A Proposed Feature Extraction Method for EEG-based Person Identification

Phuoc Nguyen, Dat Tran, Xu Huang and Dharmendra Sharma
Faculty of Information Sciences and Engineering
University of Canberra, ACT 2601, Australia

Abstract— *We propose in this paper a feature extraction method to extract brain wave features from electroencephalography (EEG) signal. The proposed feature extraction method is based on an assumption that EEG signal could be considered as stationary if the time window is sufficiently short. With this assumption, EEG signal has some similar properties to speech signal and hence a feature extraction method that is currently used to extract speech features can be applied to extract brain wave features from EEG signal. Mel-frequency cepstral coefficients are features extracted and evaluated in EEG-based person identication. Experimental results show that the proposed method could provide very high recognition rate.*

Keywords: EEG, Person Identification, Brain Computer Interface

1. Introduction

Brain Computer Interface (BCI) has been considered as a new communication channel that uses brain activity as reflected by electric, magnetic or hemodynamic brain signals to control external devices such as computers, switches, wheelchairs, or neuroprosthetic extensions [1]. A BCI system can be classified as an invasive or non-invasive BCI according to the way the brain activity is being measured within this BCI. If the sensors used for measurement are placed within the brain, i.e., under the skull, the BCI is said to be invasive. On the contrary, if the measurement sensors are placed outside the head, on the scalp for instance, the BCI is said to be non-invasive [2]. The non-invasive BCI systems avoid health risks and associated ethical concerns. A typical non-invasive BCI system includes the following stages: data acquisition, data pre-processing, feature extraction, classification, device controller and feedback [3].

In invasive BCIs, electrodes or a multiunit electrode array will be placed directly inside the cortex to record electrical potentials for subsequent analysis of the electrocorticogram (ECoG). Brain signals obtained have a superior signal-to-noise ratio, need little user training, and are suitable for replacing or restoring lost motor functions in patients with damaged parts of the neuronal system [3]. Noninvasive BCIs, on the other hand, can use a variety of brain signals as input, for example, electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and near infrared spectroscopy (NIRS). MEG, fMRI, and NIRS are expensive or bulky, and fMRI and NIRS present long time constants in that they do not measure neural activity directly—relying instead on the hemodynamic coupling between neural activity and regional changes in blood flow—they cannot be deployed as ambulatory or portable BCI systems [4]. As a result, the majority of promising non-invasive BCI systems to date exploit EEG signal, mainly due to its fine temporal resolution, ease of use, portability and low set-up cost.

The use of brain wave patterns for person authentication has been investigated at IDIAP in Switzerland [5]. It has been shown that the brain-wave pattern of every individual is unique and that the EEG can be used for biometric identification [6]. The use of brain wave patterns as a new modality for person authentication has several advantages: 1) it is confidential because it corresponds to a mental task; 2) it is very difficult to mimic because similar mental tasks are person dependent; and 3) it is almost impossible to steal because the brain activity is sensitive to the stress and the mood of the person, an aggressor cannot force the person to reproduce his/her mental pass-phrase [5].

It has been shown that under the same recording condition EEG signals are low intra subject variability and high inter subject variability. In [7], [9], [10], [11] the subjects were asked to image moving hand, finger, foot or tongue while EEG data was recorded. Then the person recognition tasks were performed for each task separately. In [12] the subjects were asked to look at black and white drawings of common objects, extracted from the Snodgrass and Vanderwart picture set, while EEG signals were recording.

As indicated in [2], although numerous pre-processing, feature extraction and classification methods have been proposed and explored for BCI systems, none of them has been identified as the best method for EEG-based BCIs.

We have found that the EEG signal is a slowly time varying signal in the sense that, when examined over a sufficiently short period of time, depending on the signal between 5 and 100 ms, its characteristics are approximately stationary. However over longer periods of time, the signal characteristics are non-stationary. They change to reflect the sequence of different brain activities. Based on this "quasi-stationarity" which is also observed in speech signal, a reasonable brain wave model should have the following components. First, short-time measurements at an interval

of the order of 10 ms are to be made along the pertinent speech dimensions that best carry the relevant information for linguistic or speaker distinction. Second, because of the existence of the quasi-stationary region, the neighbouring short-time measurements on the order of 100 ms need to be simultaneously considered, either as a group of identically and independently distributed observations or as a segment of a non-stationary random process covering two quasi-stationary regions.

Based on this "quasi-stationarity", a feature extraction method that is currently used to extract speech features can be applied to extract brain wave features from EEG signal. In this paper we consider mel-frequency cepstral coefficients (MFCCs) which are the most popular speech features and extract them from EEG signal. We evaluated this feature extraction method in EEG-based person identication using support vector machine (SVM). Experimental results show that the proposed method could provide very high recognition rate.

2. Speech Features

Speech is a time varying signal. In a long period, speech signals are non-stationary but in a short interval between 5 and 100ms, the speech signals are "quasi-stationary", and the articulatory configuration stays nearly constant. Therefore, speech features are extracted for short frames. The sampled waveform is analysed in frames with short window sizes so that the signals are "quasi-stationary". The frames overlap by setting the frame period smaller than the window size. Each frame is then investigated to extract parameters. This process results in a sequence of parameter blocks [19]. source rate and target rate are the number of samples of the wave source and the number of extracted feature vectors, respectively.

In speech signal processing, the window size is typically between 15 ms and 35 ms long with a period of 10 ms. When adapting to EEG signal processing, we multiply the window size by a factor of 10 due to low frequency signal.

We define a frame of speech to be the product of a shifted window with the speech sequence [21]:

$$f_s(n;m) = s(n)w(m-n) \tag{1}$$

where $s(n)$ is the speech signal and $w(m-n)$ is a window of length N ending at sample m.

2.1 Mel-Frequency Cepstral Coefficients

The filterbank models the ability of the human ear to resolves frequencies non-linearly across the audio spectrum and decreases with higher frequencies. The filterbank is an array of band-pass filters that separates the input signal into multiple components. The filters used are triangular and they are equally spaced along the mel-scale defined by [19]:

$$\mathrm{Mel}(f) = 2595\log_{10}(1 + \frac{f}{700}) \tag{2}$$

Mel-Frequency Cepstral Coefficients (MFCCs) are calculated from the log filterbank amplitudes m_j using the Discrete Cosine Transform

$$c_i = \sqrt{\frac{2}{N}} \sum_{j=1}^{n} m_j \cos\left(\frac{\pi i}{N}(j-0.5)\right) \tag{3}$$

where N is the number of filterbank channels, c_i are the cepstral coefficients.

2.2 Energy

These features model intensity based on the amplitude. The energy is computed as the average of the signal energy, that is, for speech samples $s(n), n = 1, \ldots, N$, the short-term energy of the speech frame ending at m is [20]

$$E_s(m) = \frac{1}{N} \sum_{n=m-N+1}^{m} f_s(n;m)^2 \tag{4}$$

2.3 Pitch

The pitch signal, or the glottal waveform, is produced from the vibration of the vocal folds. Two common features related to the pitch signal are the pitch frequency and the glottal air velocity [20]. The vibration rate of the vocal folds is the fundamental frequency of the phonation F_0 or pitch frequency. The air velocity through glottis during the vocal fold vibration is the glottal volume velocity. The most popular algorithm for estimating the pitch signal is based on the autocorrelation [20]. At first, the signal is low filtered at 900 Hz and then it is segmented to short-time frames of speech $f_s(n;m)$. Then the nonlinear clipping procedure that prevents the first formant interfering with the pitch is applied to each frame $f_s(n;m)$ giving

$$\hat{f}_s(n;m) = \begin{cases} f_s(n;m) - C_{thr} \text{if } |f_s(n;m)| > C_{thr} \\ 0 \text{if } |f_s(n;m)| < C_{thr} \end{cases} \tag{5}$$

with C_{thr} is about 30% of the maximum value of $f_s(n;m)$. Next the short-term autocorrelation is determined by

$$r_s(\eta;m) = \frac{1}{N} \sum_{n=m-N+1}^{m} \hat{f}_s(n;m)\hat{f}_s(n-\eta;m) \tag{6}$$

where η is the lag. Finally, the pitch frequency of the frame ending at m can be given by

$$\hat{F}_0(m) = \frac{F_s}{N} \mathrm{argmax}_\eta \{|r(\eta;m)|\}_{\eta=N(F_l/F_s)}^{\eta=N(F_h/F_s)} \tag{7}$$

where F_s is the sampling frequency, and F_l, F_h are the lowest and highest perceived pitch frequencies by humans, respectively. Normally, $F_s = 8000$ Hz, $F_l = 50$ Hz, and $F_h = 500$ Hz [20]. The maximum value of the autocorrelation $\max\{|r(\eta;m)|\}_{\eta=N_w(F_l/F_s)}^{\eta=N_w(F_h/F_s)}$ represents the glottal velocity volume.

2.4 Zero Crossing Measure

The number of zero crossings, or number of times the sequence changes sign, is also a useful feature in speech analysis. The short-term zero crossing measure for the N-length interval ending at $n = m$ is given by [21]:

$$Z_s(m) = \frac{1}{2N} \sum_{n=k}^{m} |sign\{s(n)\} - sign\{s(n-1)\}| w(m-n) \tag{8}$$

where $k = m - N + 1$ and

$$\text{sign}\{s(n)\} = \begin{cases} +1 & if \quad s(n) \geq 0 \\ -1 & if \quad s(n) < 0 \end{cases} \tag{9}$$

2.5 Probability of Voicing

Pitch detection has high accuracy for voiced pitch hypotheses but the performance degrades significantly as the signal condition deteriorates. Pitch extraction for telephone speech is more difficult because the fundamental is often weak or missing. Therefore it is more useful to provide F_0 value and probability of voicing at the same time. The hypothesis is that first, voicing decision errors will not be manifested as absent pitch values; second, features such as those describing the shape of the pitch contour are more robust to segmental misalignments; and third, a voicing probability is more appropriate than a "hard" decision of 0 and 1, when used in statistical models [22].

2.6 Jitter and Shimmer

Jitter and shimmer are micro fluctuations in vocal fold frequency and amplitude. They are correlated to rough or hoarse voice quality. The major difference is that shimmer has irregular amplitude at regular frequency while in contrast jitter has irregular frequency at regular amplitude. The wave in the top picture has irregular amplitude at the third peak and the wave in the bottom picture has irregular frequency at the second peak.

Jitter indicates cycle-to-cycle changes of the fundamental frequency and is approximated as the first derivative of the fundamental frequency [23]. These changes are considered as variations of the voice quality.

$$\text{jitter}(n) = \frac{|F_0(n+1) - F_0(n)|}{F_0(n)} \tag{10}$$

where $F_0(n)$ is the fundamental frequency at sample n. Shimmer indicates changes of the energy from one cycle to another.

$$\text{shimmer}(n) = \frac{|\text{en}(n+1) - \text{en}(n)|}{\text{en}(n)} \tag{11}$$

where en(n) is energy of sample n.

3. Support Vector Machine (SVM)

3.1 Binary Case

Consider the training data $\{x_i, y_i\}$, $i = 1, ..., n$, $x_i \in R^d$, where label $y_i \in \{-1, 1\}$. The support vector machine (SVM) using C-Support Vector Classification (C-SVC) algorithm will find the optimal hyperplane [24]:

$$f(x) = w^T \Phi(x) + b \tag{12}$$

to separate the training data by solving the following optimization problem:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \tag{13}$$

subject to

$$y_i \left[w^T \Phi(x_i) + b \right] \geq 1 - \xi_i \text{ and } \xi_i \geq 0 \text{ , } i = 1, ..., n \tag{14}$$

The optimization problem (13) will guarantee to maximize the hyperplane margin while minimizing the cost of error, where $\xi_i, i = 1, ..., n$ are non-negative slack variables introduced to relax the constraints of separable data problems to the constraint (14) of non-separable data problems. For an error to occur the corresponding must exceed unity (see Eq. (14)), so $\sum_i \xi_i$ is an upper bound on the number of training errors. Hence an extra cost $C \sum_i \xi_i$ for errors is added to the objective function where C is a parameter chosen by the user.

The Lagrangian formulation of the primal problem is:

$$L_P = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_i \alpha_i \{ y_i ({x_i}^T w + b) - 1 + \xi_i \} - \sum_i \mu_i \xi_i \tag{15}$$

We will need the Karush-Kuhn-Tucker conditions for the primal problem to attain the dual problem:

$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \Phi(x_i)^T \Phi(x_i) \tag{16}$$

subject to

$$\begin{aligned} 0 \leq \alpha_i \leq C \\ \sum_i \alpha_i y_i = 0 \end{aligned} \tag{17}$$

The solution is given by

$$w = \sum_i^{N_S} \alpha_i y_i x_i \tag{18}$$

where N_S is the number of support vectors. Notice that data only appear in the training problem, Eq. (15) and Eq. (16), in the form of dot product and can be replaced by any kernel K

with $K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j), \Phi$ is a mapping to map the data to some other (possibly infinite dimensional) Euclidean space. One example is Radial Basis Function (RBF) kernel $K(x_i, x_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2}$

In test phase an SVM is used by computing the sign of

$$f(x) = \sum_{i}^{N_S} \alpha_i y_i \Phi(s_i)^T \Phi(x) + b = \sum_{i}^{N_S} \alpha_i y_i K(s_i, x) + b \tag{19}$$

where s_i is the ith support vector.

3.2 Multi-class Support Vector Machine

The binary SVM classifiers can be combined to handle the multi-class case: One-against-all classification uses one binary SVM for each class to separate their members from other classes, while one-against-one or pairwise classification uses one binary SVM for each pair of classes to separate members of one class from members of the other. In one-against-one approach, there are $n(n-1)/2$ class pairs decision functions were trained. In test phase, the voting stategy was used as follow: each binary classification was considered to be a voting where votes could be cast for all data points x. The final result was the class with maximum number of votes [25].

4. Datasets

The Graz dataset IIIa in the BCI Competition 2005 (Graz IIIa 2005), Graz dataset A (Graz A 2008) and Graz dataset B (Graz B 2008) in the BCI Competition 2008 come from the Department of Medical Informatics, Institute of Biomedical Engineering, Graz University of Technology for motor imagery classification problem in BCI Competition 2005 and 2008 [26], [27], [28]. The Graz IIIa 2005 dataset contains EEG recordings of 3 subjects. Each subject was required to do cued 4 motor imagery task (left hand, right hand, foot, tongue). The recording was made with a 64-channel EEG amplifier from Neuroscan at 250 Hz with time length 7s for each trial. The Graz A 2008 dataset contains EEG recordings of 9 subjects. This dataset use the same cue-based BCI paradigm as the Graz IIIa 2005. Two sessions on different days were recorded for each subject with 288 trials per session. Twenty-two Ag/AgCl electrodes were used and the signals were sampled with 250 Hz and bandpass-filtered between 0.5 Hz and 100 Hz. The Graz B 2008 dataset consists of EEG data from 9 subjects. The subjects paticipated in two sessions contain training data without feedback (screening), and three sessions were recorded with feedback. It consisted of two classes: the motor imagery (MI) of left hand and right hand. Three bipolar recordings (C3, Cz, and C4) were recorded at sampling frequency of 250 Hz.

The dataset searched and downloaded from the web-based de-identified searchable database Australian EEG Database used in this research consists of EEG recordings of 40 patients. The database consists of EEG records recorded at the John Hunter Hospital [29], near University of Newcastle, over an 11-year period. The recordings were downloaded with the search criteria that the recordings come from the both man and women in various age. The recordings were made by 23 electrodes placed on the scalp sampled at 167 Hz for about 20 minutes.

Table 1: Dataset descriptions (#ses: number of sessions, and length: measured in seconds)

Dataset	#subjects	#tasks	#trials	#ses	length
Graz IIIa 2005	3	4	60	1	7
Graz A 2008	9	4	288	2	7.5
Graz B 2008	9	2	120	5	7.5
Australian EEG	40	free	1	1	1200
Alcoholism (large)	20	2	120	1	1
Alcoholism (full)	122	2	120	1	1

The Alcoholism datasets come from a study to examine EEG correlates of genetic predisposition to alcoholism [30]. The datasets contain EEG recordings of control and alcoholic subjects. Each subject was exposed to either a single stimulus (S1) or to two stimuli (S1 and S2) which were pictures of objects chosen from the 1980 Snodgrass and Vanderwart picture set. When two stimuli were shown, they were presented in either a matched condition where S1 was identical to S2 or in a non-matched condition where S1 differed from S2. The 64 electrodes placed on the scalp sampled at 256 Hz for 1 second. The Alcoholism large dataset contains training and test data for 10 alcoholic and 10 control subjects. The Alcoholism full dataset contains 120 trials for 122 subjects. The summary of those datasets is listed in Table 1.

5. Feature Extraction

We used the open-source Emotion and Affect Recognition toolkit's feature extracting backend openSMILE [31] for extracting brain wave features. Each channel is extracted features individually and then features from all channels are merged together. The features include MFCCs, spectral features, energy, pitch (F0), zero crossing rate, probability of voicing, jitter and shimmer and their statistics functionals [32].

The 8 channels selected are C3, Cz, C4, P3, Pz, P4, O1, O2 which are an extension of [9], except for The Graz B 2008 dataset which has only three channels C3, Cz and C4. Because of the resulting high dimensionality the Correlation-based Feature Selection with Sequential Floating Forward Selection is used [33].

Except for the Graz datasets which have separated training and testing sets, the Australian EEG dataset and the Alcoholism datasets used 1/3 for cross validation training and 2/3 for testing. Linear SVM classifiers [25] is trained in 3-folds

cross validation scheme with parameter C ranges from 1 to 1000 in 5 steps.

6. Experimental Results

Person identification rates in test phase are 99% on Graz IIIa 2005, 80.8% on Graz B 2008, 46.24% on Graz A 2008, 92.8% on Alcoholism large and 61.7% on Alcoholism full datasets. Tables 2, 3 and 4 show the confusion matrices for the Graz datasets. Results show high identification rate on Graz IIIa 2005 and Graz B 2008. However it is not high on Graz A 2008. The reasons could be that the data for training is not enough when there is high variation in data (4 imagery tasks), and the training data should come from different sessions in order to have good performance in test phase.

Table 2: Confusion matrix in test phase on Graz dataset IIIa BCI Competition 2005. Identification rate is 99%.

Classified as →	k3b	k6b	l1b
k3b	149	0	0
k6b	0	83	0
l1b	0	3	80

Table 3: Confusion matrix in test phase on Graz data set A, BCI Competition 2008. Identification rate is 46.24%.

→	A1	A2	A3	A4	A5	A6	A7	A8	A9
A1	219	9	42	0	11	0	0	0	0
A2	0	96	108	0	0	0	36	43	0
A3	5	9	244	0	3	0	0	12	0
A4	10	12	11	148	43	4	0	0	0
A5	0	0	0	0	0	0	276	0	0
A6	0	6	0	2	0	194	0	0	13
A7	221	14	0	0	15	23	0	4	0
A8	0	6	86	0	0	1	0	178	0
A9	0	0	0	0	0	217	0	31	16

Table 4: Confusion matrix in test phase on Graz dataset B, BCI Competition 2008. Identification rate is 80.8%.

→	B1	B2	B3	B4	B5	B6	B7	B8	B9
B1	221	2	0	0	5	0	0	0	0
B2	20	66	0	2	91	0	12	3	51
B3	0	0	230	0	0	0	0	0	0
B4	0	0	0	141	1	0	146	16	3
B5	1	30	0	0	242	0	0	0	0
B6	0	0	0	0	1	250	0	0	0
B7	0	0	0	0	0	0	206	0	26
B8	4	2	0	6	6	0	0	212	0
B9	0	0	0	2	0	0	0	0	243

Figure 1 shows the identification rate on Australian EEG database using a single channel Cz and 8 channels C3, Cz, C4, P3, Pz, P4, O1 and O2. For a single channel case, the EEG recording length should be at least 7 seconds for the person to be recognizable. The identification rate has a peak of 94% at 15 seconds length then slowly decreases and climbs up again. The use of 8 channels shows stable performance 97% at recording length from 3 to 60 seconds and shows higher performance of 99% at recording length longer than 90 seconds.

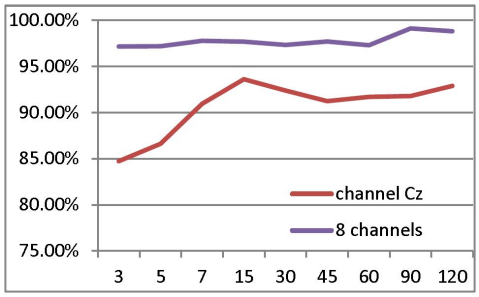

Fig. 1: Identification rate (in %) versus segment length on Australian EEG dataset

Person identification rates on Alcoholic large and Alcoholic full show that a single channel can hardly recognise the subjects. The use of 8 channels have high performance 92.8% in identification of 20 subjects and it can classify 122 subjects.

Table 5: Accuracies on dataset Alcoholic large and Alcoholic full using a single channel Cz and 8 channels C3, Cz, C4, P3, Pz, P4, O1, O2.

	Alcoholic large		Alcoholic full	
	channel Cz	8 channels	channel Cz	8 channels
Cross validation	52.50%	100%	21.87%	99.68%
Test	34.50%	92.83%	18.36%	61.59%

7. Conclusion

We have proposed a speech-based approach to brain wave feature extraction and used the proposed features to identify person. A single channel with appropriate recording time interval can be used in person identification task. The set 8 channels can give very good identification performance. For future investigation, other speech feature extraction methods will be applied to brain wave feature extraction.

References

[1] F. Babiloni, A. Cichocki, and S. Gao, "Brain-Computer Interfaces: Towards Practical Implementations and Potential Applications", Computational Intelligence and Neuroscience, Volume 2007, Article ID 62637, 2 pages, 2007.

[2] F. Lotte, M. Congedo, A. LÃľcuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brainâĂŞcomputer interfaces", Journal of Neural Engineering, 4 R1, 2007.

[3] R. Krepki, G. Curio, B. Blankertz, K.-R. Muller, "Berlin Brain-Computer Interface-The HCI communication channel for discovery", International Journal of Human-Computer Studies, vol. 65, no. 5, pp. 460-477, 2007.

[4] A. Cichocki, Y. Washizawa, T. Rutkowski, H. Bakardjian, A. Phan, S. Choi, H. Lee, Q. Zhao, L. Zhang, and Y. Li, “Noninvasive BCIs: Multiway Signal-Processing Array Decompositions”. Computer vol. 41, no. 10, pp. 34-42, 2008.

[5] Marcel, S. and Millan, J. d., Person Authentication Using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation. IEEE Trans. Pattern Analysis and Machine Intelligence, 29(4), 743-752, 2007.

[6] Wang, Z., A hybrid SVMâĂŞGLM approach for fMRI data analysis, Neuro Image, vol. 46, pp. 608âĂŞ615, 2009.

[7] S. Sun , "Multitask learning for EEG-based biometrics," Pattern Recognition, 2008. ICPR 2008. 19th International Conference on , vol., no., pp.1-4, 8-11 Dec. 2008

[8] M. A. Lebedev, and M. A.L. Nicolelis, Brain-machine interfaces: past, present and future. TRENDS in Neurosciences: Elsevier, 2006.

[9] X. Bao, J. Wang, J. Hu, "Method of individual identification based on electroencephalogram analysis, New Trends in Information and Service Science", 2009. NISS ’09. International Conference on, pp. 390-393, 2009.

[10] J. F. Hu, "New biometric approach based on motor imagery EEG signals, BioMedical Information Engineering", 2009. FBIE 2009. International Conference on Future, 2009

[11] D. Jiang, J. Hu, "Research of Computing in EEG Password based on Wavelet". International Conference on Future BioMedical Information Engineering, 2009

[12] K. Bringham and B. V. K. V. Kumar, "Subject Identification from Electroencephalogram (EEG) Signals During Imagined Speech", 4th BTAS, 2010.

[13] R. Palaniappan and D. P. Mandic, "Biometrics from brain electrical activity: A machine learning approach," IEEE Trans. Pattern Anal., vol. 29, no. 4, 2007.

[14] R. Palaniappan, "Identity Verification using Resting State Brain Signals". Encyclopedia of Information Ethics and Security 2007

[15] R. Palaniappan, "Two-stage biometric authentication method using thought activity brain waves, International Journal of Neural Systems", vol. 18, no. 1, 2008.

[16] F. Bimbot et. al., "A Tutorial on Text-Independent Speaker Verification". EURASIP Journal on Applied Signal Processing, 2004

[17] D.A. Reynolds, R.C. Rose; , "Robust text-independent speaker identification using Gaussian mixture speaker models," Speech and Audio Processing, IEEE Transactions on , vol.3, no.1, pp.72-83, Jan 1995

[18] "The INTERSPEECH 2009 Emotion Challenge". B. Schuller, S. Steidl, A. Batliner, INTERSPEECH, 2009

[19] S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. C. Woodland. "The HTK Book (for HTK Version 3.4)", 2009.

[20] Dimitrios Ververidis and Constantine Kotropoulos. "Emotional speech recognition: Resources, features, and methods". Speech Communication, **48**(9):1162–1181, 2006.

[21] J. R. Deller, J. H. L. Hansen, and J. G. Proakis. "Discete-Time Processing of Speech Signals". Wiley, N.Y, 2000.

[22] W. Chao and S. Seneff. "Robust pitch tracking for prosodic modeling in telephone speech". In Acoustics, Speech, and Signal Processing, 2000. ICASSP ’00. Proceedings. 2000 IEEE International Conference on, pages 1343–1346 vol.3, 2000.

[23] S. Steidl, "Automatic Classification of Emotion-Related User States in Spontaneous Children’s Speech". PhD thesis, 2009.

[24] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition". Data Mining and Knowledge Discovery, 1998.

[25] Chih-chung Chang and Chih-Jen Lin. LIBSVM: a Library for Support Vector Machines, 2001.

[26] G. Pfurtscheller, A. SchlÃűgl, Dataset IIIa: 4-class EEG data. Available on line at http://www.bbci.de/competition/iii/

[27] C. Brunner, R. Leeb, G. R. MÃijller-Putz, A. SchlÃűgl, and G. Pfurtscheller, BCI Competition 2008 - Graz data set A. Available online at http://www.bbci.de/competition/iv/

[28] R. Leeb, C. Brunner, G. MÃijller-Putz, A. SchlÃűgl, G. Pfurtscheller, BCI Competition 2008 - Graz data set B. Available online at http://www.bbci.de/competition/iv/

[29] M. Hunter, R. Smith, W. Hyslop, O. Rosso, R. Gerlach, J. Rostas, D. Williams, and F. Henskens, "The Australian EEG database", Clin EEG Neurosci, vol. 36, no. 2, pp. 76âĂŞ81, 2005.

[30] H. Begleiter, EEG Database. Available online at http://kdd.ics.uci.edu/databases/eeg/eeg.data.html

[31] F. Eyben, M. WÃűllmer, and B. Schuller. OpenEAR - Introducing the Munich open-source emotion and affect recognition toolkit. Proceedings - 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009, 2009

[32] B. Schuller, S. Steidl, A. Batliner, F. Schiel, J. Krajewski, "The Interspeech 2011 Speaker State Challenge, Interspeech (2011)", ISCA, Florence, Italy, 2011.

[33] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. "The WEKA data mining software: an update". SIGKDD Explor, 2009.

[34] L. I. Kuncheva and J. J. Rodriguez, “Classifier ensembles for fMRI data analysis: an experiment, Magnetic Resonance Imaging”, vol. 28, no. 4, pp. 583-593, 2010.

[35] A. Riera, A. Soria-Frisch, M. Caparrini,C. Grau, and G. Ruffini, "Unobtrusive Biometric System Based on Electroencephalogram Analysis". EURASIP Journal on Advances in Signal Processing, 2008.

[36] A. ZÂt’uquete, B. Quintela, J. P. S. Cunha, "Biometric authentication using brain responses to visual stimuli". Proceedings of the International Conference on Bio-inspired Systems and Signal Processing, 2010.

[37] C. Miyamoto, S. Baba and I. Nakanishi, "Biometric person authentication using new spectral features of electroencephalogram (EEG)". International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS2008), 2008

[38] S. Sanei and J. Chambers, EEG signal processing, 2007.

[39] R. B. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, "The electroencephalogram as a biometric, Proc. Canadian Conf". Electrical and Computer Eng., vol. 2, 2001.

[40] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou, "Person identification based on parametric processing of the EEG", Electronics, Circuits and Systems, 1999. Proceedings of ICECS ’99. The 6th IEEE International Conference on, vol.1, 1999.

[41] M. A. Lebedev, and M. A.L. Nicolelis, Brain-machine interfaces: past, present and future. TRENDS in Neurosciences: Elsevier, 2006.

[42] B. Blankertz et.al., The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG Single Trials, 2003.