Person Authentication using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation

Sébastien Marcel [a]   José del R. Millán [a]

IDIAP–RR 05-81

December 2005

TO APPEAR IN
IEEE Transactions on Pattern Analysis and Machine Intelligence
– Special Issue on Biometrics 2007

[a] IDIAP Research Institute

IDIAP Research Institute **www.idiap.ch**
Rue du Simplon 4 P.O. Box 592 1920 Martigny – Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

IDIAP Research Report 05-81

Person Authentication using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation

Sébastien Marcel José del R. Millán

December 2005

to appear in
IEEE Transactions on Pattern Analysis and Machine Intelligence – Special Issue on Biometrics 2007

Abstract. In this paper, we investigate the use of brain activity for person authentication. It has been shown in previous studies that the brain-wave pattern of every individual is unique and that the electroencephalogram (EEG) can be used for biometric identification. EEG-based biometry is an emerging research topic and we believe that it may open new research directions and applications in the future. However, very little work has been done in this area and was focusing mainly on person identification but not on person authentication. Person authentication aims to accept or to reject a person claiming an identity, i.e comparing a biometric data to one template, while the goal of person identification is to match the biometric data against all the records in a database.

We propose the use of a statistical framework based on Gaussian Mixture Models and Maximum A Posteriori model adaptation, successfully applied to speaker and face authentication, which can deal with only one training session. We perform intensive experimental simulations using several strict train/test protocols to show the potential of our method. We also show that there are some mental tasks that are more appropriate for person authentication than others.

1 Introduction

An *authentication* (or *verification*) system involves confirming or denying the identity claimed by a person (one-to-one matching). In contrast, an *identification* system attempts to establish the identity of a given person out of a closed pool of N people (one-to-N matching). Authentication and identification share the same preprocessing and feature extraction steps and a large part of the classifier design. However, both modes target distinct applications. In authentication mode, people are supposed to cooperate with the system (the claimant wants to be accepted). The main applications are access control systems (airport checking, monitoring, computer or mobile devices log-in), building gate control, digital multimedia access, transaction authentication (in telephone banking or remote credit card purchases for instance), voice mail, or secure teleworking. On the other hand, in identification mode, people are generally not concerned by the system and often even do not want to be identified. Potential applications includes video surveillance (public places, restricted areas) and information retrieval (police databases, video or photo album annotation/identification). Such authentication systems are based on the characteristics of a person, such as face, voice, fingerprint, iris, gait, hand geometry or signature. A good introduction to person authentication can be found in [15].

In this paper, we investigate the use of brain activity as a new modality for person authentication. This modality has several advantages: (1) it is confidential (as it corresponds to a mental task), (2) it is very difficult to mimic (as similar mental tasks are person dependent) and (3) it is almost impossible to steal (as 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).

Monitoring the brain activity in order to design future man-machine interfaces is the aim of Brain Computer Interfaces (BCI) [6, 17]. A BCI may monitor brain activity via a variety of methods, which can be coarsely classified as invasive and non-invasive. Given the risks generated by permanent surgically implanted devices in the brain, and the associated ethical concerns, we concentrate only on non-invasive approaches, in particular electrical brain signals as measured by electroencephalogram (EEG); i.e., the electrical brain activity recorded from electrodes placed on the scalp. The main source of the EEG is the synchronous activity of thousands of cortical neurons. Measuring the EEG is a simple non-invasive way to monitor electrical brain activity, but it does not provide detailed information on the activity of single neurons (or small brain areas). Moreover, it is characterized by small signal amplitudes (a few Volts) and noisy measurements (especially if recording outside shield rooms). Besides electrical activity, neural activity also produces other types of signals, such as magnetic and metabolic, that could be used in a BCI. Magnetic fields can be recorded with magnetoencephalography (MEG), while brain metabolic activity – reflected in changes in blood flow – can be observed with positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and optical imaging. Unfortunately, such alternative techniques require sophisticated devices that can be operated only in special facilities. Moreover, techniques for measuring blood flow have long latencies and thus are less appropriate for interaction.

It has been shown in previous studies that the brain-wave pattern of every individual is unique and that the electroencephalogram (EEG) can be used for biometric identification. We believe that EEG-based biometry is an emerging research topic and that it may open new research directions and applications in the future. Unfortunately, EEG signal is known to be very noisy and difficult to process.

Very little work has been done in this area [12, 8, 9] and was focusing mainly on person identification but not on person authentication. Poulos and al. [12] have proposed to model the EEG signal using autoregressive (AR) models and then to use the parameters of the AR model for the identification. The classification is performed using Kohonen's Vector Quantizer (VQ). Poulos and al. tried to differentiate four subjects individually from a pool of different individuals. Paranjape and al. [8] proposed also to represent the EEG signal (from the single P4 electrode) using AR models, then discriminant analysis is employed to perform the classification. More recently, Palaniappan and al [9] investigated features based on the spectral power of the signal together with a fuzzy Neural Network for the classification.

The paper is structured as follow. In the next section, we first introduce the reader to the problem of person authentication and we present the proposed approach based on Gaussian Mixture Models and Maximum A Posteriori model adaptation. Then, we describe the database we used and the different experiment protocol. Finally, we present the results obtained using our approach and conclude.

2 The Proposed Approach

2.1 Problem Description

An identity authentication system has to deal with two kinds of events: either the person claiming a given identity is the one who he claims to be (in which case, he is called a *client*), or he is not (in which case, he is called an *impostor*). Moreover, the system may generally take two decisions: either *accept* the *client* or *reject* him and decide he is an *impostor*.

We propose to adopt a statistical framework widely used in other biometric authentication approaches such as speaker authentication [13] or face verification [2]. In this framework, one first needs a probabilistic model (see section 2.2) of *anybody*'s biometric data, often called a *world model* and trained on a large collection of recordings of several people. From this generic model, a more specific, client-dependent model, is then derived using adaptation techniques (see section 2.4), built on data from a particular client. One can then estimate the ratio of the likelihood of the data corresponding to some access with respect to the model of the claimed client identity, with the likelihood of the same data with respect to the *world model*. The access is accepted or rejected (see section 2.3) if the likelihood ratio is higher or lower than a given threshold, selected in order to optimize either a low rejection rate, a low acceptance rate, or a combination of both.

2.2 Gaussian Mixture Models

Let us note the biometric data (extracted from the EEG signal) as a sequence ($X_1^T = \{\mathbf{x}_1 ... \mathbf{x}_T\}$) of frames, where $\mathbf{x}_t \in \mathbb{R}^D$ and D is the number of features per frame.

In the Gaussian Mixture Model (GMM) approach, all feature vectors are assumed to be independent. Given the GMM parameter set λ, the likelihood of a set of T feature vectors $X = \{\mathbf{x}_t\}_{t=1}^{T}$ is found with

$$P(X|\lambda) = \prod_{t=1}^{T} P(\mathbf{x}_t|\lambda) \tag{1}$$

where

$$\begin{aligned} P(\mathbf{x}|\lambda) &= \sum_{k=1}^{N} w_k \, \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k) && (2) \\ \lambda &= \{w_k, \mu_k, \Sigma_k\}_{k=1}^{N} && (3) \end{aligned}$$

Here, $\mathcal{N}(\mathbf{x}|\mu, \Sigma)$ is a D-dimensional Gaussian density function [4] with mean μ and diagonal covariance matrix Σ. N is the number of Gaussians and w_k is the weight for Gaussian k (with constraints $\sum_{k=1}^{N} w_k = 1$ and $\forall \; k : w_k \geq 0$).

2.3 Application to Person Authentication

Let us denote the parameter set for client C as λ_C, and the parameter set describing a generic non-client as $\neg\lambda_C$. Given a claim for client C's identity and a set of feature vectors X supporting the claim, we find an opinion $\Lambda(X)$ on the claim using:

$$\Lambda(X) = \log P(X|\lambda_C) - \log P(X|\neg\lambda_C) \tag{4}$$

where $P(X|\lambda_C)$ is the likelihood of the claim coming from the true claimant and $P(X|\neg\lambda_C)$ is the likelihood of the claim coming from an impostor.

The above probabilities are represented by diagonal Gaussian Mixture Models. The generic EEG model is trained using data from many people. Finally, the authentication decision is reached as follows: given a threshold τ, the claim is accepted when $\Lambda(X) \geq \tau$ and rejected when $\Lambda(X) < \tau$.

2.4 Training

We can use different ways to train each client model. Traditional Maximum Likelihood (ML) training, such as Expectation-Maximization, can be used [3, 4]. Maximum A Posteriori (MAP) training [5] can also be used to adapt a generic model using client data. Indeed, it has been previously shown that the traditionally used ML training approach has problems estimating robust model parameters when there are only a few training data available. More precise models can be obtained through the use of MAP.

Given a set of training vectors, X, the probability density function (pdf) $P(X|\lambda)$ and the prior pdf of λ, $P(\lambda)$, the MAP estimate of model parameters, $\lambda_{\texttt{MAP}}$, is defined as:

$$\begin{aligned} \lambda_{\texttt{MAP}} &= \arg\max_{\lambda} P(\lambda|X) && (5) \\ &= \arg\max_{\lambda} P(X|\lambda)P(\lambda) && (6) \end{aligned}$$

Assuming λ to be uniform is equivalent to having a non-informative $P(\lambda)$, reducing the solution of $\lambda_{\texttt{MAP}}$ to the standard ML solution. Thus, the difference between ML and MAP training is in the definition of the prior distribution for the model parameters to be estimated. It has been observed that MAP based training obtains best performance when only the means are adapted (rather than adapting the covariance matrices and weights). We thus choose to adapt only the means.

An implementation of MAP training for client model adaptation consists of using a global parameter to tune the relative importance of the prior. The equation for adaptation of the means is:

$$\hat{\mu}_k = \alpha \mu_k + (1-\alpha) \frac{\sum_{t=1}^{T} P(k|\mathbf{x}_t)\,\mathbf{x}_t}{\sum_{t=1}^{T} P(k|\mathbf{x}_t)} \quad (7)$$

here $\hat{\mu}_k$ is the new mean of the k-th Gaussian, μ_k is the corresponding parameters in the generic model, $P(k|\mathbf{x}_t)$ is the posterior probability of k-th Gaussian (from the client model from the previous iteration) and $\alpha \in [0,1]$ is the adaptation factor chosen empirically.

3 Experimental Protocol

3.1 Database

EEG signals were recorded with a Biosemi system using a cap with 32 integrated electrodes located at standard positions of the International 10-20 system. The sampling rate was 512 Hz. Signals were acquired at full DC. No artifact rejection or correction was employed.

This dataset contains data from 9 normal subjects during 12 non-feedback sessions over 3 days (4 sessions per day). The subject sat in a normal chair, relaxed arms resting on their legs. There are 3 tasks:

1. Imagination of repetitive self-paced left hand movements, (left),
2. Imagination of repetitive self-paced right hand movements, (right),
3. Generation of words beginning with the same random letter, (word).

For all sessions of a given subject acquired on the same day (each lasting 4 minutes with 5-10 minutes breaks in between them), the subject performed a given task for about 15 seconds and then switched randomly to another task at the operator's request. EEG data can then be splitted into segments corresponding to a given mental task. Each segment is considered as a record. There are 3 records per sessions.

3.2 Preprocessing and Feature extraction

Raw EEG potentials are too noisy and variable to be analyzed directly. Thus the first step is to preprocess them to increase their signal-to-noise ratio and extract relevant features that better describe the mental states to be recognized. The raw EEG potentials were first spatially filtered by means of a surface Laplacian (SL). This operation yields new potentials that represent better the cortical activity due only to local sources below the electrodes. The superiority of SL-transformed over raw potentials for the recognition of mental tasks has been demonstrated in different studies [1, 7]. Specifically, we first interpolated using spherical splines of order 2 and then took the second spatial derivative which is sensitive to localized sources of electrical activity [10, 11]. The second derivative is evaluated only at the 8 locations of the electrodes.

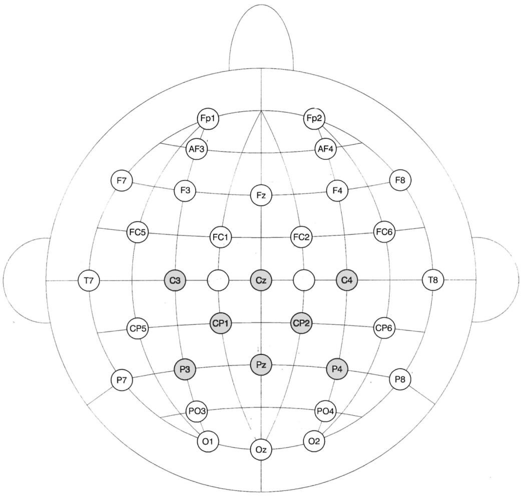

Figure 1: Illustration of the location of electrodes on the scalp. Electrodes we are using are indicated in gray.

Then, every 62.5 ms –i.e., 16 times per second– the power spectral density (PSD) in the band 8-30 Hz was estimated for the 8 centro-parietal channels C3, Cz, C4, CP1, CP2, P3, Pz, and P4 (Fig. 1). The PSD features we extract from the 8 SL-transformed electrode signals are based on a temporal Fourier transform. To estimate the power spectrum of each channel over the last second we used the Welch periodogram algorithm [16]. Specifically, we averaged the FFT of 3 segments of 0.5 second with 50% overlap, which yields a frequency resolution of 2 Hz. The values in the frequency band 8-30 Hz were normalized according to the total energy in this same band. As a result, an EEG sample is a 96-dimensional vector (8 channels times 12 frequency components). It is worth noting that, for our experimental protocol, PSD features lead to better or similar performances than more elaborated features such as parameters of autoregressive models and wavelets [14].

The choice of the electrodes and frequency band is based on the expertise available in the BCI community that shows that they contain most of the relevant information for the recognition of the mental tasks used for this study (for a review see [6, 17]). Similarly, the reason for the fast computation of the PSD-based EEG samples (16 times per second using windows of 1 second) is to fit the real-time constraints of a BCI.

3.3 Experimental Methodology

Regarding the fact that our database is small, we have (1) to design carefully several experimental protocols based on distinct training/validation/evaluation sets and (2) to perform several simulations. Therefore, we propose 4 different protocols:

- to evaluate the potential of our method for person authentication on a small dataset (*protocol 1*),
- to confirm the previous findings on a larger dataset and to measure the performance degradation over days (*protocol 2*),
- to demonstrate that training with data spawn over several days improves the performance (*protocol 3*),
- to show the benefit of incremental learning (*protocol 4*).

3.4 Performance Evaluation

Authentication systems make two types of errors: a False Acceptance (FA), which occurs when the system accepts an impostor, or a False Rejection (FR), which occurs when the system refuses a true claimant. The performance is generally measured in terms of False Acceptance Rate (FAR) and False Rejection Rate (FRR) expressed in percentages. To aid the interpretation of performance, the two error measures are often combined using the Half Total Error Rate (HTER), defined as:

$$\text{HTER} = (\text{FAR} + \text{FRR})\,/2$$

The verification decision is then reached as follows:

- the claim is accepted when $\Lambda(X) \geq \tau$,
- the claim is rejected when $\Lambda(X) < \tau$.

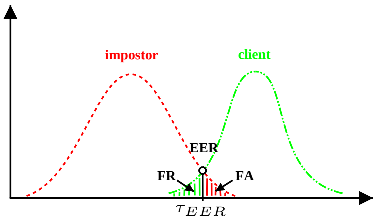

Figure 2: Illustration of typical errors of a biometric system. An impostor above the threshold is a false acceptance. A client below the threshold is a false rejection.

Since in real life the decision threshold τ has to be chosen *a priori*, this threshold is chosen to optimize a given criterion, such as the Equal Error Rate (EER), i.e when $FAR = FRR$ (Fig. 2), on the validation set. This threshold is then used on the evaluation set to obtain a HTER figure.

3.5 Experimental Protocols

3.5.1 Protocol P1

The purpose of this first experimental protocol is to evaluate the potential of the proposed method on a small dataset. It is based on a cross-validation scheme with distinct training/validation and evaluation sets. Among the 9 initial subjects, 3 subjects are kept. Only the sessions of the first day are used in this protocol.

Table 1 describes the usage of different sessions in each configuration. The notation C/I means that a session can be used to access a model as a client as well as an impostor. As an example, let us consider Kfold1. Data from the session 1 of person 1 and 2 are used to train the world model and the client specific models (1 and 2). Data from the session 2 of person 1 and 2 are used to compute client and impostor scores (validation set) when testing against client models (1 and 2). Additionally to supplement the validation set, data from the session 1 and 2 of impostor 3 are used to compute impostor scores against client models (1 and 2). Finally, the evaluation set is obtained in a similar way but this time using sessions 3 and 4.

Table 1: Usage of sessions for the 3-Kfold protocol P1.

person	session	Kfold 1			Kfold 2			Kfold 3		
		T	V	E	T	V	E	T	V	E
1	1	C			C				I	
	2		C/I			C/I			I	
	3			C/I			C/I			I
	4			C/I			C/I			I
2	1	C				I		C		
	2		C/I			I			C/I	
	3			C/I			I			C/I
	4			C/I			I			C/I
3	1		I		C			C		
	2		I			C/I			C/I	
	3			I			C/I			C/I
	4			I			C/I			C/I

Then, for each Kfold, we have the following number of accesses:

- validation set: 8 accesses made of 2 client accesses and 6 impostor accesses (including 2 sessions of an impostor unseen during the training to access the 2 client models),
- evaluation set: 12 accesses made of 4 client accesses and 8 impostor accesses.

It is worth noting that despite the small number of available subjects, we have designed a hard experiment protocol where one of the subject, out of three, was always removed from the training data and used as an impostor during evaluation. We decided to design an experiment protocol based on a 3 K-folds scheme. In each K-fold, every person is, in turn, considered as a client or an impostor.

Table 2: Usage of sessions for protocols P2, P3 and P4.

person	session	P2			P3			P4		
		T	V	E	T	V	E	T	V	E
clients (2, 3, 4, 5, 7, 8)	1	C			C			C		
	2	C			C			C		
	3		C/I			C/I			C/I	
	4			C/I		C/I				C/I
	5			C/I	C			C^{d+1}		
	6			C/I	C					C/I
	7			C/I		C/I				C/I
	8			C/I		C/I				C/I
	9			C/I			C/I	C^{d+2}		
	10			C/I			C/I			C/I
	11			C/I			C/I			C/I
	12			C/I			C/I			C/I
impostor 1	1		I			I			I	
	2		I			I			I	
impostors 6, 9	1			I			I			I
	2			I			I			I

3.5.2 Protocol P2

The goal of this protocol is to confirm the findings of protocol P1 on a larger dataset and to measure the performance degradation over days. Among the 9 subjects (Table 2), 3 are considered as real impostors (persons 1, 6 and 9) and the 6 remaining are considered as clients. Real impostors are used to compute impostor accesses. Impostor 1 is used on the validation set and impostors 6 and 9 are used on the evaluation set. Sessions 1-2 are used for client training, session 3 for client/impostor validation and session 4 for client/impostor evaluation. Sessions 5 to 8 will be used for client/impostor day 2 evaluation. Sessions 9 to 12 will be used for client/impostor day 3 evaluation.

3.5.3 Protocol P3

We expect protocol P3 to demonstrate that training with data spawn over several days increases the performance. Therefore, we will use half of day 1 sessions (1-2) and half of day 2 sessions (5-6) for client training. The second half of days 1 and 2 will be used for client/impostor validation (sessions 3-4 and 7-8). All sessions from day 3 (9 to 12) will be used for client/impostor evaluation.

3.5.4 Protocol P4

Finally, the protocol P4 will try to show the benefit of incremental learning. This protocol is very similar to protocol P2. The only difference is that sessions 5 and 9 are kept for incremental client training and then are not available for client/impostor evaluation.

4 Results

4.1 Results on Protocol P1

We provide in Table 3, HTER results obtained on the described database according to the above experiment protocol P1 on the evaluation set. The authentication was performed for each mental task. We present also the results for each K-fold (K1, K2 and K3) and the average over the 3 K-fold using 5 different values for the number of Gaussians in the mixture. Each value is the average of 100 simulations with different initial conditions[1].

Table 3: HTER performance (in %) for each mental task and protocol P1

Mental tasks	Number of Gaussians	K-folds			
		K1	K2	K3	Avg
left	4	15.6	9.5	8.6	11.2
	8	13.8	5.5	4.0	7.8
	16	15.4	2.1	2.4	**6.6**
	32	20.1	0.5	5.5	8.7
	64	14.6	1.2	12.9	9.5
right	4	22.8	6.2	28.7	19.2
	8	12.4	3.0	20.5	12.0
	16	23.7	6.6	7.4	12.6
	32	29.1	10.1	9.7	16.3
	64	27.7	13.6	20.3	20.5
word	4	12.6	50.0	15.8	26.1
	8	19.0	5.9	11.4	12.1
	16	27.6	2.1	19.6	16.4
	32	25.7	0.0	16	13.9
	64	22.3	0.0	23.4	15.2

These results suggest that EEG signal is an effective modality for person authentication and that the GMM/MAP framework can be a good choice for this task. These results also show that not all mental tasks are equally appropriate for person authentication. Results can even improve if, for each person, a different mental task were used – as if each person had his/her individual "mental password". The best result was obtained with the "left" mental task. Interestingly, the three persons in the database were right-handed. It is also worth noting that the optimal number of Gaussians is rather small (8 or 16): a small number fails to capture the complexity of the data distribution while a larger number seems to model noise. However, no conclusions can be drawn on such a small number of individuals.

[1] This is also true for all experiments in this paper.

4.2 Results on Protocol P2

We provide in Table 4, FAR/FRR and HTER results obtained according to the experiment protocol P2 on the evaluation set. The authentication was only performed for mental tasks “left” and “right” as the previous experiment (Section 4.1) suggested that those tasks were more appropriate than the mental task “word”. We present also the results for each day (d1, d2 and d3) using 4 different values for the number of Gaussians in the mixture.

First of all, these results confirm that EEG signal is an effective modality for person authentication and that the GMM/MAP framework is a good choice for this task. We have also the confirmation that the mental task “left” is better suited than the mental task “right” on this database. Also, we observe the degradation of performance over days 2 and 3.

Obviously, the mismatch between testing and training increases from days to days. Therefore, data collected only over one day is not enough for training robust models.

Interestingly, we see also on days 2 and 3 that the FAR is much lower than the FRR, while the decision threshold was optimized at the EER on the validation set of day 1 only. This is a clear indication of the robustness of the system because despite the high false rejection rate of clients, it keeps a small false acceptance rate of impostors. The system, however, needs a better fine tuning to model intra-class variability over time.

4.3 Results on Protocol P3

We provide in Table 4, FAR/FRR and HTER results obtained according to the experiment protocol P3 on the evaluation set. Again, the authentication was only performed for mental tasks “left” and “right”. The reader should keep in mind that in this protocol the evaluation set corresponds to the day 3 as parts of days 1 and 2 were used for training and validation (Section 3.5.3).

From the results, we can conclude that the performance can be improved by using training/validation data over 2 days. Both GMM parameters and decision threshold can be estimated more accurately. We reached nearly 2.5 improvement between protocol P2 (35.5 % average HTER over P2-d2 and P2-d3) and protocol P3 (12.9 % HTER). This suggests that even much better results can be achieved by using training data over all days and that there might be a potential for incremental learning.

4.4 Results on Protocol P4

We provide in Table 5, FAR/FRR and HTER results obtained according to the experiment protocol P4 on the evaluation set. The authentication is only performed on mental task “left” and using 2 values for the number of Gaussians (the one providing the best results in the previous experiment). The purpose of this protocol is to convince the reader that there is a potential for incremental learning. Here of course, we are making a strong assumption, i.e. that the first session of days 2 and 3 can be trusted (the identity of the claimant is known) and used for training (Section 3.5.4). This training is called incremental because client models

Table 4: FAR/FRR/HTER performance (in %) for mental tasks “left” and “right”, and protocol P2 and P3

.

Mental tasks	Number of Gaussians	Protocol											
		P2-d1			P2-d2			P2-d3			P3		
		FAR	FRR	HTER	FAR	FRR	HTER	FAR	FRR	HTER	FAR	FRR	HTER
left	4	15.1	17.2	16.1	19.6	50.3	**34.9**	24.8	47.6	**36.2**	18.6	32.3	25.4
	8	12.4	17.6	15.0	17.9	64	40.9	25.6	56.6	41.1	23.8	25.15	24.5
	16	9.0	15.9	12.4	11.1	79.9	45.5	13.8	71.6	42.7	19.3	19.65	19.5
	32	5.7	8.5	**7.1**	7.2	82.2	44.7	8.3	93.7	51.0	13.7	24.9	**19.3**
right	4	14.3	8.5	11.4	21.3	49.3	**35.3**	24.3	60.9	**42.6**	18.4	40.5	29.4
	8	10.2	13.9	12.0	14.6	73.8	44.2	18.8	74.0	46.4	20.6	29.5	25.0
	16	7.2	9.7	**8.4**	10.7	75.9	43.3	12.4	86.3	49.3	15.0	23.6	**19.3**
	32	4.8	15.9	10.3	4.8	82.6	43.7	7.6	95.5	51.5	13.0	30.15	21.6

are re-trained completely from all training data (incrementally stored) and are not re-adapted using new data samples. Furthermore, the decision threshold is not re-estimated.

We first report the error rate on the evaluation set of days 1, 2 and 3 (depicted as P4-d1, P4-d2, P4-d3). Second, we report the error rate on the evaluation set of days 2 and 3 (depicted as P4^{d+1}-d2, P4^{d+1}-d3) after re-training of the client models using sessions 1, 2 and 5 (session 5 being the first of day 2). Finally, we report the error rate on the evaluation set of day 3 (depicted as P4^{d+2}-d3) after re-training of the client models using sessions 1, 2, 5 and 9. We should notice first that results for P4-d1, P4-d2 and P4-d3 are very similar to P2. It is logical, since results should not be much affected by removing one testing session from days 2 and 3.

Table 5: FAR/FRR/HTER performance (in %) for mental task “left” and protocol P4

Number of Gaussians	Protocol								
	P4-d1			P4-d2			P4-d3		
	FAR	FRR	HTER	FAR	FRR	HTER	FAR	FRR	HTER
4	15.1	17.2	16.1	20.0	50.5	**35.3**	24.7	46.8	**35.7**
32	5.7	8.5	**7.1**	7.3	82.7	45.0	8.3	96.0	52.1
	Protocol								
	P4^{d+1}-d2			P4^{d+1}-d3			P4^{d+2}-d3		
	FAR	FRR	HTER	FAR	FRR	HTER	FAR	FRR	HTER
4	24.9	2.7	13.8	29.4	10.6	20.0	29.3	1.2	15.25
32	16.0	0.2	**8.1**	17.8	28.3	**23.0**	24.5	0.02	**12.3**

Secondly, we observe the effectiveness of incremental learning. Indeed, a day-to-day comparison of results under protocols P4^{d+1}-d2 and P4-d2 or under protocols P4^{d+2}-d3 and P4-d3 shows an improvement of the HTER of a factor 2. A closer look shows that this improvement is mainly due to the reduction of the FRR. Therefore, intra-class variability is better modeled. Furthermore, we can notice that the results obtained under P4^{d+1}-d3 are nearly as good as for protocol P3 (Table 4). Again, this shows the effectiveness of incremental learning because in protocol P3, sessions 5-6 are used for training (in addition to sessions 1-2) and sessions 7-8 are used also for validation, while in P4^{d+1}-d3 the session 5 only is used for model training and the decision threshold is not re-estimated. Therefore, we can confirm that there is a large potential for incremental learning. Its benefit should be even larger in the case of doing also decision threshold re-estimation.

5 Conclusion and Future Work

In this paper, we investigated the use of brain activity for person authentication. We proposed the use of a statistical framework based on Gaussian Mixture Models and Maximum A Posteriori model adaptation. We performed intensive experimental simulations using strict train/test protocols to show the potential of our method. We also show (1) that there are some mental tasks that are more appropriate for person authentication than others, (2) that the performance degrades over days, (3) that using training data over two days increases the performance and (4) that there is a potential for incremental learning.

However, the database we used is still small and no definite conclusive lessons can be learned for the task of person authentication from the results reported here. We plan to collect a more appropriate database with more clients and impostors, and where various real-world scenarios and mental tasks will be investigated. We will be able to test several state-of-the-art biometric authentication algorithms and to propose others on the light of experimental findings.

It should be noted also that all the choices made for the preprocessing and feature selection algorithm used here were based on studies seeking a different goal, namely recognition of mental tasks from EEG. Thus, a subject that deserves further investigation is the exploration of alternative choices better suited for person authentication.

Acknowledgment

The authors would like to thank the Swiss National Science Foundation for supporting this work through the National Center of Competence in Research (NCCR) on ”Interactive Multimodal Information Management (IM2)”.

References

[1] F. Babiloni, F. Cincotti, L. Lazzarini, J.d.R. Millan, J. Mourino, M. Varsta, J. Heikkonen, L. Bianchi and M.G. Marciani, “Linear classification of low-resolution EEG patterns produced by imagined hand movements”, *IEEE Trans. on Rehabilitation Engineering*, vol. 8, pp. 186-188, 2000.

[2] F. Cardinaux, C. Sanderson and S. Marcel, “Comparison of MLP and GMM Classifiers for Face Verification on XM2VTS,” *Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication*, pp. 911-920, 2003.

[3] A.P. Dempster, N.M. Laird and D.B. Rubin. “Maximum-likelihood from incomplete data via the EM algorithm” *Journal of Royal Statistical Society*, Series B (Methodological), vol. 39, no. 1, pp. 1-38, 1977.

[4] R.O. Duda, P.E. Hart and G.S David, *Pattern Classification*, Wiley, 2001.

[5] J.L. Gauvain and C.-H. Lee, “Maximum a posteriori estimation for multivariate Gaussian mixture observation of Markov chains,” *IEEE Transactions on Speech Audio Processing*, vol. 2, pp. 291-298, 1994.

[6] J.d.R. Millán, “Brain-computer interfaces,” *Handbook of Brain Theory and Neural Networks*, ed. M.A. Arbib: MIT Press, Cambridge Massachusetts, 2002.

[7] J. Mourino, “EEG-based analysis for the design of adaptive brain interfaces”, *Ph.D. thesis, Centre de Recerca en Enginyeria Biomedica,* Universitat Politecnica de Catalunya, Barcelona, Spain, 2003.

[8] R.B. Paranjape, J. Mahovsky, L. Benedicenti and Z. Koles. “The Electroencephalogram as a Biometric,” *Proceedings of the Canadian Conference On Electrical And Computer Engineering,* vol. 2, pp. 1363-1366, 2001.

[9] R. Palaniappan and K.V.R. Ravi, “A new method to identify individuals using signals from the brain,” *Proceedings of the 4th International Conference on Information Communications and Signal Processing,* Singapore, pp. 15-18, 2003

[10] F. Perrin, J. Pernier, O. Bertrand and J. Echallier, “Spherical spline for potential and current density mapping,” *Electroencephalography and Clinical Neurophysiology,* vol. 72, pp. 184-187, 1989.

[11] F. Perrin, J. Pernier, O. Bertrand and J. Echallier, “Corrigendum EEG 02274,” *Electroencephalography and Clinical Neurophysiology,* vol. 76, pp. 565, 1990.

[12] M. Poulos, M. Rangoussi, V. Chrissicopoulos and A. Evangelou, “Person identification based on parametric processing on the EEG,” *Proceedings of the Sixth International Conference on Electronics, Circuits and Systems,* vol. 1, pp. 283-286.

[13] D.A. Reynolds, T.F. Quatieri and R.B. Dunn, “Speaker Verification Using Adapted Gaussian Mixture Models,” *Digital Signal Processing,* vol. 10, no. 1-3, 2000.

[14] M. Varsta, J. Heikkonen, J.d.R. Millan and J. Mourino, “Evaluating the performance of three feature sets for brain-computer interfaces with an early stopping MLP”, *Proceedings 15th Int. Conf. on Pattern Recognition,* pp. 911-915, 2000.

[15] P. Verlinde, G. Chollet and M. Acheroy, “Multi-modal identity verification using expert fusion,” *Information Fusion,* vol. 1, pp. 17-33, 2000.

[16] P.D. Welch, “The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms,” *IEEE Trans. Audio Electroacoustics,* vol. AU-15, pp. 70-73, 1967.

[17] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller and T.M. Vaughan, “Brain-computer interfaces for communication and control, ” *Clinical Neurophysiology,* vol. 113, pp. 767-791, 2002.