negative connections. Next, the calculated F value is used to select 200 connections that differ significantly between different categories. Similar classifier and cross-validation frame was used as in the aforementioned section.

III. Result

The classification results are shown in the table below. Higher classification accuracy is obtained based on functional connectivity features.

Table 1 Classification results based on brain activation and functional connectivity features

	Brain activation features	**Brain connectivity features**
Frequency decoding	70.7%	73.7%
Direction decoding	71.6%	77.7%

The weights of SVM are read to get the most representative support vectors in the sa-mple. The BrainNet Viewer (http://www.nitrc.org/projects/bnv/) is used to draw weight maps.

According to the distribution of weights in SVM shown in Fig. 2 and Fig. 3, we can see activation strengths in superior temporal gyrus and cuneus were significantly contributive in classify different frequencies of stimulation. For the different direction of the sound categories, the significant contribution of the region is mainly concentrated in the superior temporal gyrus. Other regions in frontal cortex also contribute to the auditory decoding based on brain activations.

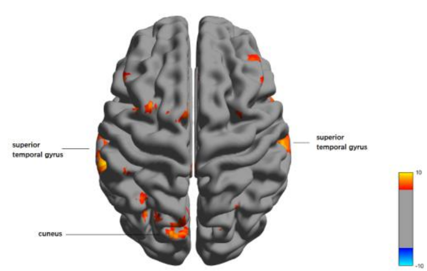

Fig.2 Brain activation features that contributed to auditory frequency decoding (cluster size>20, p<0.001)

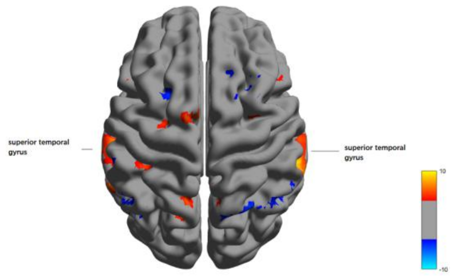

Fig.3 Brain activation features that contributed to direction decoding (cluster size>20, p<0.001)

Figure 4 and figure 5 show the functional connections that have significant contributions to differentiating sounds of different frequencies and directions. Each subject had 13456 functional connections under each condition. We chose the top 200 distinctive connections as features in auditory decoding. The results show that the cuneus and precuneus have obvious weights in discrimination of different frequency conditions. The more significant connections are those between bilateral precuneus and bilateral cuneus and their connection to left superior temporal gyrus, middle occipital gyrus, and right superior frontal gyrus. Under the directional condition, bilateral middle occipital gyrus and left superior temporal gyrus showed significant contributions. The contributing connections are primarily distributed in the connections between left superior temporal gyrus and middle occipital gyrus and their connections to paracentral lobe, cuneus and precuneus.

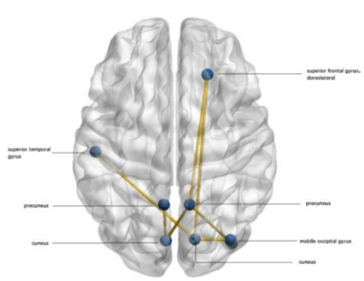

Fig.4 Function connection features that contributed to auditory frequency decoding

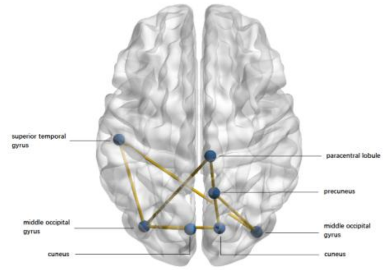

Fig.5 Function connection features that contributed to auditory direction decoding

IV. Summary and Prospect

Previous auditory studies were mostly carried out to sp ecify the regions, and most of them are based on univariate analysis of stimuli encoding process. The univariate analysis hypothesizes that each voxel in the brain is independent. However, more and more evidence shows that the cognitive behavior of human brain usually requires the cooperative participation of multiple brain regions. Studies based on individual brain regions fail to identify the intrinsic connections between brain regions. In order to fully explore the information interaction of multiple brain regions under auditory stimulation, this experiment studied the processing mechanism of auditory information in the brain from activation characteristics and connection characteristics, respectively. We used a decoding model to study the multi-variate representative pattern of auditory stimuli, which can more sensitively detect the recruited activation regions and connection patterns during auditory processing compared with univariate encoding analysis.

Based on the results of statistical analysis and functional connection analysis, two decoding models of different frequencies and different directions were constructed by using lib-SVM. Significant decoding accuracies were obtained, suggesting the effectiveness of decoding models. In decoding research based on brain activation, the results showed that the activation of the superior temporal gyrus and the cuneus region was the main change of different frequency information. Compared with the activation results based on encoding analysis [17], it was found that the cuneus was also significantly activated under this condition. Results from different directions also showed significant effects on the superior temporal gyrus region. In the decoding research based on connection characteristics, for frequency conditions, three significant regions were obtained, namely, the cuneus, the precuneus and the superior temporal gyrus. The region of the superior temporal gyrus was found in addition to connection results based on coding analysis. For directional conditions, the regions with significant changes include the middle occipital gyrus, the cuneus and the superior temporal gyrus. From these contributed regions and connections in auditory decoding, we can see they not only include regions in the auditory dual-pathway but also recruit regions outside the dual-pathway, demonstrating that the auditory stimuli may be represented in a more distributive way.

In future research, the following questions can be further discussed. First of all, the processing patterns of different kinds of sounds by human brain can be further discussed, so as to find more characteristics and rules [16]. Secondly, there is noise in fMRI collection of auditory signals, which may affect the decoding results. In the future research, sparse acquisition can be considered to improve the SNR. Finally, we can study the temporal process of auditory information processing. High time-resolution acquisition technology combined with dynamic analysis method can be considered in the future study.

ACKNOWLEDGEMENT

This study was supported by the National Nature Science Foundation of China (No. 61503278).

REFERENCES

[1] Bizley JK,Cohen YE.The what, where and ho-w of auditor y-object perception.Nat Rev Neu-rosci 14(10):693-707.

[2] Ida C.Zündorf,Jörg Lewald,Hans-Otto Karnath.Testing the d ual-pathway model for auditory processing in human cortex [J]. Neuroimage,2016,124(Pt A):672-681.

[3] J. Fridriksson J,Yourganov G,Bonilha L,et al.Revealing the dual streams of speech processi-ng[J].Proceedings of the Na tional Academy o-f Sciences,2016,113(52):15108-15113.

[4] Arnott S R,Binns M A,Grady C L, et al.Asse-ssing the au ditory dual-pathway model in hu-mans[J].Neuroimage,2004,2 2(1):401-408.

[5] Clarke S,Thiran A B,Maeder P,et al.What andWhere in hu man audition: selective deficits fo-llowing focal hemispheric lesions[J].Experime-ntal Brain Research,2002,147(1):8-15.

[6] Alain C,Arnott S R,Hevenor S,et al."What" a-nd "where" i n the human auditory system[J].Proc Natl Acad Sci USA, 2001,98(21):12301-12306.

[7] Bizley J K,Cohen Y E.The what, where and how of audito ry-object perception[J].Nature R-eviews Neuroscience,2013,1 4(10):693-707.

[8] Furukawa S.Coding of sound-source location by ensembles of cortical neurons[J].J.Neurosci.2000（20）.

[9] J. C. Middlebrooks,A. Clock,L. Xu,et al.A pa-noramic code for sound location by cortical n-eurons[J].Science.1994（26 4）:842-844.

[10] Zimmer U,Lewald J,Erb M,et al.Processing ofauditory spati al cues in human cortex: An fM-RI study[J].Neuropsycholo gia,2006(3):454-461.

[11] Okada K, Rong F, Venezia J, Matchin, Hsieh IH, Saberi K, Serences JT, Hickok G. (2010): Hierachical Organizatio n of Human Audito-ry Cortex: Evidence from Acoustic Inv arianc-e in the Response to Intelligible Speech. Cere-bral C ortex 20(10): 2486-2495.

[12] Vetter P, Smith FW, Muckli L. (2014): Deco-ding sound a nd imagery content in early vis-ual cortex: Current Biology 24(11):1256-1262.

[13] Haynes JD,Rees G.Decoding mental states fro-m brain activ ity in humans.Nat Rev Neurosci- 7(7):523-534.

[14] Fox M D,Zhang D,Ssnyder,et al.The Global S-ingal and O bserved Anticorrelated Resting State Brain Networks. Jou rnal of Neurophysio-logy[J],2009,101（6）:3270-3283.

[15] Tal Z,Geva R,Amedi A.The Origins of Meta-modality in V isual Object Area Lo:Bodily To-pographical Biases and Inc reased Functional Connectivity to SI.Neuroimage[J],2016,12 7:363-375.

[16] Zhou Y, Xiang H, Chen J Y, et al. Functiona-l magnetic re sonance imaging of brain durin-g Stroop task in adolescent s with online gami-ng disorder [J]. Chin J Psychiatry, 2008, 51(5):329-334.

[17] Liang Yaping. Study on the processing mecha-nism of soun d in human brain [D]. Tianjin: Tianjin University,2019.

[18] Jinliang Zhang,Gaoyan Zhang,Xianglin Li,et al.Decoding so und categories based on whole-br-ain functional connectivit y patterns[J].Brain I-maging and Behavior,2018.

[19] Zhang haimin, Chen shengzu. A new methodof brain functi on imaging: statistical parametri-c graph (SPM). Chin med imaging technolog-y,2002,18 (7) :711-3.

[20] Lei Wei, Yang Zhi, Zhan Minye et al. Neura-l characteriza tion of cognitive decoding usingbrain imaging multi-voxel model: Principles a-nd applications. Advances in Psychologi cal S-cience, (12) :1934-41.

[21] Xiang jie, Chen junjie. SVM based fMRI dat-a classificatio n: a method for decoding thinki-ng. Computer research and development,2010,47 (2) :286-291.

[22] Francisco Pereira,Tom Mitchell,Matthew Botvi-nick.Machine learning classifiers and fMRI:Atutorial overview.NeuroImag e,2009,45（2009）:199-209.

[23] Li Hang. Statistical methods [D]. Beijing: Tsi-nghua Univer sity Press,2012.