User authentication systems based on brain finger-prints

Soo-Young Lee*[a], Eun-Soo Jung[a]
[a]Dept. of Electrical Engineering, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea 305-701

ABSTRACT

We propose to use EEG signals to make user authentication for requiring high security. EEG signals were measured while the subjects saw several images in sequences. Since subjects' EEG signals are different for known and unknown images, these EEG sequences may be used to identify each subject. Correlation analysis and classification results show the feasibility of user authentication from EEG signals.

Keywords: EEG, brain finger-prints, security, authentication, single trial analysis

1. INTRODUCTION

1.1 Motivation and goals

As the importance of information and security has been arisen, the safer authentication or identification systems are desired. The codes or keys should be unique and complicated. Therefore, individual's biometric signatures such as fingerprints, irises, or voices are used for authentic systems in these days. Most of people have those features, and they are different from each other. Moreover, it is comfortable and convenient because biometric signatures themselves are used as keys or codes, which have no worries of being lost or forgotten. However, there are still limitations. Most of the user identification systems are not safe against rubber hose attacks, which are applying mental or physical force to obtain password or keys of an authentic system. Also, most of the biometric signatures have chances to be duplicated or stolen, and attackers can break into the system. Therefore, an identification system using brain signal with memory, which can be considered as the most complex biometric signature, is proposed. It is impossible to generate certain pattern of brain signals with today's technology. Brain signals are very complicated to be generated artificially, and no one can control his or her brain signals.

1.2 Background

The human brain is a very important body part and has tremendous cognitive functions. Memorizing and containing the information from experiences can be considered as the most important function. Every human has different experiences and emotions through his or her whole life, therefore, each brain contains different information. Reading people's mind or thinking has been dreamed and even tried since a long time ago. These days, human got closer to read minds with science and technology. Yet we cannot fully understand others' thinking and emotion, however, can get clues from signals such as facial expression, pupil diameter, skin conductance, and so on. Moreover, the technology such as electroencephalography (EEG), functional magnetic resonance imaging (fMRI), or positron emission tomography (PET) leads us to get signals from brain with noninvasive methods. Among those technologies, we used EEG to get brain signals and tried to show the possibilities of using brain signals with memories as biometric signature.

Since brain signal measuring technologies have been developed, the interest in using brain signals had been increased. Recently few studies reported the use of implicit memory in motor-related skills in human identification and 'pass-thoughts' to replace passwords [1, 2, 3]. However, the time for gathering those skills from individuals is too long, and still there are chances for the authentic system to be cracked if the fixed tasks are revealed. On the other hand, human memory can be detected faster, as well as it can act like an embedded key. Moreover, we are proposing to use brain signals which are produced in passive way by presenting stimuli.

*sy-lee @kaist.ac.kr; phone 82 42 350-3431; fax 82 42 350-8490; cnsl.kaist.ac.kr/

Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering XII, edited by Harold H. Szu, Liyi Dai, Proc. of SPIE Vol. 9118, 91180M

A technique for determining whether the specific information is contained in a person's brain with EEG is called brain fingerprinting. It is yet controversial but has been used in forensic science and the advocators of the technique argue about its low error rate and high statistical confidence [4, 5]. According to the research on brain fingerprinting, event related potentials (ERPs), especially p300 can be detected when subjects are shown to the stimuli, which are known or familiar. Also, ERP differences between familiar and unfamiliar faces have been reported with latencies of ERPs such as 250ms, 300ms, 400ms and so on [6, 7]. A study on attacking revealing user's private information, which is opposite to security, was also attempted by presenting visual stimuli and the feasibility was shown even with inexpensive EEG cased BCI devices [8]. Inversely, we can eavesdrop other's memories by evoking brain signals with stimulus and they can be formed as a code. Brain signals can be formed into the most complicating code but does not need to be memorized. It means that even the users do not know the exact code. The codes or keys are already implemented in users' brains and the system will draw them out. In other words, the system drives out brain signals from the users with stimuli and they would produce the signals in passive way. In this research, series of visual stimulus were used to evoke brain signals from the subjects.

Achieved brain signals from experiment were passed through few processes, than the users were identified. In some stages, we presented several methods for single-trial analysis, which is a main issue of current BCI study [9, 10].

2. DATA ACQUISITION

2.1 Experiments

We recruited fourteen healthy volunteers were recruited under the approval of the Institutional Review Board (IRB). All of the participants were right handed and the ages were ranged in 20 to 28 years with a mean age of 23 and variance of 8.

Visual stimuli were used in this research, and the categories for the stimuli were facial pictures, pictures of sites, and flags. Those categories are suitable for the stimuli because individuals without certain disabilities can all recognize tremendous of distinctive faces or places with familiarity, and they are relatively easier to obtain. The visual stimuli were carefully selected and arranged so that the composition of known stimuli can be different for each participant. Therefore, we had collected some pictures before the experiments from all the participants; two pictures of themselves, two pictures of a man or woman they know in person, such as family or friend. Some pictures of celebrities were selected and added for the visual stimuli. Too famous faces, which can be recognized by anyone, were avoided to control the number of known and unknown faces for each subject. The ideal case was each subject knows half of the faces with different composition from others.

Conditions of pictures such as brightness or size were controlled. Finally, each picture became a monochrome picture with 400 by 400 pixels. Figures of flags were remained colored and resized with original shapes.

Overall 150 face pictures, 50 site pictures, and 50 flags were gathered for an experiment. Thus, the total number of visual stimuli was 250, and 250 trials composed a session. A visual stimulus was presented in a trial. Each subject participated two sessions per a day, and the experiments were repeated twice again with a week interval. Overall six sessions were conducted for each subject. In a session, 250 trials were separated into four blocks. Different categories of pictures were included in different blocks, therefore, it makes two blocks of facial pictures (half for each), a block for sites, and a block for flags.

All the sessions had same stimuli, because the responses from same stimuli needed to be compared. If it was needed, participants could have recess between blocks. When the sessions were repeated on same day, the order of stimuli was same as the previous one. However, the orders were different for different days. We gave fake information to the subjects that only some of the stimuli were same as previous day. Especially, for sessions on the third day, four to five 'dummy trials' were added to each block to prevent subjects' stereotype that all of the stimuli were same as before. Those additional stimuli were not counted as trials and neither included in analysis.

In a trial, each visual stimulus was shown for 1s. However, to prevent sudden pop up of figure from black screen, pre-stimulus of white noise with same size as stimulus was inserted for 0.5s before a stimulus. Using pre-stimulus is a popular tactic for cognitive research, such as observing EEG signals or pupillometry [12, 13]. Also, after showing a picture, a question was shown for 2s on the screen, such as 'do you know the face?', and subjects had to speak out the

answers; yes or no. The oral responses would unnecessary in practical security system, however, were collected and used as labels for analysis and single-trial classification.

The EEG signals were recorded with a cap from Brain Vision with 32 integrated electrodes located at standard positions of the International 10-20 system [11]. Among thirty-two channels, thirty are placed on scalp potentials, one below left eye for Electrooculogram (EOG) and the other near left collarbone for Electrocardiogram (ECG). A separate channel for reference was placed on FCz location. The sampling rate was 500Hz and various filters were used; notch filter at 60Hz, low pass filter at 70Hz, and high pass filter at 0.531Hz. Individual sensors were adjusted under 20kΩ of impedances. The signals were recorded with software, BrainVision Recorder 1.10 and BrainVision Analyzer 1.05 was used for data export.

2.2 Preprocessing

The methods of EEG signal measuring and preprocessing can be emphasized as analyzing and can effect on the result significantly. At the same time, they are very tricky, because it is very hard to define or revealing the best methods. In this research, various methods for removing noise or artifacts and enhance meaningful signals were shown. First of all, average-reference transformation was adopted to enhance signals of channels near reference (channel FCz). Then, independent component analysis (ICA) was used, and the components for eye blinking or heartbeats were removed. Low pass filter (high cut off frequency at 35Hz) and high pass filter (low cut off frequency at 0.95Hz) were also used to reduce high frequency noise and slow voltage changes. Finally, filtered and artifact removed data was segmented into epochs. Each epoch includes from 0.2s before to 0.6s after stimulus onset. Averaged amplitude of signal before a stimulus onset was used as a baseline.

3. FEATURE EXTRACTION

Since the EEG signals were measured with 500Hz from 31 scalp channels, the data was needed to be effectively reduced. In this research, we aimed to analyze signals with single trials. Therefore, efficient feature extraction was necessary. Even though ICA and filters were used to reduce noise, each signal from a trial still contains noise that can be hardly ignored. Signals' peak amplitudes and latency are important information, however, often covered by noise. To observe such information and reduce meaningless frequency parts at the same time, we adopted discrete cosine transform (DCT).

3.1 Discrete cosine transform (DCT)

The DCT is a transform which expresses a finite sequence of data points with a sum of cosine functions [14]. In particular, DCT can be considered as a Fourier-related transform similar to the discrete Fourier transform (DFT) with real numbers only. However, the difference from magnitude of DFT is that DCT is not shift invariant [15]. With this characteristic, information with time domain also can be observed. Therefore, the coefficients related to latencies of ERPs can be preserved. At the same time, the data of high frequency range, which can be considered as noise, were efficiently removed. In this research, DCT type-II was used, and the length of each DCT signal was 1024.

3.2 Observation of DCT signals

We defined two classes of stimulus as 'known' and 'unkown'. With the oral responses from subjects, the EEG signals were separated into the two classes. Then, the grand averages of DCT signals for the two classes were examined. Some channel strongly showed that the grand averages for 'known' were distictive from that from 'unknown' class. Those channels were F7, T7, Pz, CP1 and CP2. Espacially, the difference could be detected in the first nine DCT coefficients. However, the first coefficients had great variences. Therefore, the DCT coefficients from 2^{nd} to 9^{th} from the five channels were slected as the feature to represent response type.

4. RESULTS

4.1 Behavioral results

As it was mentioned, the stimuli were same for all the sessions and subjects. If the subjects were always naïve for the stimuli sets, the identification would be easier with less change in the EEG responses. However, subjects usually recognized many of visual stimuli from the last experiments, and the portion of 'known' increased. Answers from the first session of each day were counted. We considered that if a subject could recall a stimulus that was presented a week ago, the information was stored in long-term memory. Some of the subjects hardly remembered what was shown in the last week's experiment and some called back most of the stimuli. The number of 'known' stimulus was ranged from 59 to 117 per a subject, and the average was 90. Since the total number of stimulus was 250, the ratio was 36%. For day-2, average number of 'known' stimulus per a subject was 144, which was more than half (57.6%). Finally, some of the subjects answered that they remember most of the stimulus on day-3, and the average was 183.5 per a subject (73.2%).

4.2 Single trial classification

To support the assumption that the different resonses related to memory can be observed through EEG signals, we tried single-trial classification with the two classes; 'known' and 'unknown'. For classifier, nonlinear support vector machine (SVM) with radial basis function (RBM) was used. SVM is a binary learning machine which selects discriminant hyper-plane that maximizes the margins between classes [16, 17]. Parameters for SVM were adjusted for maximum recognition rate for each case. EEG responses from single-trials were classified subject dependently. In other words, single-trials from each subject were classified separately.

It was shown that channel F7, T7, Pz, CP1, and CP2 has more discriminative signals for the two classes than other channels. Also, DCT coefficients from 2^{nd} to 9^{th} included useful data for the classification. Therefore, 8 points from those 5 channels were concatenated, and a feature vector of 40 points was obtained for each trial. Each trial was classified with the feature.

In each session, there were 250 trials. For within session classification, 3-fold cross validation was used. Trials for class 'known' and 'unknown' were separately divided into three groups randomly. A group from each class was selected as test set in turrn, and the rest groups became training set. Subject-6 memorized too many stimuli, thus removed from the classification. Among 78 sessions (except 6 sessions from subject-6), the worst session showed error rate of 50% and the best session showed error rate of 24.74%. The overall mean of the error rate was 38.55%.

4.3 Identification of session data

There were 6 sessions for a subject, thus the total numbers of sessions from all subjects was 84. Each of the session includes 250 trials, and features from trials can be concatenated and form a vector for each session. To identify the users, the data need to be compared. For efficient comparison, we checked correlation between signals. In other words, each row vectors were normalized by dividing with 2-norm of the vector, and the inner product between two normalized vectors stands for their similarity. The maximum value would be 1, when the two vectors are identical. The correlation between all the pairs can be presented with correlation matrix, which is descrived in Figure 1. The diagonal components of the correlation matrix are always 1, because it presents the similairity between same vectors.

The optimal correlation matrix would be the block-diagonal matrix. The block matrices size of 6 by 6, which has 1 for all components, should be placed on the diagonal and prove that all sessions from a subject share relatively higher correlation values. We expected the correlation matrix to be similar as the optimal correlation matrix. One way of evaluate the identification result was observing the correlation matrix.

For the identification, it was required to use both memory information and subject distinguishability. Features were extracted from 2^{nd} to 16^{th} DCT coefficients. The range covers both features for memory information and subject distinguishability. The first DCT coefficient was not included because we had filtered the signals with high pass filter in preprocessing. Channels were selected for the highest identification rate. Among various compositions of channels, T7, T8, CP5, CP6, CP1, CP2, Pz, and Oz drew out the best identification rate (15.72%). Fifteen DCT coefficients from selected channels were concatenated for a trial. Then, features from 250 trials were merged and formed a vector of the session.

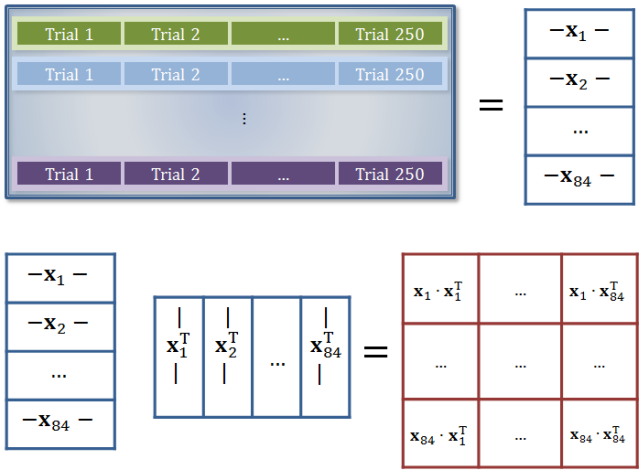

Figure 1: The matrix that represents correlation between all the pairs of sessions. Normalized vector x_n stands for the concatenated data for n^{th} session. With simple matrix product, the correlation between all the pairs can be avhieved into correlation matrix, size of 84 by 84.

5. DISCUSSION

An identification system using brain signals related to memory was suggested, and potentials of the system were shown. The research was covered from the experiment for gathering data to identification results. From the experiment, subjects passively generated brain responses, and the responses were obtained with EEG signals. Also, experiments were repeated to understand memory and brain signal variation. We assumed that EEG signals can reveal the person's memory information, and individuals have unique patterns of brain signals. The assumptions were checked in various ways, such as grand average observation or single-trial classification. DCT was used for the feature extration, and similarities between session data were observed with correlation coefficients. Those methods are unique methods for EEG signal analysis.

Single-trials were classified into 'known' and 'unknown' classes to confirm that those passively induced EEG signals contain the subject's memory information about the stimuli. SVM classifier was used, and the results with various compositions of train and test set were observed. Single-trials were possibly classified into the two classes. However, when the EEG signals from different days were used as train and test data, the classification error rate was increased.

After it was clarified that single trials can be discriminated into 'known' and 'unknown' classes, the whole sessions were identified. Data from a session was identified by comparing correlation with other sessions. The best result showed false positive of 15.71%. However, we need more study for channel selection, and which channels strongly produce brain signals for identification.

Also, subjects started to know about the stimuli, and the familiarity of overall stimuli increased as the experiment repeated. Therefore, it was obvious that the brain responses for stimuli had been changed during the whole period of experiment, and it influences on the results. To overcome the brain signal variation with time passing, the database of identification system needs to be regularly updated. Both of the data for stimuli and the users' brain signals can be updated. The signals from previous identification can be stored and used to update brain signal database. However, not only users' brain signals, but also users' memories can be varied rapidly. To catch up with what the user learned and experienced, life logging can be applied. If the system can be implemented with commercial grade EEG measuring tools, this idea can be applied in tremendous situations, such as web access or authentication of the smart phone user. For this research, we needed to collect data for stimuli, but information in smart phones or social networks can be gathered and presented in further use. Then the amount of database for stimuli would be huge and can be updated rapidly. The user also upload or share photoes and articles, thus automatically update the data in social networks. Yet it might sound cumbersome, using brain activities and memories can be a strong method of human identification and authentication.

Acknowledgment: This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2012-0008681; 2012.09.01-2013.08.31).

REFERENCES

[1] Bojinov, H.,Sanchez, D., Reber, P., Boneh, D.,and Lincoln, P., "Neuroscience meets cryptography: designing crypto primitives secure against rubber hose attacks," Proc. 21st USENIX Security Symp. (2012).

[2] Marcel, S. and Millan, J.R., "Person Authentication using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation," IEEE trans.Pattern.Anal. Machine Intell. 29(4), 743-748 (2007).

[3] Thorpe, J., van Oorschot, P. C., and Somayaji, A., "Pass-thoughts: authenticating with our minds," Proc. 2005 workshop on New security paradigms, 45-56 (2005).

[4] Farwell, L.A. and Donchin,E., "The Truth Will Out: Interrogative Polygraphy ("Lie Detection") With Event-Related Brain Potentials," Psychophysiology, 28(5), 531-547 (1991).

[5] Rosenfeld, Peter, J., and Labkovsky, E., "New P300-based protocol to detect concealed information: Resistance to mental countermeasures against only half the irrelevant stimuli and a possible ERP indicator of countermeasures." Psychophysiology, 47.6, 1002-1010 (2010).

[6] Tanaka J.W., Curran, T., Porterfield, A.L., and Collins,D., "Activation of Preexisting and acquired face representations. The N250 Event-related Potential as an Index of Face Familiarity," J. Cognitive Neurosci.18(8), 1-10 (2006).

[7] Bentin, S., and Deouell, L. Y., "Structural encoding and identification in face processing: ERP evidence for separate mechanisms,"Cognitive Neuropsychology, 17(1-3), 35-55 (2000).

[8] Martinovic, I., Davies, D., Frank, M., Perito, D., Ros, T., and Song, D., "On the feasibility of side-channel attacks with brain-computer interfaces," Proc. 21st USENIX Security Symp. (2012).

[9] Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K. R., "Single-trial analysis and classification of ERP components—a tutorial," NeuroImage, 56(2), 814-825 (2011).

[10] Neuper, C., Scherer, R., Reiner, M., and Pfurtscheller, G., "Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial EEG." Cogn. Brain Res., 25(3), 668-677 (2005).

[11] Deuschl, G., and Eisen, A., "Recommendations for the practice of clinical neurophysiology (guidelines of the international federation of clinical neurophysiology)," Electroencephalogr. Clin. Neurophysiol., Suppl. (1999).

[12] Luck, S.J., [An introduction to the event-related potential technique (Cognitive neuroscience)], MIT press, Messachusetts (2005).

[13] Klingner, J., Kumar, R., and Hanrahan, P., "Measuring the task-evoked pupillary response with a remote eye tracker." Proc. 2008 symp. Eye tracking research & appli., 69-72, ACM (2008).

[14] Chen, W. H., Smith, C. H., and Fralick, S., "A fast computational algorithm for the discrete cosine transform," Comm., IEEE Transac. 25(9), 1004-1009 (1977).

[15] Heckmann, M., Kroschel, K., Savariaux, C., and Berthommier, F., "DCT-based video features for audio-visual speech recognition," Proc. INTERSPEECH (2002).

[16] Haykin, S., [Neural networks and learning machines], Upper Saddle River, New Jersey (2009).

[17] Vapnik, V., [Statistical Learning Theory], Wiley-Interscience, New York (1998).