Subject Independent Classification of Implicit Intention Based on EEG Signals

Sang-Hoon Oh

Division of Information Communication Convergence Eng.
Mokwon University, Daejon, 302-729, Korea

ABSTRACT

Brain computer interfaces (BCI) usually have focused on classifying the explicitly-expressed intentions of humans. In contrast, implicit intentions should be considered to develop more intelligent systems. However, classifying implicit intention is more difficult than explicit intentions, and the difficulty severely increases for subject independent classification. In this paper, we address the subject independent classification of implicit intention based on electroencephalography (EEG) signals. Among many machine learning models, we use the support vector machine (SVM) with radial basis kernel functions to classify the EEG signals. The Fisher scores are evaluated after extracting the gamma, beta, alpha and theta band powers of the EEG signals from thirty electrodes. Since a more discriminant feature has a larger Fisher score value, the band powers of the EEG signals are presented to SVM based on the Fisher score. By training the SVM with 1-out of-9 validation, the best classification accuracy is approximately 65% with gamma and theta components.

Key words: Implicit Intention, EEG, Subject Independent BCI, Support Vector Machine, Fisher Score.

1. INTRODUCTION

Alphago, which composed of "value network" to evaluate board positions and "policy networks" to select moves, is a machine learning system to play the game of Go [1]. The value and policy networks are deep convolutional neural networks [2] trained by a novel combination of supervised learning from human expert games and reinforcement learning from games of self-play. The matches of Go between Lee Sedol and Alphago extremely increase the interests of artificial intelligence (AI), especially AI based on machine learning.

Machine learning can be defined as a "Field of study that gives computers the ability to learn without being explicitly programmed" [3]. Owing to splendid progress of machine learning [2], [4], it is widely used to identify objects in images, generate captions of images, transcribe speech into text, match new items, and select relevant results of search [2]. Furthermore, the application area of machine learning spreads to bioinformatics [5] and hydrology [6], [7] etc.

Emotion is a very important characteristic of human and developing a human-like machine learning system requires emotion recognition by analyzing people's speech, gesture, and facial expression [8]. The other important aspect of human is "implicit intention". In many cases, humans explicitly express their intention in various ways. However, sometimes in a very important moment, humans do not explicitly express their

intention but implicitly do.

Brain computer interfaces (BCIs) have focused on recognizing and responding to explicitly-expressed intentions [9], [10]. Surprisingly, there have been a few research to recognize humans' implicit intention [11]-[13]. The implicit intention can be categorized into unexpressed and hidden intention. The hidden intention corresponds to whether or not a human's explicitly-expressed intention is the same as the actual intention [11], [12]. The unexpressed intention is whether or not a human agrees with the others during conservation or sentence reading [13], [14].

Since electroencephalography (EEG) has non-invasive nature and high temporal resolution, many BCI technologies are based on EEG signals. Most recently, Dong et al. focused on the second category of implicit intention and reported the EEG-based classification of implicit intention during self-relevant sentence reading [13]. However, the best classification accuracy of 75.5% belongs to a subject dependent classification. This means that we can only classify the implicit intention of users whose EEG signals are already recorded and analyzed. For classifying implicit intention of unanalyzed users, we need a subject independent classifier. In this paper, we develop a subject independent classification of EEG signals to decide whether or not users agree with contents of self-relevant sentences.

2. EEG SIGNALS FOR IMPLICIT INTENTION

Cognitive processing of information in human brain is

closely related to EEG signals [15]. Specifically, there are power changes in frequency bands during language processing. Grammatical violations increase the theta band power, most at the temporal areas [16]. Also, the power increase in the gamma band is found at the frontal area in sentence processing [17]. The alpha band power decreases with semantic congruency [18], whereas the beta band power decreases owing to semantic violations [19]. These EEG-based studies are closely related to self-relevant sentences with strong responses. In this point of view, Dong et al. with computational neuro-systems laboratory (CNSL) at Korea Advanced Institute of Science and Technology (KAIST) presented the EEG-based subject dependent classification of implicit intention during self-relevant sentence reading [13].

The EEG signals for implicit intention by CNSL, KAIST, were measured from nine healthy right-handed Korean subjects [13]. For each subject, an experimental session consisted of 74 trials. The seventy four stimulating sentences for trials were selected from the Minnesota multiphasic personality inventory-II (MMPI-II) [20], one of the most frequently used standardized psychometric tests. All sentences were converted to contents block with SOV topology like '*The experience of worrying over money*' and sentence-ending block such as '*do exist*' or '*do not exist*'. After showing the contents block for four seconds and the sentence-ending block for two seconds, an asterisk image was presented for two seconds and subjects were instructed to answer whether or not they agreed with the statement. Here, the mission is to classify the subject's implicit intention of '*agreement*' and '*disagreement*' by EEG signals during reading of contents block.
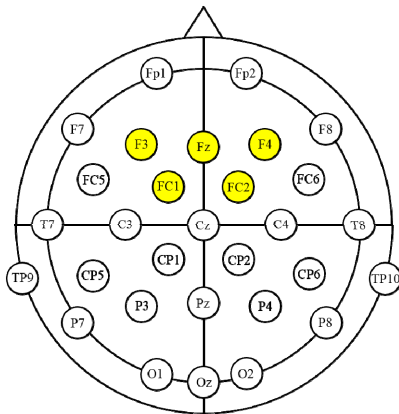

Fig. 1. The location of thirty electrodes on a subject's scalp. The yellow electrodes were used to find effective time-frequency components

The EEG signals were recorded using the BrainAmp system and an EEG cap with thirty two electrodes [13]. Thirty electrodes were placed on each subject's scalp as shown in Fig. 1 and two electrodes were positioned below subject's left eye and left collarbone to record electrooculogram (EOG) and electrocardiogram (ECG), respectively. After removing artifacts from eye blinking, eye movement, and heartbeats by independent component analysis (ICA) algorithm [21], the time-frequency representations of the thirty EEG signals from subjects' scalp were extracted by Morlet wavelet transform.

Based on fMRI studies with similar experiments [14], the yellow electrodes in Fig. 1 were used to find effective time-frequency components. Consequently, five time-frequency components were selected as (1) the gamma component (35Hz-45Hz) in 350~550ms, (2) the beta2 component (20Hz-26Hz) in 300~450ms, (3) the beta1 component (14Hz-17Hz) in 800~1,000ms, (4) the alpha component (9Hz-12Hz) in 300~700ms, and (5) the theta component (5Hz-7hz) in 400~1,000ms after the onset of the contents [13].

3. SUBJECT INDEPENDENT CLASSIFICATION OF IMPLICIT INTENTION

We use the EEG signals explained in the previous section for a subject independent classification of implicit intention [13]. A total of nine subjects were participated voluntarily and none of them had a history of mental disorder, significant physical illness, head injuries, neurological disorder, or alcohol or drug dependencies. Written informed consent was obtained from all subject and the study was approved by the institutional review board at KAIST. Throughout testing experiments for reliable and honesty responses, one trial data from Subject 5 and eleven trial data from Subject 6 were removed from the dataset. So, we have totally 653 trial data.

Since there are not separate sets of training and test data, we use 1-out of-9 validation method for training and performance evaluation of a classifier. That is, EEG signals from randomly chosen eight subjects among nine ones are used for training of a classifier and those from another subject are used for evaluating the test performance of classifier. Finally, the average of nine cases is the result of performance evaluation.

3.1 Fisher Score

Firstly, we extract the five time-frequency components in each electrode from Morlet wavelet transform of EEG signals. Since we do not know which component at which electrode is a good discriminant feature for classification, there must be a feature selection procedure before training a classifier. We adopt Fisher score to measure the discriminant ability of features [22]. For each time-frequency component, Fisher score for the i-th electrode is evaluated as

$$F_i = \frac{\sum_{k=1}^{K} n_k \left(\mu_k^i - \mu^i \right)^2}{\sum_{k=1}^{K} n_k \left(\sigma_k^i \right)^2} , \qquad (1)$$

where n_k is the number of data in the k-th class, μ_k^i and σ_k^i are the mean and standard deviation of the k-th class at the i-th electrode, respectively. Also, μ^i is the mean of whole data in the i-th electrode. Because we classify agreement or disagreement of EEG data, $K=2$. For each component, we evaluate Fisher score of training data and the data of top T rank electrodes are used as inputs of a subject independent classifier.

3.2 Support Vector Machine

Nowadays, there have been great improvement of machine learning field, especially with deep architecture of neural networks [2]. However, deep neural networks (DNNs) have so many parameters to be tuned during training and the number of training samples also should be increased proportional to the number of parameters.

As explained in the first paragraph of the Section 3, it is very difficult to gather EEG data because of life-ethics problem and we have only 653 trial data supplied by CNSL, KAIST. So, among many machine learning models, we select the support vector machine (SVM) with radial basis function (RBF) kernels because of its' better performance with less training samples [23]. SVM is deterministic and has few parameters to be tuned, while DNN is nondeterministic and has a lot of parameters.

Let's assume that we have a data set $\chi = \{\mathbf{x}^t, r^t\}_t$ where $r^t = +1$ if $\mathbf{x}^t \in C_1$ and $r^t = -1$ if $\mathbf{x}^t \in C_2$. Here, C_1 and C_2 denote the class 1 and class 2, respectively. Then, the key idea of SVM is to find an optimal discriminant hyperplane $\mathbf{w}$ and its bias b as shown in Fig. 2, where black circles are samples in C_1 and white circles are samples in C_2. For concrete separation of C_1 and C_2, the distance from the discriminant hyperplane to the closest sample on either side should be maximized. Thus, maximizing the margin subject to $r^t(\mathbf{w}^T \mathbf{x}^t - b) \geq +1, \forall t$ is the quadratic programming problem. Here, as shown in Fig. 2, samples on the margin are support vectors.

If the data set $\chi = \{\mathbf{x}^t, r^t\}_t$ cannot be separable by a hyperplane, we adopt a nonlinear kernel function which transforms the input data to another feature space. This allows the quadratic programming technique to fit the maximum margin hyperplane in a transformed feature space. Although the classifier is a hyperplane in the transformed feature space, it may be nonlinear in the original input data space. Since the classification of implicit intention with EEG signals is a nonlinear problem, we adopt a RBF as a kernel function and implement SVM using LIBSVM toolbox [24].
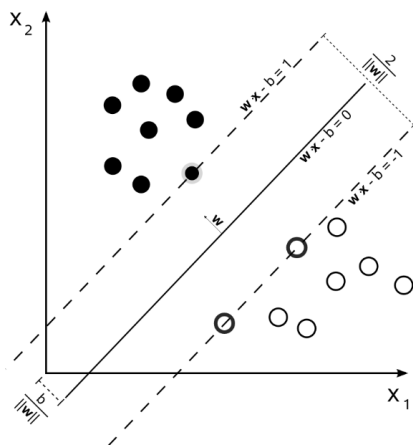

Fig. 2. Optimal maximum margin hyperplane of two classes. Samples on the margin correspond to support vectors

3.3 Simulations

Let's summarize the simulation procedure for the subject independent classification of implicit intention as follows: **1) Extraction of time-frequency components**: the five time-frequency components are extracted from Morlet wavelet transform of EEG signals. **2) 1-out of-9 validation**: Among whole time-frequency components of EEG signals for nine subjects with seventy four trials, we randomly select eight subjects' time-frequency component data for training of SVM with RBF kernel and the other subject's time-frequency component data for validating the test performance of SVM. Accordingly, we have nine sets of training and test data (named "training/test dataset1 ~ dataset9). **3) Evaluation of Fisher score (Table. 1)**: We evaluate the discriminant ability of each time-frequency component at thirty electrodes based on Fisher score of a training dataset and the data of top T rank electrodes are presented to SVM. **4) Training of SVM:** SVM with RBF kernel is trained using LIBSVM toolbox. **5) Averaging of performance evaluation (Fig. 3):** After training of SVM for each training dataset, we evaluate the classification performance of SVM for the corresponding test dataset. Finally, we average the nine cases of classification performances for training and test datasets.

Among the nine cases, Table. 1 shows the six best electrodes for each time-frequency component based on Fisher score of training dataset1. For the theta component, the selected electrodes are located at the frontocentral regions (FC5, F4, and FC6), the parietal region (P7), temporoparietal region (TP9), and the centroprietal region (CP6). For the alpha component, the selected electrodes are at the occipital regions (O2, Oz, and O1), the frontocentral region (FC1 and Fz), and the parietal region (P7). For the beta1 component, the selected electrodes are at the frontocentral regions (F4 and FC2), the parietal region (P7), temporoprietal region (TP10), the temporal region (T8), and the central region (C4). For the beta2 component, the selected electrodes are at the occipital region (Oz), the frontal regions (Fp1 and Fp2), and the frontocentral regions (FC6, F8 and Fz). For the gamma component, the selected electrodes are located at the frontocentral regions (F3, and Fz), centroparietal regions (Cp2 and CP5), and the parietal regions (P7 and Pz). Based on the above investigation, the frontocentral electrodes are the common locations to contain discriminant features of the five components. However, the importance of electrode position is different in each component since the rank orders are different component by component.

Table. 1. Six best electrodes for each time-frequency component based on Fisher score of training data set 1

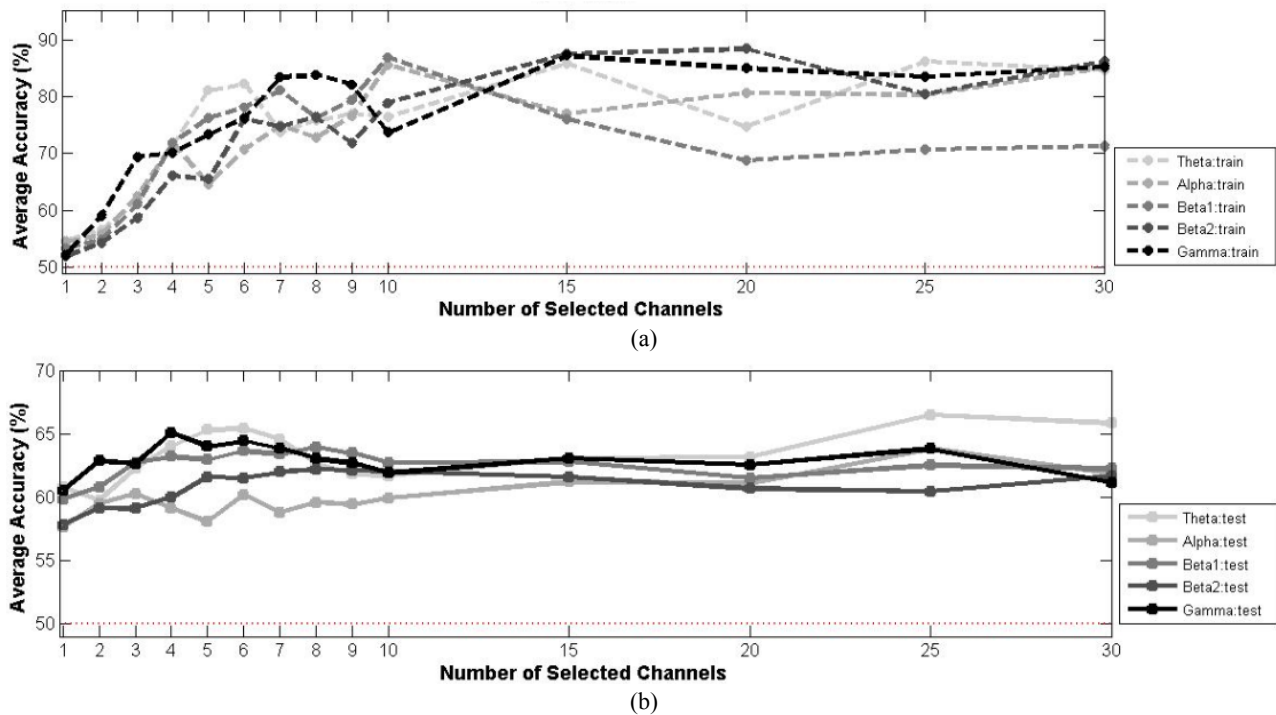| Rank	Theta	Alpha	Beta1	Beta2	Gamma
1	'FC5'	'O2'	'F4'	'Oz'	'F3'
2	'P7'	'Oz'	'FC2'	'Fp2'	'Fz'
3	'F4'	'FC1'	'P7'	'Fp1'	'CP2'
4	'TP9'	'Fz'	'TP10'	'FC6'	'CP5'
5	'CP6'	'P7'	'T8'	'F8'	'P7'
6	'FC6'	'O1'	'C4'	'Fz'	'Pz'

(a)

(b)

Fig. 3. Classification performance of implicit intention for each time-frequency component with increasing number of selected electrodes. (a) Training accuracy, (b) Test accuracy

Fig. 3 shows the classification performance after average of nine cases. Fig. 3(a) is the classification performance for training data (training accuracy) and (b) is that for test data (test accuracy). We attain 65.05% of test accuracy for four selected electrodes of the gamma component. Regarding the theta component, we attain 65.32% for five selected electrodes, 66.48% for twenty five selected electrodes, and 65.84% for whole thirty electrodes. The other components show inferior test accuracies. Also, there is a big discrepancy between the training and test accuracies, since the subject independent classification of implicit intention is a very difficult problem.

Dong et al. reported that they attained the best test accuracy of 75.5% with five gamma components in the subject dependent classification of implicit intention [13]. Comparing our result with the subject dependent case, there is approximately 10% of difference and this supports the difficulty of classifying EEG signals in subject independent case. So, there should be more efforts to improve the subject independent classification of EEG signals.

4. CONCLUSION

In this paper, we developed the subject independent classification of EEG signals to decide whether or not the user agrees with contents of self-relevant sentences. The classifier was implemented with the electrode selection of each time-frequency component based on Fisher score followed by training of SVM with RBF kernels using LIBSVM toolbox. The classification performance was evaluated by averaging the nine cases of 1-out of-9 validation method and test accuracy was 65.05% with four gamma components and 65.32% with five theta components. The best test accuracy is 66.48% with

five theta components.

REFERENCES

[1] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis, "Mastering the Game Go with Deep Neural Networks and Tree Search," Nature, vol. 529, Jan. 2016, pp. 484-489.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, May 2015, pp. 436-444.

[3] P. Simon, *Too Big To Ignore: The Business case for Big Data*, Wiley, 2013.

[4] S. H. Oh and H. Wakuya, "Comparison of Objective Functions for Feed-Forward Neural Network Classifiers Using Receiver Operating Characteristics Graph," Int. Journal of Contents, vol. 10, no. 1, Mar. 2014, pp. 23-28.

[5] S. H. Oh, "Protein Disorder Prediction Using Multilayer Perceptrons," Int. Journal of Contents, vol. 9, no. 4, Dec. 2013, pp. 11-15.

[6] C. W. Dawson and R. L. Wilby, "Hydrological modelling using artificial neural networks," Progress in Physical Geography, vol. 25, 2001, pp. 80-108.

[7] S. H. Oh and H. Wakuya, "Hydrological Modelling of Water Level near "Hahoe Village" Based on Multilayer Perceptron," Int. Journal of Contents, vol. 12, no. 1, pp. 49-53, Mar. 2016.

[8] B. K. Kim, J. Roh, S. Y. Dong, and S. Y. Lee, "Hierarchical Committee of Deep Convolutional Neural Networks for Robust facial Expression Recognition," J. Multimodal User Interfaces, vol. 10, issue. 2, pp. 173-189, Jun. 2016.

[9] Jonathan R. Wolpaw, Niels Birbaumerc, Dennis J. McFarlanda, Gert Pfurtschellere, and Theresa M. Vaughan, "Brain-Computer Interfaces for Communication and Control," Clin. Neurophysiol., vol. 113, 2002, pp. 767-791.

[10] Benjamin Blankertz, Guido Dornhege, Matthias Krauledat, and Klaus-Robert Müller, "The Berlin Brain-Computer Interface: EEG-based Communication without Subject Training," IEEE Trans. Neural Sys. and Rehabil., vol. 14, no. 2, 2006, pp. 147-152.

[11] L. A. Farwell and E. Donchin, "The truth Will Out: Interrogative Polygraphy ("Lie Detection") with Event-related brain Potentials," Psychophysiology, vol. 28, 1991, pp. 531-547.

[12] C. Davatzikos, K. Ruparel, Y. Fan, D. G. Shen, M. Acharyya, J. W. Loughead, R. C. Gur, and D. D. Langleben, "Classifying Spatial Patterns of Brain Activity with Machine Learning Methods: Application to Lie Detection," Neuroimage, vol. 28, 2005, pp. 663-668.

[13] S. Y. Dong, B. K. Kim, and S. Y. Lee, "EEG-based Classification of Implicit Intention during Self-relevant Sentence Reading," IEEE Trans. Cybernetics, Accepted for publication.

[14] S. Y. Dong, B. K. Kim, and S. Y. Lee, "Implicit Agreeing/Disagreeing Intention While Reading Self-relevant Sentences: A Human fMRI Study," Social Neuroscience, vol. 11, no. 3, 2016, pp. 221-322.

[15] E. Basar, M. Schürmann, T. Demiralp, C. Basar-Eroglu, and A. Ademo, "Event-related oscillations are 'real brain responses'—Wavelet analysis and new strategies," Int. J. Psychophysiol., vol. 39, nos. 2-3, Jan. 2001, pp. 91-127.

[16] D. Roehm, M. Schleschewsky, I. Bornkessel, S. Frisch, and H. Haider, "Fractionating Language Comprehension via Frequency Characteristics of the Human EEG," Neuroreport, vol. 15, no. 3, Mar. 2004, pp. 409-412.

[17] L. A. Hald, M. C. Bastiaansen, and P. Gagoort, "EEG Theta and gamma Responses to Semantic Violations in Online Sentence processing," Brain Lang, vol. 96, no. 1, Jan. 2006, pp. 90-105.

[18] R. M. Willems, R. Oostenveld, and P. Hagoort, "Early Decreases in Alpha and Gamma Band Power Distinguish Linguistic from Visual Information during Spoken Sentence Comprehension," Brain Res., vol. 1219, Jul. 2008, pp. 78-90.

[19] Y. Luo, Y. Zhang, X. Feng, and X. Zhou, "Electroencephalogram oscillations differentiate semantic and prosodic processes during sentence reading," Neuroscience, vol. 169, no. 2, Aug. 2010, pp. 654-664.

[20] J. N. Butcher, W. G. Dahlstrom, J. R. Graham, A. Tellegen, and B. Kaemmer, *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for Administration and Scoring,* Minneapolis, MN, USA: University of Minnesota Press, 1989.

[21] T. W. Lee, M. Girolami, A. J. Bell, and T. J. Sejnowski, "A unifying information-theoretic framework for independent component analysis," Comput. Math. Appl., vol. 39, Jun. 2000, pp. 1-21.

[22] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, Wiley-Interscience Publication, 2001.

[23] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer Science+Business Media, LLC, New York, 2006.

[24] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, Apr. 2011, pp. 1-27.

[25] http://www.csie.ntu.edu.tw/~cjlin/libsvm

Sang-Hoon Oh

He received his B.S. and M.S degrees in Electronics Engineering from Pusan National University in 1986 and 1988, respectively. He received his Ph.D. degree in Electrical Engineering from Korea Advanced Institute of Science and Technology in 1999. From 1988 to 1989, he worked for LG semiconductor, Ltd., Korea. From 1990 to 1998, he was a senior research staff in Electronics and Telecommunication Research Institute (ETRI), Korea. From 1999 to 2000, he was with Brain Science Research Center, KAIST. In 2000, he was with Brain Science Institute, RIKEN, Japan, as a research scientist. In 2001, he was an R&D manager of Extell Technology Corporation, Korea. Since 2002, he has been with the Department of Information Communication Engineering, Mokwon University, Daejon, Korea, and is now an associate professor. Also, he was with the Division of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, USA, as a visiting scholar from August 2008 to August 2009. His research interests are machine learning, speech signal processing, pattern recognition, and bioinformatics.