



Phase Synchronization Indices for Classification of Action Intention Understanding Based on EEG Signals

Xingliang Xiong¹, Xuesong Lu², Lingyun Gu¹, Hongfang Han¹,
Zhongxian Hong², and Haixian Wang¹(✉)

¹ Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, Jiangsu, People's Republic of China

hxwang@seu.edu.cn

² Department of Rehabilitation, Zhongda Hospital, Southeast University, Nanjing 210009, Jiangsu, People's Republic of China

Abstract. The classification of action intention understanding based on EEG signals is very important for human-robot and social interaction studies. In order to classify the action intention understanding brain signals efficiently, we first use three kinds of phase synchronization indices, phase locking value (PLV), phase lag index (PLI) and weight phase lag index (WPLI), to construct functional connectivity matrices in multiple micro time windows, and then extract the sum of significant edge values of each time window matrix as the classification feature, finally apply support vector machine (SVM) classifier to implement action intention understanding data classification task. Classification result shows that new method performs well on three datasets (alpha, beta and fusion frequency bands), and brain network statistical analysis demonstrates that many significant edges appear on the alpha frequency band. We conclude that the phase synchronization indices are extremely useful for the classification task, the sum of significant edge values is an effective classification feature, and the action intention understanding closely correlates with the alpha frequency band.

Keywords: Phase synchronization · Classification · Action intention understanding

1 Introduction

In recent years, the study of action intention understanding has attracted extensive attention [1–6]. The classification of action intention understanding based on EEG signals is one of the most important branches of the study, which is viewed as a key factor for human-robot interaction [1, 7–9]. Many researchers carried out a lot of experiments on the action intention understanding classification by different methods [1, 5, 9]. However, the classification accuracy is often

unsatisfactory. This is due to two important reasons: (1) it is difficult to extract the most useful classification features, (2) it is hard to collect a large number of training samples. As for the former question, some researchers use different brain signal collection techniques to solve it [1, 9], and some other researchers mainly focus on extracting further feature [5]. Due to the popularity of deep learning, many people begin to consider using neural network (e.g., convolution neural network (CNN) and recurrent neural network (RNN)) technique to extract features. As for the latter question, people usually spend much time and money to recruit subjects and collect a certain amount of neural information data.

Brain network is an efficient tool to study neuroscience, which has some comprehensive merits (e.g., feature extraction, brain region position) [1, 5]. There are a lot of algorithms to construct the brain network [10]. The synchronous oscillation of a neural network is the main potential mechanism of brain information integration and processing, and the synchronization of multiband signals is the key feature of information exchange between different brain regions [11]. In the EEG synchronization analysis algorithms, the phase synchronization analysis method can directly separate the signal phase information in a given frequency range from the amplitude of nonstationary information. Then, the signal phase information can be used for the synchronization analysis of EEG narrowband signals, such as mu rhythm [11–13]. It is noteworthy that some previous studies point out the action intention understanding correlates with the alpha and beta rhythms [14, 15]. The advantage of phase synchronization is that it has nothing to do with the amplitudes of two neural oscillatory activities, but only with the phases.

Considering that, we first use three phase synchronization indices, phase locking value (PLV) [11], phase lag index (PLI) [12] and weighted phase lag index (WPLI) [13], to construct brain networks in this study, then apply these brain networks to solve the former two problems (feature extraction and sample collection). As for the first problem (feature extraction), we adopt t-test and FDR correction to select the edges of a functional connectivity matrix and sum for these edges as a feature. And as for the second problem (sample collection), inspired by the literature [16], we use an idea that one divides into three to model more samples. Each subject has three kinds of brain networks that are calculated by the PLI, WPLI and PLV under a certain stimulus condition. We view the different brain networks as different samples. Because our final goal is to classify action intention understanding, the sample model method is feasible. More details about how to solve these two problems are presented out in next section.

The main contribution of this study is that we design a novel method which can effectively solve the problem in feature selection for classifying action intention understanding EEG signals. Many other neural information data classification studies (e.g., epilepsy, emotion, and mathematical genius classification tasks) also can draw lessons from the novel feature extraction method. The training sample collection idea based on the phase synchronization indices shows its advantage under the condition of limited manual sampling. With the analyses of

signals classification and brain network statistical test, we found that the alpha frequency band easily obtains significant achievements in experiments, which further supports the conclusion that action intention understanding closely correlates with the alpha frequency band in some previous studies [5, 14, 15].

2 Materials and Method

2.1 Subjects

After deleting 5 subjects of which EEG data were seriously noised (e.g., someone has abnormal channels that there were no signals in them), we totally retained 25 healthy subjects (17 males, 8 females; aged 19–25 years, mean \pm SD: 22.96 ± 1.54 ; right-handed). This research was approved by the Academic Committee of the School of Biological Sciences and Medical Engineering, Southeast University, China.

2.2 Experimental Paradigm

In the progress of EEG data acquisition, all subjects were asked to see three kinds of hand-cup interaction pictures that were performed by an actor. They only needed to judge the actor’s intention, but not to implement any concrete operations. The three action intentions were drinking water (Ug), moving the cup (Tg) and simply contact the cup (Sc). Figure 1 shows the experimental stimuli and procedure. This design comes from Ortigue et al.’s experimental paradigm [17].

2.3 Data Collection and Preprocessing

In this study, we use 64 AgCl electrodes that were arranged with the international 10–20 system to record the EEG signals. The sampling rate was set to 500 Hz, and the reference electrode was set as M1 that was placed on the left mastoid. All the data collection tasks were completed with the Neuroscan 4.3.

In order to obtain useful data, we clean the raw EEG signals by Neuroscan 4.3 and EEGLAB 14.0 [18]. In the light of some previous data preprocessing experiences, it can’t efficiently clear noises of the raw EEG data by independent component analysis (ICA). We adopted ocular processing in the Neuroscan to replace the ICA in the EEGLAB. Mastoid reference is efficient in somatosensory evoked potentials, which correlates with action behaviors. Thereupon, we transformed the unilateral mastoid reference (M1) into bilateral mastoid (M1, M2) re-reference. Both the ocular processing and re-reference are completed with the Neuroscan.

When completed the ocular processing and re-reference, we used the EEGLAB to select 60 scalp electrodes that cover frontal, partial, central, occipital and temporal areas. Then, we adopted the Basic FIR filter in the EEGLAB to extract the 1–30 Hz data. And then, we segmented the data with event types

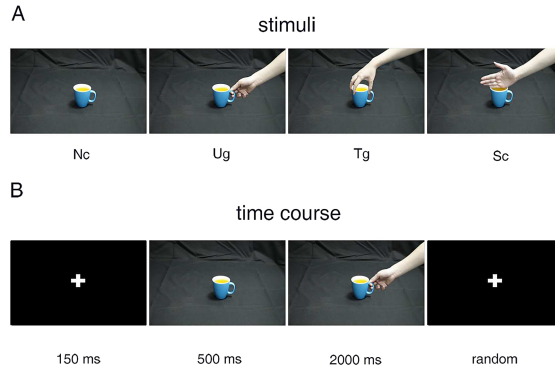


Fig. 1. Experimental stimuli. (A) An example of stimulus materials. Ug (use grip), Tg (transport grip), and Sc (simple contact) denote that the actor grasps the cup for drinking water, moving it, and touching it without any clear intention, respectively. (B) An example of the experimental procedure in a trial. In the stimuli, a symbol ‘+’ was first presented on the screen with 150 ms. Then, a cup was shown for 500 ms to keep consistent sense. The formal action intention stimulus that was denoted by a hand-cup interaction picture starts after 650 ms, which was sustained for 2000 ms. When the hand-cup interaction stimulus appears, the subjects need to guess what the actor want to do immediately. Before the next trial, the ‘+’ was presented again with a random time that varied from 1000 to 2000 ms.

in a time window (-0.65 s to 2.5 s) and removed the baseline by setting the baseline at -0.65 s. In the end, we deleted artifacts with a threshold range that varied from -75 to $75 \mu\text{V}$. A total of 679 trials were removed and an average of 267 trials were kept for each subject. It is noteworthy that the alpha and beta frequency sub-bands used in this study were extracted by low resolution electromagnetic tomography (LORETA) in the source space.

2.4 Phase Synchronization Indices

The advantage of phase synchronization is that it has nothing to do with the amplitudes of any two neural oscillatory activities, but only with the phases. Considering that, we use three phase synchronization indices to construct the functional connectivity matrices. Before giving out the concrete formulas of the three indices, it needs to explain two important concepts in this study. One is that each brain network is represented by a functional connectivity matrix. Another is that the node in the brain network is defined by the region of interest (RIO) in the source space, i.e., each node corresponds to a RIO. Our experiments mainly based on the whole brain, which has 84 RIOs that are defined in the LORETA. The preprocessed 60 channels EEG data were converted into 84 RIO time series with the LORETA. Hence, the size of a functional connectivity matrix (brain network) is 84×84 . The mathematical expressions of the three algorithms that are used to construct the functional connectivity matrices are defined as follows:

- Phase locking value (PLV)

$$PLV_{xy} = \frac{1}{N} \left| \sum_{n=1}^N e^{j\{\Phi_{n,x}(t) - \Phi_{n,y}(t)\}} \right| \quad (1)$$

where t is the time point, N is the total number of sample points, $\Phi_{n,x}(t)$ and $\Phi_{n,y}(t)$ are two instantaneous phases that come from channel x and channel y at the n th time point respectively. The instantaneous phases are computed by Hilbert transform.

- Phase lag index (PLI)

$$PLI_{xy} = \frac{1}{N} \left| \sum_{n=1}^N \text{sign}(\sin(\Delta\Theta(t_n))) \right| \quad (2)$$

where $\Delta\Theta$ is also the instantaneous phase difference that is between the time series $x(t)$ and $y(t)$ at the n th sample time point as demonstrated in the Eq. (1).

- Weight phase lag index (WPLI)

$$WPLI_{xy} = \frac{\left| \langle \tilde{S}(w) \rangle \right|}{\left| \langle |\tilde{S}(w)| \rangle \right|} = \frac{\left| \langle |\tilde{S}(w)| \text{sign}(\tilde{S}(w)) \rangle \right|}{\left| \langle |\tilde{S}(w)| \rangle \right|} \quad (3)$$

where $\tilde{S}(w)$ denotes the imaginary component of cross-spectrum between time series $x(t)$ and $y(t)$. The symbols $\langle \cdot \rangle$, $|\cdot|$ and function sign denote mean, absolute, signum function, respectively.

After obtaining the functional connectivity matrices that are calculated by PLV, PLI and WPLI, we first use paired t-test and false discovery rate (FDR) correction ($p < 0.05$) to find the positions of which edges are significantly different between two kinds of action intention understandings (Ug-vs-Tg, Ug-vs-Sc and Tg-vs-Sc). The final positions are determined by the edges that are all significantly different on the three paired comparisons. Then, we use the final positions to select the weighted edges in each time window. Finally, we use the sum of the selected edges in each time window as the classification feature, i.e., each time window matrix is corresponding to one feature. Because there are 63 dynamical time windows in this study (We divide the full-length time series into 63 sub time series, each has a length of 50 ms.), hence, it totally has 63 features for a single frequency band dataset. Figure 2 shows out the flow chart of our novel method. It is important to note that the feature selection is completed on the condition of 5-fold cross validation.

As for the problem of increasing data samples that is mentioned in the introduction, we let the feature vectors which are from the three functional connectivity matrices (PLV, PLI and WPLI) of one subject be three samples. Therefore, we can make 75 data samples for the original 25 subjects on one kind of action intention stimulus. A total of 225 data samples (Ug, Tg, and Sc all have 75 samples) are collected in this experiment.

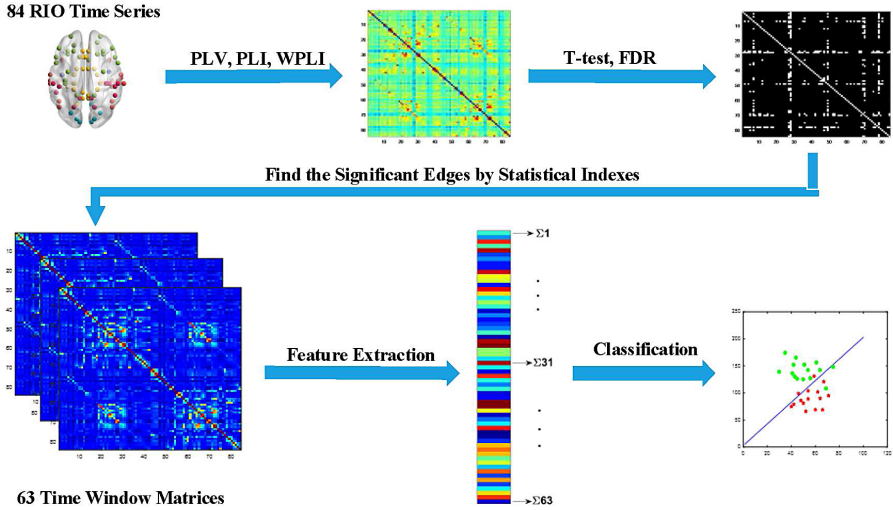


Fig. 2. The flow chart of new method.

3 Results

The experimental results mainly include two parts: action intention understanding classification and brain network statistics. In the following, we give out the demonstrations of the results.

3.1 Action Intention Understanding Classification

The classification pattern is one-versus-one, i.e., Ug-vs-Tg, Ug-vs-Sc and Tg-vs-Sc. The classifier is the classical support vector machine (SVM), of which kernel function and order number parameters are set as polynomial, 1, respectively. We use three datasets, alpha, beta and fusion, to carry out the action intention understanding classification. That is to say, the action intention understanding based on EEG signals come from the alpha, beta frequency sub-bands. Fusion dataset is constructed by merging the features of alpha and beta datasets into a big dataset. The classification tasks based on the group level, not a single subject. In order to avoid random factors, we implemented 1000 times 5-fold cross validation and calculated the mean of these 1000 times experiments as the final classification accuracy.

Figure 3 shows the average classification accuracies on the alpha, beta and fusion datasets. We can see that the lowest accuracy is over 65%, while the highest accuracy is even over 95%. The classification accuracies on the Ug-vs-Tg are all more than 85%, which performs the best compared with the other two one-versus-one patterns. Among the three datasets, beta performs the worst and

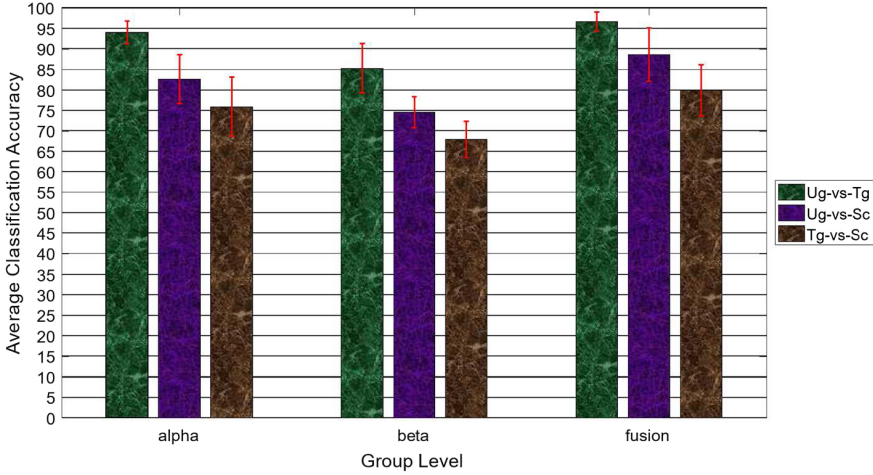


Fig. 3. Average classification accuracies on three datasets. The red stick denotes the standard deviation. The alpha, beta and fusion are three different datasets, which come from the alpha, beta and fusion frequency bands, respectively. (Color figure online)

fusion performs the best. Especially, the average classification accuracies on the fusion dataset are all greater than or equal to 80%. Table 1 displays the details of four classification estimation metrics under different conditions. We can see that most of the numerical values of mean (average classification accuracy), sensitivity and specificity are very high, while the numerical values of standard deviation are very low.

3.2 Brain Network Statistics

In this part, we give out the edge value differences of the average brain network that is based on 63 dynamical time windows. Each functional connectivity edge is tested by serious t-test and FDR correction. The statistics are carried out on the three one-versus-one patterns.

Figure 4 shows the statistical results on the alpha and beta frequency subbands. We can see that there are many significant connectivity edges in the alpha band, while the beta band is very sparse, especial in both Ug-vs-Sc and Tg-vs-Sc. Additionally, we also can see that many important nodes appear in the temporal, frontal and occipital lobes. The positions of the important nodes are uncommon on different frequency bands and one-versus-one patterns.

Table 1. Classification estimation metrics. The mean denotes the average classification accuracy.

Condition		Estimation metrics			
Frequency band	Pattern	Mean	Standard deviation	Sensitivity	Specificity
Alpha	Ug-vs-Tg	94.00	2.79	93.33	94.67
	Ug-vs-Sc	82.67	5.96	76.00	89.33
	Tg-vs-Sc	76.00	7.23	61.33	90.67
Beta	Ug-vs-Tg	85.33	6.06	89.33	81.33
	Ug-vs-Sc	74.67	3.8	66.67	82.67
	Tg-vs-Sc	68.00	4.47	44.00	92.00
Fusion	Ug-vs-Tg	96.67	2.36	98.67	94.67
	Ug-vs-Sc	88.67	6.50	81.33	96.00
	Tg-vs-Sc	80.00	6.24	64.00	96.00

4 Discussion

From the experimental results of action intention understanding classification and brain network statistics, we have some important findings. In this section, we present out the elaborations of these findings.

The results of classification accuracies display the novel method performs well on the alpha, beta and fusion datasets (see Fig. 3 and Table 1), which suggest that the feature extraction and sample collection that are dealt with by the phase synchronization indices are effective and satisfactory. Some previous studies [1, 5] which use a single index of phase synchronization (PLV, PLI or WPLI) to decode action intention understanding point out that the phase synchronization is a useful tool. The experimental results based on multiple indices of phase synchronization (PLV, PLI and WPLI) further prove the previous viewpoint. Both the alpha and beta frequency sub-bands show extremely high classification accuracies, which indicate that the action intention understanding brain activities easily occur at these two bands. These are consistent with some previous studies [5, 15]. Figure 3 demonstrates that the fusion dataset obtains the best classification accuracy. Because the fusion dataset contains twice the classification features compared with a single alpha or beta dataset (both alpha and beta have 63 features, fusion has 126 features), it more easily obtains a better result. Another feasible explanation is that the fusion dataset not only captures the alpha but also contains the beta frequency band information, therefore, the fusion dataset can take the advantages of both frequency bands at the same time. Briefly speaking, feature fusion is an effective method for classification. It is noteworthy that when compared with the previous studies about action intention understanding classification [1, 5, 9], our novel method obtains more higher classification accuracies.

In the Fig. 4, many functional connective edges are still retained after serious FDR correction ($p < 0.05$). Because the difference brain networks are

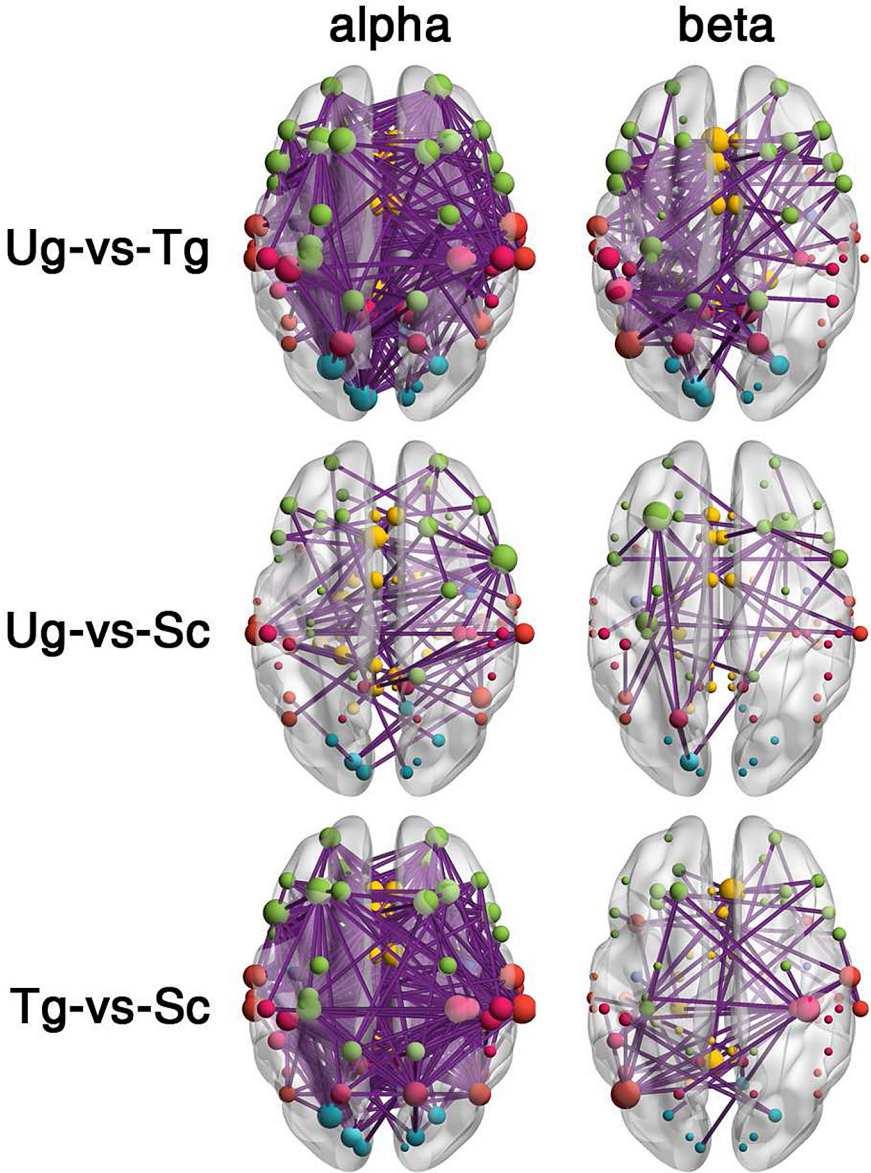


Fig. 4. Difference brain network. Each edge is determined by T-test and FDR correction ($p < 0.05$). The red, yellow, green, cyan, blue and purple-red nodes are from the temporal lobe, limbic lobe, frontal lobe, occipital lobe, sub-lobar and parietal lobe, respectively. The size of the node denotes the degree, the larger the size is, the higher the degree is. (Color figure online)

constructed by the mean of the 63 sub brain networks, hence, we conclude that the dynamical network effectively catches the difference information of action intention understanding brain activities in each microstate (50 ms). The alpha frequency band displays more significant edges than the beta frequency band, which suggests that the action intention understanding brain activities more easily occur at the alpha band [5, 16]. The classification accuracies in the alpha dataset are better than the results in the beta dataset (see Fig. 3) also support the viewpoint that action intention understanding brain activities closely correlate with the alpha frequency band [5]. Some big size nodes appear at the temporal, parietal, frontal and occipital lobes, which indicate that different action intention stimuli easily lead to different extents brain activities in these areas. These are consistent with some previous studies [6, 16, 17, 19]. According to distribution of the big size nodes in both alpha and beta frequency bands, we conclude that human beings need the cooperation of multiple brain regions to complete the correct understanding of action intention. If one big size node were destroyed, the structure of brain network were destroyed, then the corresponding function will be changed. This is why some people with brain injuries can not understand other people's intention perfectly, e.g., can not understand others' language and gesture.

Although the novel method obtains very high classification accuracies, there are still some limitations in this study. For instance, the classifications are implemented on the group level (we view each subject from one stimulus condition, Ug, Tg or Sc, as a classification sample), while the number of raw subjects is only 25. Hence, it still needs to expand the sample size in the future. Additionally, the classification accuracy based on integral level, i.e., the classification accuracy is equal to the number of correctly classified samples divided by the total number of samples in the test set. For the prediction of a real sample with unknown stimulus condition, the integral level calculate strategy can not complete effectively. The reason is that the classification task is carried out under the 5-fold cross validation condition, which easily leads the sample labels to be confused. Of course, we have a satisfactory way to solve this problem. With the idea of Leave-One-Out (LOO) cross validation, we can use the real sample to be predicted as the test set and all other samples as the training set. Because the real sample is divided into three (they come from PLV, PLI and WPLI, respectively), hence, it needs to predict three sample labels. When the labels of the three indices are consistent, then we can get the predict class of the real sample.

5 Conclusion

In summary, the novel method improves the classification accuracies of action intention understanding based on EEG signals effectively. It has some merits of generality, e.g., it can be applied in other state-of-the-art neuro informatics data recording technologies, such as MRI, fMRI, MEG, NIRS and so on, meanwhile, it also can be introduced into other neuroscience study fields, such as emotion, motor imagery, Alzheimer's disease, epilepsy, autism, alcohol addiction, etc. When adopt the LOO cross validation strategy, the novel method will

have some potential practical application values for the human-robot interaction. In the future, we will try our best to weaken the deficiencies in the novel method so as to obtain more satisfactory classification results and decode more complex neural mechanisms about the action intention understanding brain signals.

Acknowledgement. This work was supported in part by the National Nature Science Foundation of China under Grant 61773114, the Foundation of Hygiene and Health of Jiangsu Province under Grant H2018042, and the Key Research and Development Plan (Industry Foresight and Common Key Technology) of Jiangsu Province under Grant BE2017007-3.

References

1. Zhang, Z., Yang, Q., Leng, Y., Yang, Y., Ge, S.: Classification of intention understanding using EEG-NIRS bimodal system. In: 12th International Computer Conference on Wavelet Active Media Technology & Information Processing, Chengdu, pp. 67–70. IEEE (2015)
2. Brune, C., Woodward, A.L.: Social cognition and social responsiveness in 10-month-old infants. *J. Cogn. Dev.* **8**(2), 133–158 (2007)
3. Catmur, C.: Understanding intentions from actions: direct perception, inference, and the roles of mirror and mentalizing systems. *Conscious. Cogn.* **36**, 426–433 (2015)
4. Avanzini, P., Fabbridesiro, M., Volta, R.D., et al.: The dynamics of sensorimotor cortical oscillations during the observation of hand movements: an EEG study. *PLoS ONE* **7**, e37534 (2012)
5. Zhang, L., Gan, J.Q., Zheng, W., Wang, H.: Spatiotemporal phase synchronization in adaptive reconfiguration from action observation network to mentalizing network for understanding other's action intention. *Brain Topogr.* **31**(3), 447–467 (2017). <https://doi.org/10.1007/s10548-017-0614-7>
6. Ge, S., Ding, M., Zhang, Z., et al.: Temporal-spatial features of intention understanding based on EEG-fNIRS bimodal measurement. *IEEE Access* **5**, 14245–14258 (2017)
7. Dindo, H., Lopresti, L., Lacascia, M., Chella, A., Dedić, R.: Hanneket-based action classification for motor intention recognition. *Rob. Auton. Syst.* **94**, 120–133 (2017)
8. Bundy, D.T., Pahwa, M., Szrama, N., et al.: Decoding three-dimensional reaching movements using electrocorticographic signals in humans. *J. Neural Eng.* **13**(2), 026021.1–026021.18 (2016)
9. Liu, H., Zheng, W., Sun, G., et al.: Action understanding based on a combination of one-versus-rest and one-versus-one multi-classification methods. In: 10th International Congress on Image & Signal Processing, Shanghai, pp. 1–5. IEEE (2017)
10. Niso, G., et al.: HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity. *Neuroinformatics* **11**(4), 405–434 (2013). <https://doi.org/10.1007/s12021-013-9186-1>
11. Lachaux, J.P., Rodriguez, E., Martinerie, J., et al.: Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* **8**, 194–208 (1999)
12. Stam, C.J., Nolte, G., Daffertshofer, A.: Phase lag index: assessment of functional connectivity from multichannel EEG and MEG with diminished bias from common sources. *Hum. Brain Mapp.* **28**(11), 1178–1193 (2007)

13. Vinck, M., Oostenveld, R., Wingerden, M.V., Battaglia, F., Pennartz, C.M.A.: An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage* **55**(4), 1548–1565 (2011)
14. Hari, R.: Action-perception connection and the cortical mu rhythm. *Prog. Brain Res.* **159**(1), 253–260 (2006)
15. Avanzini, P., Fabbri-Destro, M., Volta, R.D., et al.: The dynamics of sensorimotor cortical oscillations during the observation of hand movements: an EEG Study. *PLoS ONE* **7**(5), e37534 (2012)
16. Xiong, X., Yu, Z., Ma, T., et al.: Weighted brain network metrics for decoding action intention understanding based on EEG. *Front. Hum. Neurosci.* **14**, 232 (2020)
17. Ortigue, S., Sinigaglia, C., Rizzolatti, G., Grafton, S.T.: Understanding actions of others: the electrodynamics of the left and right hemispheres. A high-density EEG neuroimaging study. *PLoS ONE* **5**(8), e12160 (2010)
18. Arnaud, D., Scott, M.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**(1), 9–21 (2004)
19. Fogassi, L., Ferrari, P.F., Gesierich, B., et al.: Parietal lobe: from action organization to intention understanding. *Science* **308**(5722), 662–667 (2005)