Subspace techniques for task-independent EEG person identification

Subspace techniques for task-independent EEG person identification

Mari Ganesh Kumar[1], Saranya M S[1], Shrikanth Narayanan[2], Mriganka Sur[3] and Hema A. Murthy[1]

***Abstract*— There has been a growing interest in studying electroencephalography signals (EEG) as a possible biometric. The brain signals captured by EEG are rich and carry information related to the individual, tasks being performed, mental state, and other channel/measurement noise due to session variability and artifacts. To effectively extract person-specific signatures present in EEG, it is necessary to define a subspace that enhances the biometric information and suppresses other nuisance factors. i-vector and x-vector are state-of-art subspace techniques used in speaker recognition. In this paper, novel modifications are proposed for both frameworks to project person-specific signatures from multi-channel EEG into a subspace. The modified i-vector and x-vector systems outperform baseline i-vector and x-vector systems with an absolute improvement of 10.5% and 15.9%, respectively.**

***Index Terms*— EEG, biometric, i-vector, x-vector**

I. Introduction

In the literature, different elicitation protocols have been a primary focus for building reliable systems. A detailed review of these different techniques can be found in [1], [2]. Although biometric systems are typically trained and tested on the same set of tasks, it has been observed that the person-specific information is present in the EEG signal across various elicitation protocols [3], [4]. This leads to a conjecture that these person-specific signatures should be present in the EEG signal irrespective of tasks or state of the brain. In [5], this conjecture was verified under closed eyes condition by using different tasks/elicitation protocols during train and test condition. In this work, we extend it to the open eye condition as well and show that individuals can be identified irrespective of the stimuli shown. To accomplish this, we build upon and extend the latest advances in total variability modeling in speaker recognition, namely, i-vectors and x-vector based signal representations [6], [7] to EEG biometrics.

In any EEG based biometric system, the efficacy of the system should be cross-validated by retaining a few sessions for testing purpose. Different days of recording in which EEG sensors are removed and placed afresh are referred to as sessions in this paper. This testing across sessions is essential because the electrodes are unlikely to be placed at the same place on the scalp across sessions. In addition to this, the plastic nature of the brain and the mental state of the individual can also change the EEG signatures across sessions. Person identification systems using EEG perform poorly when tested on sessions that were not used for training [5], [8]. This paper evaluates the scalability of person identification irrespective of the tasks/elicitation protocol with only mismatched train and test sessions.

Universal background model-Gaussian mixture model (UBM-GMM) [9] systems have been widely used for EEG person identification [5], [10]–[13]. In addition to the person-specific information present in the EEG, the UBM also models the task-specific information along with the artifacts. This paper explores different techniques to define a subspace in which the person-specific biometric information is enhanced. i-vector system is one such technique proposed for speaker recognition in [6] and adopted for EEG person identification in [14]. Similarly, x-vector is a deep neural network (DNN) based speaker recognition approach proposed in [7]. i-vector extracts the person-specific signatures from the UBM space, whereas x-vector derives the same directly from the input spectrograms. However, both of these approaches are defined for speech data where speaker information is typically bound within a channel.

EEG collects signals from multiple regions through the sensors placed all over the scalp. It is well-known that different parts of the brain respond to different types of stimuli. Hence, person-specific signatures might be scattered across multiple channels. Moreover, the effective use of multichannel information for EEG person identification is still not well established. The objective of this paper is two fold. First, to incorporate multi-channel information in the i-vector and the x-vector framework to better identify individuals from EEG. Second, to show that biometric signatures can be reliably identified from EEG in a subspace where the task related information is suppressed.

The rest of the paper is organized as follows. Section II discusses the details of the baseline and proposed systems. The EEG dataset and the experimental setup that is used to evaluate the person identification systems are briefly discussed in Section III and Section IV, respectively. The results of the evaluation are analyzed in Section V, followed by conclusions in Section VI.

II. Methods

A. Baseline: UBM-GMM

UBM is a GMM trained using feature vectors pooled from training sessions of all the subjects. Person-specific models are obtained by adapting the UBM to a particular individual's data using maximum-a-posteriori (MAP) adaptation. Similar to [5], [14], feature vectors are pooled from every channel, and no explicit channel information is given to the model.

[1]Mari Ganesh Kumar, Saranya M S and Hema A. Murthy are with the Department of Computer Science and Engineering, Indian Institute of Technology Madras {mari,saranms,hema}@cse.iitm.ac.in

[2]Shrikanth Narayanan is with the Viterbi School of Engineering, University of Southern California shri@ee.usc.edu

[3]Mriganka Sur is with the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology msur@mit.edu

TABLE I

DETAILS OF DATA COLLECTION PROTOCOLS.

No.	Experiment Type	Description	No. of subjects
1	Odd Ball	Subjects were presented with frequent non-target stimuli and infrequent target stimuli. The target and non-target stimuli consist of 1. Audio beeps of two different frequencies; 2. Audio beeps played in left and right ear and 3. Visual objects of varying shape and color.	23
2	Familiar and Unfamiliar words	Subjects were presented with common words and uncommon words. They were expected to respond with a mouse click on hearing a familiar word.	7
3	Imagining binary answers	A set of questions with the answer being either yes or no were presented to subjects. They were asked to imagine the answer and then respond with a mouse click. Left click was used for positive responses and right click for negative responses.	7
4	Motor and Mental imaginary	Subjects were asked to imagine motor-movements like left and right fist rotation. Followed by this, they were asked to count numbers in reverse for mental activity. This experiment aimed to classify motor vs. mental task.	6
5	Passive audio	In this experiment type, subjects listened passively to a variety of audio stimuli such as words, sentences, stories, music, and sounds that trigger attention (for example sirens and the scattering of glass).	24
6	Steady State Visually Evoked Potential	Visual objects were displayed at various frequencies to subjects. At the end of each trial, a question about the shape or color of the object was asked.	10
7	Passive audio-visual	Subjects were asked to watch audio-visual clips. At the end of each clip, a question was asked based on the stimuli shown.	7

During testing, Top-C scoring is used to determine the identity of the EEG segment [9].

B. i-vector

i-vector is a lower dimensional representation of the person-specific information modeled in UBM space. A total variability space (T) matrix is used to project the information from UBM space to i-vector space.

Let λ denote the parameters of the UBM and $Y = \{\, \bar{y}_n^c \mid n = 1 \textit{ to } N \;\&\; c = 1 \textit{ to } C\}$ denote an EEG segment with N time frames and C channels. The zeroth and first statistics for every mixture component for given Y can be estimated as given in Equation 1 and 2, respectively.

$$N_k(Y) = \sum_{c=1}^{C}\sum_{n=1}^{N} P(k \mid \bar{y}_n^c,\ \lambda) \tag{1}$$

$$F_k(Y) = \sum_{c=1}^{C}\sum_{n=1}^{N} P(k \mid \bar{y}_n^c,\ \lambda)(\bar{y}_n^c - \bar{m}_k) \tag{2}$$

where $k = 1 \textit{ to } K$ represent the mixture IDs of the UBM and $P(k \mid \bar{y}_n^c,\ \lambda)$ corresponds to the posterior probability of the k-th mixture component given the feature vector y_n^c. $\bar{m}_k$ is the mean of the k-th UBM component. These first order statistics (F_k) are concatenated to form a supervector of dimension $Kd \times 1$ where d is the dimension of the input feature vector. A T-Matrix of dimension $Kd \times R$ is used to project the supervector to a lower dimension (R). The procedure for training the T-Matrix and thereby extracting i-vectors can be found in [6]. This system will be referred to as "baseline-i-vector" in the rest of the paper. This standard approach for multichannel EEG [14] is perhaps not appropriate, as this is not equivalent to collecting speaker data through different channels or environments. Different stimuli activate different regions of the brain. These activations are captured by the EEG sensors placed on the corresponding regions. To integrate channel information in the i-vector framework, this paper proposes a novel way of finding the zeroth and first statistics from UBM as given in Equation 3 and 4, respectively.

$$N_{kc}(Y) = \sum_{n=1}^{N} P(k \mid y_n^c,\ \lambda) \tag{3}$$

$$F_{kc}(Y) = \sum_{n=1}^{N} P(k \mid y_n^c,\ \lambda)(y_n^c - m_k) \tag{4}$$

Since in this approach the zeroth and first order statistics are calculated for every channel, the dimensions of the supervector and T-matrix increases to $KCd \times 1$ and $KCd \times R$, respectively. This system will be referred to as "modified-i-vector". Owing to the huge dimensions of the supervectors this system uses a significantly lower number of mixtures for UBM. It is better to concatenate the channel wise statistics rather than concatenating the channel feature vectors at the input level. The latter system has a huge input dimension and leads to a poor estimation of UBM statistics. After extracting the i-vectors, within class covariance normalization (WCCN) and Linear Discriminant Analysis (LDA) are performed to reduce nuisance factors in i-vector space. Finally, a one-vs-all support vector machine (SVM) [15] with a cosine kernel is trained to predict the class labels.

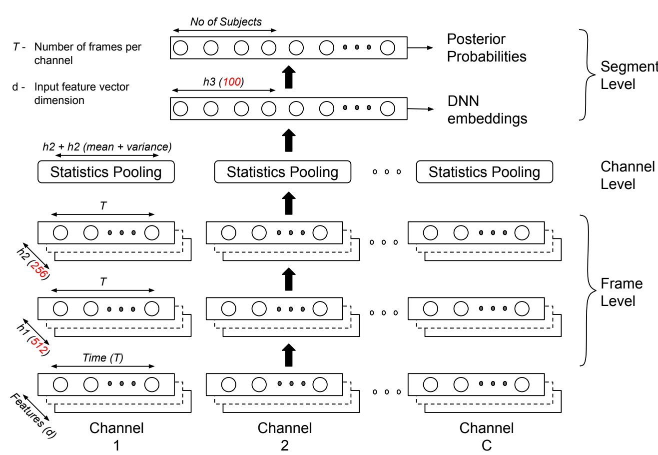

Fig. 1. Architecture of the modified x-vector model for EEG person identification. The values in red denote the specific number of hidden nodes used for this implementation.

C. x-vector

x-vector is a DNN based state-of-the-art speaker recognition technique [7]. The DNN architecture process the input speech first at frame level and then at segment level after accumulating statistics. This is analogous to the i-vector framework with UBM acting at the frame level and i-vectors operating at the segment level. We propose a modification to the x-vector architecture for EEG data by estimating the statistics channel wise similar to Section II-B. x-vectors use time delay neural network (TDNN) for modeling context in speaker recognition. As the sampling rate of EEG is low (250 Hz), a frame size of 360ms is used. Hence for EEG data, the context information did not help, and 1D convolution of a single frame is used in place of TDNN.

The modified x-vector architecture is given in Figure. 1. The model takes spectrograms of all channels as input. There are four hidden layers in this model. The initial two layers of the network act at the frame level, i.e., 1D convolution on all the frames. Followed by the convolution layers, the statistical pooling layer acts at the channel level, where the mean and variance for every channel is estimated. The fourth hidden layer takes the statistics concatenated from all the channels and projects it onto a lower dimension using a feedforward layer. All the hidden layers in this model use sigmoid activation. The final output feedforward layer use softmax activation. The network is trained using the cross-entropy error function similar to [7].

The outputs of the network are the posterior probabilities for each person given the EEG segment. Since the goal is to identify the persons, the posteriors are directly used to get the final class labels. This system will be referred to as "modified-x-vector-classifier". In addition to this, the fourth hidden layer ("DNN embeddings" in Figure. 1) gives a lower dimensional embedding of the person-specific information. LDA is computed over these embeddings, and a one-vs-all SVM classifier [15] is built using a cosine kernel to predict the person IDs. This system will be referred to as "modified-x-vector". The advantage of this system is that it can be used in a verification framework as well. The x-vector classifier without any channel concatenation is used as a baseline. This is referred to as "baseline-x-vector-classifier" in the remainder of the paper.

III. Dataset

The dataset used in this work is an expanded and modified version of the dataset used in [5]. A 128 channel EEG system provided by Electrical Geodesics, Inc (EGI) [16] was used to collect the dataset. This study was approved by the Ethics Committee of the Indian Institute of Technology Madras. All the subjects were informed about the aim and scope of the experiment, and written consent was also obtained to collect the data. It consists of EEG data from 30 participants performing multiple tasks. The dataset was collected at a sampling rate of $250Hz$. The total duration of the collected data is about 33 hours. Multiple sessions of data were collected from all the individuals. The average number of sessions per person in this dataset is ≈ 3.

Different stimuli were used to collect the dataset and details of the same are given in Table. I. Protocols 2 to 5 are experiments with only audio stimuli, and the participants were asked close their eyes. Visual stimuli were used in protocols 6 and 7, and hence the recordings were collected in open eye condition. Protocols 1 and 5 are a collection of different experiments that fall under the same type, and hence they have a significant number of participants compared to other protocols. These diverse set of protocol cover various tasks or mental states with both eyes open and closed conditions.

IV. Features and Experimental Setup

Using 128-Channel EEG system is not feasible for building real time biometrics. Hence, we use only 8 standard electrodes, namely, Fz, F7, F8, C3, C4, P7, P8 and Oz. These 8 channels were chosen such that they cover the entire scalp. Raw power spectral density (PSD) features are computed for every channel between $3Hz$ and $30Hz$. The spectrograms are estimated with a window size of $360ms$ and no overlap. This configuration was fine-tuned for the UBM-GMM system in [5]. For recordings with open eye conditions, eye blink and other artifacts are removed using [17].

The brain signals obtained from different experimental protocols in Table. I are uniformly divided into segments of uniform length. It is to be noted that, this division of segments does not take into account the cognitive state, such as, whether the person is in resting state or listening to an audio stimulus or watching a visual stimulus or doing an instructed task. Hence, identifying individuals from these segments is evidence towards the presence of person-specific signatures in EEG irrespective of tasks.

For every individual, 60% of the sessions are chosen randomly for training. Of the remaining sessions data, 10% is used for validation, and the rest is used for testing.

TABLE II

Accuracy (%) of different systems for $15s$ segments.

System	P1	P2	P3	Average
UBM-GMM	67.15	68.68	70.09	68.63
baseline-i-vector	73.57	74.52	73.98	74.02
baseline-x-vector-classifier	61.28	64.76	64.48	63.48
modified-i-vector	**82.53**	**87.54**	**83.71**	**84.59**
modified-x-vector	77.51	81.3	77.85	78.89
modified-x-vector-classifier	78.12	81.7	78.36	79.42

Pi represents the i-th random partition of train, validation, and test data.

V. Results and Discussion

A 128 mixture UBM was used to train baseline UBM-GMM and i-vector systems. For the modified-i-vector system, only 16 mixtures were used. i-vector dimension was empirically set to 100. For x-vector systems, the complete architecture used in this implementation is given in Figure. 1. All these systems were trained using task-independent EEG segments of $15s$ duration. The classification accuracy of these systems on held out sessions are compared in Table. II.

The results are reported over three random partitions of train, validation, and test to avoid bias towards a particular division.

From Table. II, it can be observed that all the modified systems that handle channel information have outperformed the baseline systems. On the averaged result, the modified-i-vector system achieves an absolute improvement of 10.5% over the baseline i-vector system. The modified x-vector systems are also robust compared to the baseline systems which takes no channel information. Although all the systems are trained on $15s$ segment length, they are capable of supporting variable length segments. To determine the optimal length of EEG required for identifying individuals, the length of the EEG segments is varied from $10s$ to $40s$. Similar to Table. II, results for various length of EEG segments are averaged across three random splits and shown in Figure. 2.

From Figure. 2 it can be seen that the performance of all systems improves when the size (duration) is increased. However, above 30 seconds, the performance get saturated. The most important observation from both Figure. 2 and Table. II is that all the modified systems that use the channel information have outperformed the simple systems borrowed directly from speaker recognition. This result suggests the importance of using the multi-channel information to model person-specific signatures. Instead of concatenating the channels at the input level, this paper proposes a novel way of handling it in i-vector and x-vector frameworks. Initially, the input feature vectors from all channels are transformed to a UBM or a higher dimensional space. In the transformed space, different orders of statistics are calculated for each channel. These statistics from different channels are then concatenated and projected to a lower dimensional space.

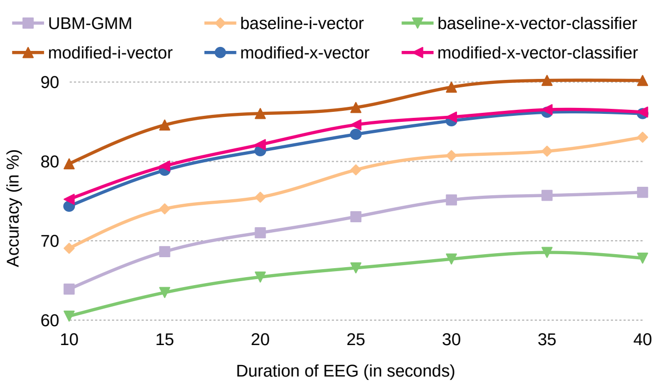

Fig. 2. Accuracy of person identification systems for various size (duration) of EEG signals.

For EEG segments greater than $30s$ in duration, the modified-i-vector system gives an accuracy of around 90% (Figure. 2). To recall, these segments were created by dividing the EEG responses irrespective of task protocols and only held out session are used for evaluation. Hence this result is a strong evidence for the presence of individual related signatures in EEG, irrespective of tasks or state of the brain.

VI. Conclusion

The paper proposes novel subspace techniques for task-independent EEG person identification. The proposed systems are inspired by state-of-the-art sub-space based voice biometric techniques. Using just 8 channels, the subspace systems are shown to reliably model the person-specific signatures in a vector of small dimension from the multi-channel EEG data. By testing on segments with mismatched task and session during testing, this paper also shows that the person-specific signatures are always present in EEG irrespective of the task.

References

[1] P. Campisi and D. L. Rocca, "Brain waves for automatic biometric-based user recognition," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 782–800, 2014.

[2] M. Del Pozo-Banos, J. B. Alonso, J. R. Ticay-Rivas, and C. M. Travieso, "Electroencephalogram subject identification: A review," *Expert Systems with Applications*, vol. 41, no. 15, pp. 6537–6554, 2014.

[3] M. Del Pozo-Banos, C. M. Travieso, J. B. Alonso, and A. John, "Evidence of a task-independent neural signature in the spectral shape of the electroencephalogram," *International Journal of Neural Systems*, vol. 28, no. 01, pp. 1750035:1–20, 2018.

[4] E. Maiorana and P. Campisi, "Longitudinal Evaluation of EEG-Based Biometric Recognition," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1123–1138, May 2018.

[5] Vinothkumar D, M. G. Kumar, A. Kumar, H. Gupta, M. S. Saranya, M. Sur, and H. A. Murthy, "Task-Independent EEG based Subject Identification using Auditory Stimulus," in *Proc. Workshop on Speech, Music and Mind 2018*, 2018, pp. 26–30.

[6] N. Dehak, P. J. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-End Factor Analysis for Speaker Verification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 4, pp. 788–798, May 2011.

[7] D. Snyder, D. Garcia-Romero, G. Sell, D. Povey, and S. Khudanpur, "X-Vectors: Robust DNN Embeddings for Speaker Recognition," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5329–5333, 2018.

[8] Y. Höller, A. C. Bathke, and A. Uhl, "Age, Sex, and Pathology Effects on Stability of Electroencephalographic Biometric Features Based on Measures of Interaction," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 459–471, Feb 2019.

[9] D. A. Reynolds, T. F. Quatieri, and R. Dunn, "Speaker verification using adapted Gaussian Mixture Models," *Digital Signal Processing*, vol. 10, no. 1-3, pp. 19–41, 2000.

[10] S. Marcel and J. D. R. Millan, "Person Authentication Using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 743–752, April 2007.

[11] P. Nguyen, D. Tran, T. Le, X. Huang, and W. Ma, "EEG-based person verification using multi-sphere SVDD and UBM," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2013, pp. 289–300.

[12] S. Altahat, M. Wagner, and E. Martinez Marroquin, "Robust electroencephalogram channel set for person authentication," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2015, pp. 997–1001.

[13] P. Davis, C. D. Creusere, and J. Kroger, "Subject identification based on EEG responses to video stimuli," in *2015 IEEE International Conference on Image Processing*, Sept 2015, pp. 1523–1527.

[14] C. Ward, J. Picone, and I. Obeid, "Applications of UBMs and I-vectors in EEG subject verification," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2016, pp. 748–751.

[15] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[16] "Electrical Geodesics, Inc," https://www.egi.com/.

[17] C. Chang, S. Hsu, L. Pion-Tonachini, and T. Jung, "Evaluation of artifact subspace reconstruction for automatic eeg artifact removal," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2018, pp. 1242–1245.