35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Classifying the speech response of the brain using Gaussian Hidden Markov Model (HMM) with Independent Component Analysis (ICA)

Jongin Kim, Suh-Kyung Lee, Boreom Lee*

Abstract— **The purpose of this paper is to determine whether electroencephalograpy (EEG) can be used as a tool for hearing impairment tests such as hearing screening. For this study, we recorded EEG responses to two syllables, /a/ and /u/, in Korean from three subjects at Gwangju Institute of Science and Technology. The ultimate goal of this study is to classify speech sound data regardless of their size using EEG; however, as an initial stage of the study, we classified only two different speech syllables using Gaussian hidden markov model. The result of this study shows a possibility that EEG could be used for hearing screening and other diagnostic tools related to speech perception.**

I. INTRODUCTION

Electroencephalography (EEG) is the electric fields produced by brain activity [2]. The ranges of its amplitude and frequency are 10~200uV and 1~50Hz, respectively. EEG was first recorded by Hans Berger in 1942, and generally is divided into delta, theta, alpha, beta, and gamma waves according to frequency range [3].

Through measured EEG values, it is possible to acquire brain activities that respond to external stimuli like sound, light, pain, and so on [2]. For this reason, there has been increasing interest in EEG classification for using the Brain Computer Interface (BCI) and various diagnostic tools.

For a good EEG classification, we should carefully consider the following two elements. First, we should choose appropriate EEG feature vectors. In order to design EEG-based BCI, a lot of features have been used. The representative features are EEG amplitudes, band powers (BP), power spectral density (PSD), autoregressive coefficient (AR), adaptive autoregressive coefficient (AAR), and so on [4]. Second, selecting an appropriate classifier is as important as a feature vector selection [4]. There are several types of classifiers, depending on their nature. A well-known standard to divide classifiers is a generative vs. discriminative classifier [4]. The generative classifier calculates the likelihood of each class and selects the class that shows the most probability. Bayes quadratic classifier is the most representative example of the generative classifier. On the other hand, the discriminative classifier does not use joint probability but directly classifies the features to each class. Support vector machine which is most popularly used in BCI is one of discriminative classifiers. Another taxonomy of classifiers is a static vs. dynamic classifier [5]. The static classifier does not use temporal information, but the dynamic classifier uses temporal information [4]. Since temporal information is a very important characteristic in almost all EEG feature vectors, a dynamic classifier such as the hidden markov model (HMM) is a suitable classifier in EEG classification problem [6]. Therefore, to classify EEG signals in this study, we used Gaussian hidden markov model (GHMM), which is a modified version of hidden markov model.

Additionally, it is necessary to remove noise before extracting feature vectors from EEG since it has a very low SNR. When we use an EEG feature vector, the noise which we should generally consider is electrooculogram (EOG), electromyogram (EMG), electrocardiogram (ECG), motion artifact and white noise [7, 8]. Band-pass filtering, fast fourier transform (FFT), autocorrelation, autoregressive modeling, adaptive filtering, Kalman filtering, and singular value decomposition (SVD) are generally used for noise cancelation in the signal processing field[9]. In the present study, we used independent component analysis (ICA) for noise removal with the Hurst exponent. In brief, ICA is an algorithm for dividing signals into statistically independent components, and the Hurst exponent automatically finds out the noise components [2, 9]. We will cover detailed descriptions of the ICA and Hurst exponent in Section III.

The objective of this paper is to classify the EEG signals which respond to two syllables, /a/ and /u/, using Gaussian hidden markov model with the independent component analysis for noise removal. In conclusion, the results of this paper present a possibility to use EEG as hearing screening and other diagnostic tools related to speech perception.

This paper is organized as follows. In Section II, we will explain why we chose such stimuli and how the stimuli were delivered to the subjects. Next, Section III will provide information on the preprocessing for rejecting noise such as EOG, ECG, and white noise. In Section IV, we will deal with in detail Gaussian hidden markov model as a classifier. Lastly, Section V will give the result of the experiment and a simple analysis.

The research was supported by the Pioneer Research Center Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0009462)

J. I. Kim is a master’s student in the Department of Medical System Engineering (DMSE), Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea (e-mail: kimji@gist.ac.kr).

S.-K. Lee is a research assistant in the Department of Medical System Engineering (DMSE), Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea (e-mail: suhkyung@alum.mit.edu).

B. Lee is an assistant professor in the Department of Medical System Engineering (DMSE), Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea (e-mail: leebr@gist.ac.kr).

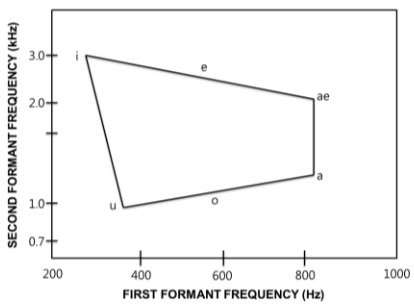

Figure 1 Difference in formant frequencies in vowels. Data points for /i u a æ/ are averages from Peterson and Barney (1952)[1].

II. STIMULUS SELECTION & DATA COLLECTION

A. Stimulus Selection

For a good classification of EEG responses to speech sounds, we should carefully select the speech stimuli. One factor we can consider when choosing a stimulus is formants. According to T. W. Picton et al., EEG waveforms which respond to speech follows the envelope of the sound [10]. In acoustics, the peaks found in the spectrum envelope of sound are called formants [11]. It is reasonable, therefore, that we select stimuli that have completely different formants to make them more classifiable. In case of vowels, they can be classified according to their articulatory characteristics such as the tongue body position, e.g., high, low, front, and back [1]. Each class of vowels, due to their vocal tract configurations, is associated with its consequent acoustic characteristics, e.g., high vowels - low first formants, back vowels - low second formants. Since /u/ and /a/, high-back and low-back vowels, respectively, show large difference in their first formant frequencies, (Figure1)[12] we expect to see better performance in classification with EEG response waveforms. For these reasons, we selected two vowels /a/ and /u/ as our test stimuli. All of the speech sounds used in this paper was provided by Naver standard pronunciation service (http://dic.naver.com/), which had been reviewed by the National Institute of the Korean Language for a six-month period.

B. Stimulus Presentation & Recording

As mentioned above, EEG is extremely vulnerable to noise. Especially EOG artifact is the most dominated noise in EEG waveform. To avoid EEG contamination by the EOG in the recording stage, we divided the recording procedure into 'eye-open' and 'eye-closed' stages and only recorded EEG waveforms during the eye-open stage. (Figure2) Since the subject cannot know when the eye-closed stage ends, a pre-test was prepared for the subject to get familiarized with the experimental procedure before the main experiment. Stimuli /a/ and /u/ were randomly presented in order to prevent the subject's prediction of the following stimulus. We presented 60 trials of vowel stimuli in a single session and two sessions of the experiment were conducted for a total of 120 trials of vowels per subject. All of the EEG data were recorded in the Department of Medical Science and Engineering (DMSE) at Gwangju Institute of Science and Technology, and the recording device was a 64 Channel EEG Net by Electrical Geodesics Inc. Volunteer subjects were three men.

III. PREPROCESSING

A. Basic Preprocessing

Preprocessing is a very important procedure for acquiring a meaningful EEG feature vector in classification problems. The first and foremost thing to consider in the preprocessing procedure is EMG artifact rejection. Generally, EMG artifact in EEG is dominated in frequencies above 20 Hz,[13] so IIR bandpass filter was adapted to the raw EEG signal (band width: 2-20 Hz, butterworth, order: 5). Next, we detrended the data to eliminate a linear trend of the signal, and adapted baseline correction using pre-stimulus data.

B. Independent Component Analysis

A lot of noise still remains in the EEG waveforms after the basic preprocessing procedure, so additional processes are needed for removing residual noise. Recently, approaches using the ICA algorithms have become popular for mitigating the effect of noise in biomedical signals [9]. According to Hoya et al. [14], classification accuracy was greatly increased using ICA. Before adapting the ICA to EEG signals, modeling the EEG procedure was imperative. It is well-known that a linear model is suitable for EEG [15]. The formula can be expressed as follows:

$$\mathbf{F}\mathbf{s}(t) + \mathbf{w}(t) = \mathbf{x}(t). \tag{1}$$

Where $\mathbf{s}(t) = [s_1(t)\ s_2(t)\ \ldots\ s_m(t)]^T$ is an m-dimensional unknown source vector; $\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_n(t)]^T$ is an n-dimensional vector of observed sensor signals; $\mathbf{F}$ is a forward model which is an (n × m) unknown mixing matrix; and $\mathbf{w}$ is a vector of white Gaussian noise. There are a number of methods to solve the inverse problem for finding the forward model. One of the methods is the ICA, which estimates the mixing matrix $\mathbf{F}$ by decomposing the EEG signals into statistically independent components [2]. According to Scott makeig et al. [16], each ICA component may represent activities generated from different biological sources. Therefore, it is possible to acquire meaningful EEG data through an ICA algorithm by reconstructing components after rejecting the components that are irrelevant to EEG. In this paper, we selected the FastICA algorithm due to its

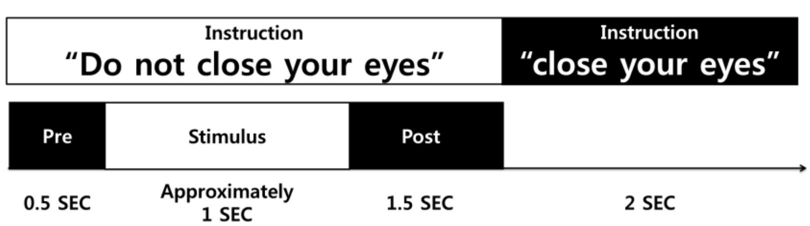

Figure 2. Trial time course for a single trial. Stimulus /a/ or /u/ was randomly presented during the stimulus session. EEG waveforms were recorded only during the 'eye-open' stage (pre-stimulus (0.5sec), stimulus (approximately 1sec), post-stimulus (1.5sec)).

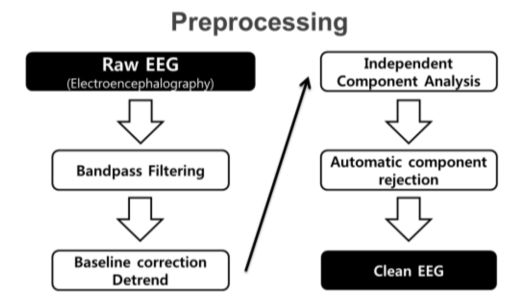

Figure 3. Preprocessing procedure.

efficient computation time[17].

$$\mathbf{s}(t) = \mathbf{F}^{-1}\mathbf{x}(t) = \mathbf{S}\mathbf{x}(t). \tag{2}$$

According to Aapo Hyvarinen, FastICA finds the separation matrix **S** that maximizes the non-Gaussianity of projection data using the fixed-point iteration scheme [17].

The following procedure is repeated until S converges:

1. Set random values as the initial **S**.

2. $E\{\mathbf{x}g(\mathbf{S}^T\mathbf{x})\} - E\{g'(\mathbf{S}^T\mathbf{x})\}\mathbf{S} \Rightarrow \mathbf{S}'$, where function g is the derivative of a non-quadratic function.

3. Let $\mathbf{S}'/\|\mathbf{S}'\| \Rightarrow \mathbf{S}$.

4. Check whether S converges or not.

After decomposing the data into independent components, we needed a procedure to identify the noise components. Since EEG data have high dimensionality, it is very difficult to manually inspect noise components. Therefore, we used the Hurst exponent to automatically identify the noise components [9]. Vorobyov and Cichocki mentioned that the Hurst exponent of the components that are contaminated by ECG and EOG signals is 0.58 - 0.69 [9]. Therefore, we evaluated the Hurst exponent of all components, then rejected the components which were within the 0.58 – 0.69 range. Figure 3 is a simple diagram of the preprocessing.

IV. Classification

For classifying the EEG data which responded to speech, HMM was used in this study because the HMM classifier is suitable for classifying time series data [4]. Performance of HMM in classifying time series data is very good, so even raw EEG data are classified well [18]. Figure4 is a basic diagram of HMM classifier. Classification steps were divided into training and testing phases. The objective of the training phase was to construct a model that shows the maximum likelihood of observed vectors. The class of the observed vectors was two types (/a/ EEG and /u/ EEG), and therefore, a total of two HMM models were needed. Each model HMM(/a/) and HMM(/u/) was updated from /a/ and /u/ EEG data, respectively, using Baum-Welch algorithms. In the testing phase, an unknown class of EEG data was used as the input for HMM /a/ and HMM /u/. Finally, HMM classified the unknown EEG to the class of the model that shows a greater likelihood than the other. Since Gaussian Hidden Markov Model (GHMM) is generally appropriate for EEG classification problem, [18] we used GHMM in this paper.

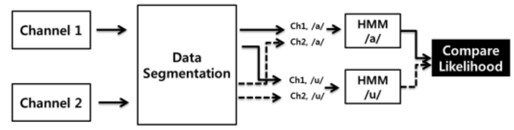

Figure 4 Schematic diagram of HMM Classifier

V. Result

Since HMM classifier is a statistical process, it was necessary to iterate each classification process ten times to evaluate general performance of the classifier. A total of 100 trials were used to evaluate the performance of the classifier and divided into 80 trials for the training phase and 20 trials for the testing phase. Figure5 shows the change of classification accuracy according to the length of a feature. As shown in Figure5, feature length of 200ms shows the best performance. Althogh performance for 200ms and 1s were similar, when considering computation time, we determined the 200 ms as the optimal feature length.

Another factor we selected is delay. To find the optimal delay, we tested the classification in 0ms, 200ms, 400ms respectively. As a result, The classifier shows the best in 200ms delay. (figure6) According to T. W. Picton et al., brain responses follow the envelope of speech with about a 200 ms delay that is consistent with the result of this study.[10]

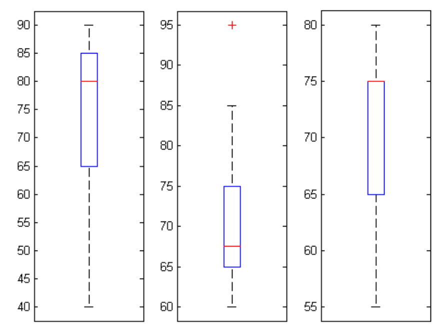

Figure 5 Classification accuracy according to the feature length, feature length (200ms, left), feature length (500ms, middle), feature length (1s, right), y-axis means the classification accuracy and red line means median of data, subject LKJ

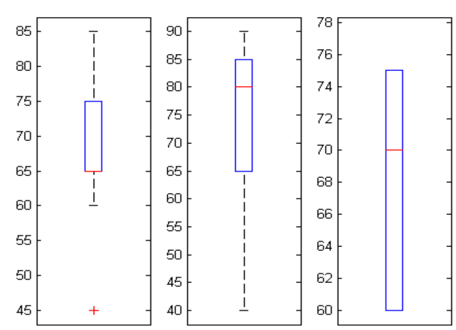

Figure 6 Classification accuracy according to the feature delay. Delay (0ms, left), Delay (200ms, middle), Delay (400ms, right). y-axis means the classification accuracy and red line means the median of data, subject LKJ

Repeating the previous process, we obtained data from three subjects. As shown in Table 1, all the results are above chance level. Especially, subject PCK shows a surprising result a median value greater than 95%.

TABLE I. CLASSIFICATION ACCURACY FOR EACH SUBJECTS

Subject	*Median*	*25th percentile*	*75th percentile*
LKJ	80 %	65 %	85 %
PCK	95%	90%	100%
JDR	85%	75%	90%

VI. DISCUSSION & CONCLUSION

Because of low SNR of the EEG, classifying EEG responses to speech sounds was very difficult; therefore for extracting speech representation from EEG, we adopted some signal processing techniques. Since there is no clear standard for selecting interesting sources related to speech stimuli, we defined a standard for selecting phoneme responses prior to using signal processing techniques. Generally, it is well known that the Hurst exponent of ECG and EOG is 0.58 -0. 69; therefore we decomposed EEG responses into independent components using ICA, and then reject the components which were within the 0.58 – 0.69 range. Althogh this is just two class classification problem, brain responses to two Korean vowels were classified in very high accuracy using hidden Markov model classifier. This means that ICA is a useful method to extract speech-related components from EEG. We consider that the components which are extracted using ICA may provide an insight for phoneme representation in the brain because they contain some distinct information about brain response to each vowel, thus enabling successful classification as reported. In future studies, we will try to classify the brain responses to many types of vowels which have very close formants to each other. We expect that this algorithm could be used for Brain Computer Interface (BCI), diagnosis and rehabilitation of diseases related to hearing.

REFERENCE

[1] G. E. Peterson, “Control Methods Used in a Study of the Vowels,” *Acoustical Society of America Journal,* vol. 24, pp. 175, 1952.

[2] B. V. K. V. K. Katharine Brigham, "Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy.", Bioinformatics and Biomedical Engineering (Icbbe), pp. 1-4, 2010

[3] G. Guan, Y. Yuan, Z. Yisheng, and Q. Yihong, "The phase analysis of ongoing EEG oscillations under face/object perception.", Biomedical Engineering and Informatics (BMEI), pp. 1063-1066, 2010.

[4] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for EEG-based brain–computer interfaces,” *Journal of Neural Engineering,* vol. 4, 2007.

[5] L. Prevost, and M. Milgram, "Static and dynamic classifier fusion for character recognition.", Document Analysis and Recognition, pp. 499-506 vol.2, 1997.

[6] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proceedings of the IEEE,* vol. 77, no. 2, pp. 257-286, 1989.

[7] Z. Weidong, and J. Gotman, "Removal of EMG and ECG artifacts from EEG based on wavelet transform and ICA.",Engineering in Medicine and Biology Society,vol.1, pp. 392-395, 2004..

[8] X. Zhaojun, L. Jia, L. Song, and W. Baikun, "Using ICA to Remove Eye Blink and Power Line Artifacts in EEG.", Innovative Computing, Information and Control, vol.3, pp. 107-110, 2006.

[9] S. Vorobyov, A. Cichocki, “Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis,” *Biological Cybernetics,* vol. 86, no. 4, pp. 293-303, 2002.

[10] T. W. Picton, S. J. Aiken, and M. Science, “Human Brain Responses to Speech Sounds,” Thesis, 2008.

[11] A. H. Benade, *Fundamentals of Musical Acoustics: Second, Revised Edition (Dover Books on Music)*: Dover Publications, 1990.

[12] K. N. Stevens, *Acoustic Phonetics (Current Studies in Linguistics)*: The MIT Press, 2000.

[13] K. J. P. Emma M. Whithama, Sean P. Fitzgibbonc Trent Lewisb, C. Richard Clarkc, Stephen Lovelessd, Marita Broberge, Angus Wallacee, Dylan DeLosAngelese, Peter Lillief, Andrew Hardyf, Rik Fronskof, Alyson Pulbrookg, John O. Willoughby, “Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG,” *Clinical Neurophysiology,* vol. 118, no. 8, pp. 1877–1888, August, 2007.

[14] G. H. Tetsuya Hoya, Hovagim Bakardjian, Tomoaki Nishimura, Taiji Suzuki, Yoichi Miyawaki, Arao Funase, Jianting Cao, "Classification of Single Trial EEG Signals by a Combined Principal + Independent Component Analysis and Probabilistic Neural Network Approach.", International Symposium on Independent Component Analysis and Blind Signal Separation, pp. 197-202, 2003.

[15] L. C. P. Lucas Parra Clay, Clay D. Spence, B Adam D. Gerson, Paul Sajda C, “Recipes for the linear analysis of EEG,” *NeuroImage,* vol. 28, pp. 326-341, 2005.

[16] Scott Makeig, Stefan Debener, Julie Onton, and Arnaud Delorme, “Mining event-related brain dynamics,” *Trends in Cognitive Sciences,* vol. 8, no. 5, pp. 204-210, May 01, 2004.

[17] A. Hyvarinen, “Fast and robust fixed-point algorithms for independent component analysis,” *Neural Networks, IEEE Transactions on,* vol. 10, no. 3, pp. 626-634, 1999.

[18] A. M. N. Soroosh Solhjoo, Mohammad Reza, Hashemi Golpayegani, "Classification of chaotic signals using HMM classifiers: EEG-based mental task classification.", 13th European Signal Processing Conference, 2005.