Gaussian Mixture Models for CHASM Signature Verification

Andreas Humm, Jean Hennebert, and Rolf Ingold

Université de Fribourg, Boulevard de Pérolles 90, 1700 Fribourg, Switzerland
andreas.humm,jean.hennebert,rolf.ingold@unifr.ch

Abstract. In this paper we report on first experimental results of a novel multimodal user authentication system based on a combined acquisition of online handwritten signature and speech modalities. In our project, the so-called CHASM signatures are recorded by asking the user to utter what he is writing. CHASM actually stands for Combined Handwriting and Speech Modalities where the pen and voice signals are simultaneously recorded. We have built a baseline CHASM signature verification system for which we have conducted a complete experimental evaluation. This baseline system is composed of two Gaussian Mixture Models sub-systems that model independently the pen and voice signal. A simple fusion of both sub-systems is performed at the score level. The evaluation of the verification system is conducted on CHASM signatures taken from the MyIDea multimodal database, accordingly to the protocols provided with the database. This allows us to draw our first conclusions in regards to time variability impact, to skilled versus unskilled forgeries attacks and to some training parameters. Results are also reported for the two sub-systems evaluated separately and for the global system.

1 Introduction

Multimodal biometrics has raised a growing interest in the industrial and scientific community. The potential increase of accuracy combined with better robustness against forgeries makes indeed multimodal biometrics a promising field. In our work, we are interested in building multimodal authentication systems using speech and signatures as modalities. Speech and signatures are indeed two major modalities used by humans in their daily transactions and interactions. On the one hand, handwritten signatures are nowadays legally and socially accepted means for user authentication and contractual terms acceptation. On the other hand, producing a speech signal is a very natural non-intrusive gesture.

1.1 Motivations

Many automated biometric systems based on signature or speech alone have been studied and developed [9] [16]. However, we still see few deployments in commercial applications. We have identified three major reasons for this:

S. Renals, S. Bengio, and J. Fiscus (Eds.): MLMI 2006, LNCS 4299, pp. 102–113, 2006.

(1) negative impact of time-variability [11], (2) degraded performances in the case of trained forgeries [10] [18], (3) decreased performances in mismatched conditions of use, such as mismatched sensors or mismatches environments [18]. While points (2) and (3) can be somehow handled with a minimum of supervision and control of the acquisition environment, the first point mentioned above has a critical impact for institutions willing to deploy such biometrics. Indeed, repeated enrollment sessions are not at all convenient for the user and generate further costs as they need to be secured.

We propose here a new approach to circumvent these problems while keeping an acceptable solution for the end user. The proposal is to record bimodal signatures by asking the user to simultaneously say and write the signature. Such bimodal signatures have already been presented in our preliminary work *Combined Handwriting and Speech Modalities for User Authentication* [6] and are referred here as *CHASM signatures*. In a similar way, we have also defined *CHASM handwriting* where the user reads what he is writing. CHASM handwriting could be used for user authentication or for enhanced content recognition, but this is out of the scope of this paper where we focus on CHASM signatures.

The main motivation of CHASM is therefore to increase performance and robustness by using two modalities instead of one. This work and our future work will attempt to assess in which degree CHASM signatures authentication systems can reach this goal. The motivation of performing a synchronized acquisition is multiple. Firstly, it avoids doubling the acquisition time. Secondly, the synchronized acquisition will probably give better robustness against intentional imposture as imitating simultaneously the voice and the writing of somebody else has a larger cognitive load. Finally, the synchronization patterns (i.e. where do users synchronize) or the intrinsic deformation of the inputs (mainly the slowdown of the speech) may be dependent to the user, therefore bringing useful biometrics information.

1.2 Related Work

Several related works have already shown that using speech and signature modalities together permits to improve significantly the authentication performances in comparison to systems based on speech or signature alone.

In [8], a tablet PC system based on online signature and voice modalities is proposed to ensure the security of electronic medical records. Tablet PCs are already used by many health care professional to have a patient's record readily available when prescribing or administering treatment. In this system, the user claims his identity by saying his first and last name that are recognized using speech recognition. The same waveform is then used with a speaker verification system based on GMMs to produce a score. In this way, the identification and verification steps are performed simultaneously. A signature is then acquired and a dynamic time warping verification system is used to produce a score. Speech and signature scores are then normalized and fused.

In [3], an online signature verification system and a speaker verification system are also combined. Both sub-systems use Hidden Markov Models (HMMs) to produce independent scores that are then fused together. Results are reported for the two sub-systems evaluated separately and for the global system. Better accuracy is reported for the fused bimodal system. For this test, fictitious users are built by randomly associating signature and speech samples from two independent databases, namely *Philips' online signature database* and *Polyphone* and *Polyvar*.

In [11], tests are reported for a system where the signature verification part is built using HMMs and the speaker verification part uses either dynamic time warping or GMMs. The fusion of both systems is performed at the score level and results are again better than for the individual systems. This last work uses the BIOMET database [4] where the speech and signature data are recorded from the same user.

The main difference between these works and our CHASM approach lies in the acquisition procedure. In our case, the speech and signature data streams are recorded simultaneously, asking the user to actually say the content of his signature. Our procedure has the advantage to shorten the enrollment and access time and will potentially allow for more robust fusion strategies upstream in the processing chain.

The remainder of this paper is organized as follows. In section 2, we give an overview of the CHASM signature database used in this work and of the evaluation protocols. In section 3 we introduce GMMs and how they are used to model the speech and signature data streams. Section 4 presents the experimental results of the evaluation of our CHASM signature verification system. Finally, conclusions, discussions and future work are presented.

2 CHASM Signature Database

In this section we describe the database that we used to conduct the evaluation. Some comments on CHASM signature data are given and the evaluation protocols are then described.

2.1 MyIDea Database

CHASM data have been acquired in the framework of the MyIDea biometric data collection [2] [5]. MyIDea database is a multimodal database that contains other modalities such as fingerprint, talking face, palm print, etc. MyIDea contains about 70 users that have been recorded over three sessions spaced in time. In MyIDea, CHASM data have been recorded according to two scenarios. In the first one, a bimodal signature with voice is acquired. In this case, the user is actually asked to say the content of his signature, - *CHASM signature*. In the second scenario, the user is asked to write and read synchronously the content of a text, - *CHASM handwriting*. The data set used to perform the experiments reported in this article has been given the reference MYIDEA-CHASM-SET1 by the distributors of MyIDea. This set should be considered as a development set. A second set of CHASM data is planned for acquisition in a near future and will be used as evaluation set.

In MyIDea, CHASM data have been acquired with a WACOM Intuos2 graphical tablet and a standard computer headset microphone (Creative HS-300). For the signature stream, x,y-coordinates, pressure and the azimuth and elevation angles of the pen are sampled at 100 Hz. The speech waveform is recorded at 16 kHz and coded linearly on 16 bits. The data samples are also provided with timestamps to allow a precise synchronization of both streams. The timestamps are especially important for the signature streams as the graphical tablet does not send data samples when the pen is out of range. Fig. 1 shows an example of CHASM signature. The grey areas on the figure correspond to inter-stroke moments, when the user lift the pen out of the range of the tablet. We have to note that these kind of events are not very frequent for signatures and are more frequent for handwriting.

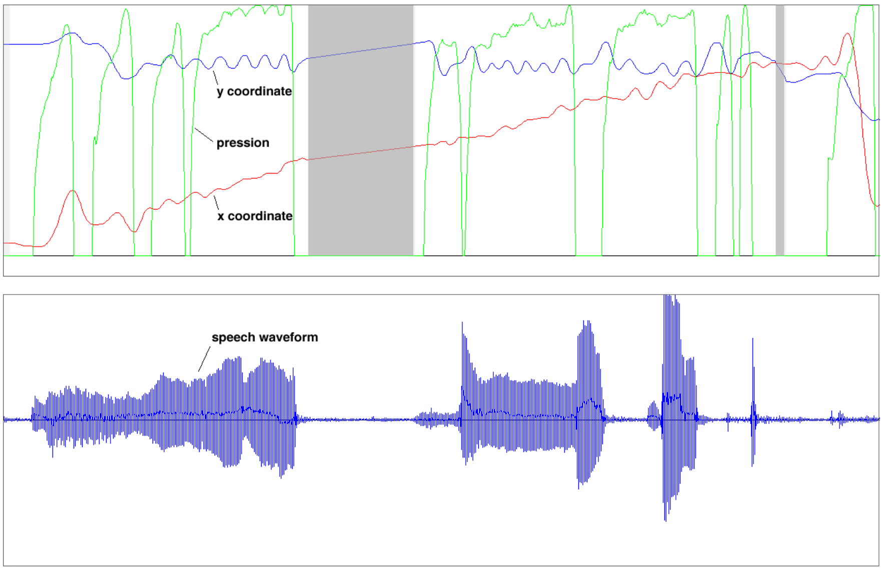

Fig. 1. Synchronized visualization of handwriting and speech signals. Azimuth and elevation angles are not displayed for sake of clarity. The upper part of the graph shows the evolution of x and y coordinates and the pression p. The bottom part shows the speech amplitude. On this visualisation, all signals are synchronized thanks to the timestamps.

2.2 Comments on CHASM Signature Data

We performed a visual inspection of CHASM signatures for several different users. Strokes and acoustic events are of course not always synchronized in the same way. In most of the cases, a given acoustic event either is synchronized with the stroke either starts slightly after the stroke (see Fig. 1). In some realizations, the speech event starts slightly before the stroke. The average synchronization times correspond roughly to syllables when the signature contains clear sequences of letters, which is the case for most of the signatures. These observations are in accordance with the acquisition protocol of MyIDea where the subjects were asked to speak in such a way that the sounds correspond roughly in time with the written symbols.

Flourishes are usually present in signatures, most frequently at the end of signatures. When flourishes are happening, a large majority of users are not producing acoustic events on top of them (as illustrated on the last stroke in the example of Fig. 1). If the signature contains only flourishes or non-readable signs, the subject was simply asked to utter his name while signing. In this case, there is no specific synchronization of acoustic and stroke events.

In our previous work [6], we report on a usability survey conducted on the subjects that took part to MyIDea recordings. The main conclusions of the survey are the following. First, all recorded users were able to perform the signature acquisition. Speaking and signing at the same time did not prevent any acquisition to happen. Second, the survey shows that simultaneous acquisitions are acceptable from the user point of view.

2.3 Evaluation Protocols

In MyIDea, six *genuine* CHASM signatures are acquired for each subject per session. This leads to a total of 18 true acquisitions after the three sessions. After acquiring the genuine signatures, the subject is also asked to imitate six times the signature of another subject. Imitations are performed by letting the subject having an access to the *static* image and to the *verbal content* of the signature to be forged. In other words, access to the voice recording is not given to perform the forgery. This procedure leads to a total of 18 *skilled forgeries*[1] after the three sessions, i.e. six impostor signatures on three different subjects.

CHASM signature assessment protocols have been defined on MyIDea [6]. The protocols have been crafted to be as realistic as possible and to put in evidence difficulties tied to time variability. Two protocols have been defined:

- **Without time variability.** For each subject in the database, models are built using three spoken signatures sampled randomly out from the six genuine accesses of the first session. For testing, the three remaining signatures of the first session are used. The same procedure is repeated for sessions two and three, leading to a total of 70 users * 3 accesses * 3 sessions = 630 *genuine* tests. Two kinds of impostor attempts are considered: *random forgeries* and *skilled forgeries*. In the case of random forgeries, impostor attempts are performed using one signature for each of the remaining subjects

[1] The term *skilled forgeries* is used here to somehow comply with the nomenclature used in the literature about signature verification systems. However, one should note that there is no trained imitation of the speech signal as only the content is reproduced with no intentions to imitate the genuine voice. For the speech part, the term *passive* or *content-based* forgeries could be then more adequate.

in the database, giving a total of 70 users * 69 accesses * 3 sessions = 14490 random forgeries. In the second case, the 18 available skilled forgeries are used against each user, giving a total of 70 users * 18 accesses * 3 sessions = 3780 skilled forgeries.
- **With time variability.** For each subject, the six signatures from the first session are used to build the models. Genuine tests are performed on the six signatures of session two and session three, giving a total of 70 users * 12 accesses = 840 genuine tests. Random and skilled impostor attempts are performed in the similar manner as for the protocol *without time variability* with the distinction that models are here trained on the first session only, giving a total of 70 users * 69 accesses = 4830 random forgeries and 70 users * 18 accesses = 1260 skilled forgeries.

The amounts of tests mentioned above are approximative as some users did not complete all sessions.

3 System Description

We have chosen to use standard GMMs to model independently both streams of data, followed by a simple fusion at the score level (see Fig. 2). While this system uses straightforward feature extraction and modelling, it will allow us to validate the evaluation protocol and to draw our first conclusions regarding the impact of time-variability and skilled vs random forgeries. Performances are also measured on the speech stream alone (1), the signature stream alone (2) and on the fused systems (3).

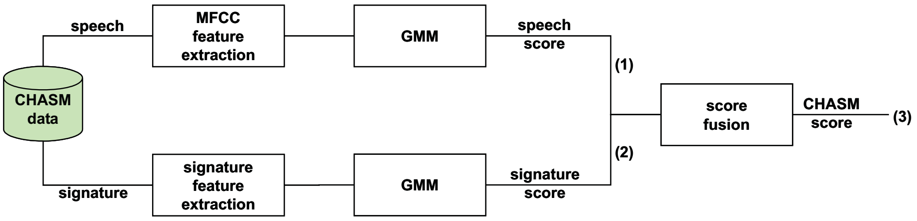

Fig. 2. Baseline CHASM signature verification system

3.1 Signature Features

For each point of the signature, we extract 25 dynamic features in a similar way as in [12]:

- the absolute speed and acceleration, the speed and acceleration in x and y directions and the tangential acceleration
- the angle α of the absolute speed vector, its cosine and sine, the derivative of α and its cosine and sine

- the pressure and the pressure derivative
- the azimuth and elevation angles of the pen and their derivatives
- the curvature radius
- the normalized coordinates $(x(n) - x_g, y(n) - y_g)$ relatively to the gravity center (x_g, y_g) of the signature
- the length to width ratio of windows of 5 and 7 points centered on the current point and the ratio of the minimum over the maximum speed on a window of 5 points centered on the current point.

The features are mean and standard deviation normalized on a per user basis.

3.2 Speech Features

We use Mel Frequency Cepstral Coefficients (MFCC) as features [14]. The frontend's frame size is 25.625 ms and the frame shift is 10 ms. The frontend extracts 12 MFCC coefficients and the energy. An energy-based speech detection module based on a bi-Gaussian model is applied to remove the silence from the data. MFCC coefficients are mean and standard deviation normalized using normalization values computed on the speech part of the data. We performed experiments including delta and delta-delta coefficients without further improvements of thr results. These features were then left apart in our baseline configuration for which results are reported here.

3.3 GMMs System

GMMs are used to model the likelihoods of the features extracted from the signature and from the speech signal. One could argue that GMMs are actually not the most appropriate models in this case as they are intrinsically not capturing the time-dependant specificities of signatures. HMMs would be potentially more adequate in this case. However, GMMs have been reported to compare reasonably well to HMMs in terms of signature verification [17] and are often considered as baseline systems in speaker verification. Furthermore, GMMs are well-known flexible modelling tools able to approximate any probability density function. With GMMs, the probability density function $p(x_n|M_{client})$ or *likelihood* of a D-dimensional feature vector x_n given the model of the client M_{client}, is estimated as a weighted sum of multivariate Gaussian densities (see e.g. [15]):

$$p(x_n|M_{client}) = \sum_{i=1}^{I} w_i \mathcal{N}(x_n, \mu_i, \Sigma_i) \tag{1}$$

in which I is the number of mixtures, w_i is the weight for mixture i and the Gaussian densities $\mathcal{N}$ are parameterized by a mean $D \times 1$ vector μ_i, and a $D \times D$ covariance matrix, Σ_i:

$$\mathcal{N}(x_n, \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} exp\left(-\frac{1}{2}(x_n - \mu_i)'\Sigma_i^{-1}(x_n - \mu_i)\right) \tag{2}$$

In our case, we make the hypothesis that the features are uncorrelated and we use diagonal covariance matrices. By making the hypothesis of observation independence, the global *likelihood* score for the sequence of feature vectors, $X = \{x_1, x_2, ..., x_N\}$ is computed with:

$$S_c = p(X|M_{client}) = \prod_{n=1}^{N} p(x_n|M_{client}) \tag{3}$$

The likelihood score S_w of the hypothesis that X is **not** from the given client is here estimated using a world model M_{world} or *universal background model* trained by pooling the data of many other users. The likelihood S_w is computed in a similar way, by using a weighted sum of Gaussian mixtures. The optimal decision whether to reject or to accept the claimed user is performed comparing the ratio of client and world score against a global threshold value T. The ratio is often computed in the log-domain with:

$$R_c = \log(S_c) - \log(S_w) \tag{4}$$

The training of the client and world models is performed with the Expectation-Maximization (EM) algorithm [1] that iteratively refines the component weights, means and variances to monotonically increase the likelihood of the training feature vectors. The client and world model are trained independently by applying iteratively the EM procedure until convergence is reached, typically after few iterations. In our setting, we apply a simple binary splitting procedure to increase the number of Gaussian components to a predefined value. The world model is trained by pooling all the available genuine accesses in the database. The skilled forgeries attempts are excluded for training the world model as it would lead to optimistic results. Ideally, a fully independent set of users would be preferable, but this is not possible considering the small number of users (≈ 70) available.

3.4 Score Fusion

In our baseline system, we fuse the two modalities by simply summing the signature and the speech log-likelihood ratios with $R_{c,CHASM} = R_{c,speech} + R_{c,signature}$ which is a reasonable procedure if we assume that the local observations of both sub-systems are independent. This is however clearly not the case as the users are intentionally trying to synchronize their speech with the signature signal. Time-dependent score fusion procedures or feature fusion followed by joint modelling would be more appropriate than the approach taken here. These approaches are part of our future work. Also, more advanced score recombination or classification strategies could also be applied such as, for example, using a weighted sum of the likelihood or using classifier-based score fusion [7]. However, such fusion methods require parameters estimation on an independent development set which is currently not available. We will then report here fusion results using the simple summation as described above. We also report results using a *z-norm* score normalization preceding the summation. The z-norm is here applied globally on both speech and signature scores, in a user-independent way. Such a normalization procedure makes sense if $R_{c,speech}$ and $R_{c,signature}$ are distributed according to a Gaussian distribution. As the mean and standard deviation of the z-norm are estimated a posteriori on the same data set, z-norm results are unrealistic but give however an optimistic estimation of what could be the performances.

4 Results

Results of biometric systems are classically measured in terms of impostor False Acceptation FA and client False Rejection FR error rates that vary as a function of the decision threshold T. Operating points (FA,FR) can then be plot on a (x,y) figure with T as parameter. Detection Error Tradeoff (DET)[13] plots are often used in which the x and y axis follow a normal deviate scale. If the scores are normally distributed the DET curve will be close to a straight line, enabling easy observation of system contrasts. We also report our results in terms of Equal Error Rates (EER) which are obtained for $FA = FR$.

For all the results reported here, we have used 64 mixtures in our world models. Experiments with different world model sizes have been conducted, leading to the conclusion that 64 mixtures give good results for all protocols. This number could probably be different if more or less training data would be available. Table 1 shows the evolution of the EER as a function of the number of mixtures in the client models, using protocol *with time variability and random forgeries*. We tested with 8, 16, 32, 64 and 128 Gaussian mixtures and concluded that, on average for all users, the optimal number of mixtures lie around 16 mixtures. Similar conclusions were obtained for the other protocols. For an improved version of the system, one could compute an optimized number of Gaussian mixtures for each user and modality such as taking into account the length of the signatures [11] or computing a cost functions that balances modelling errors and model complexity [17]. In the rest of this section, we report results with 16 Gaussian mixtures for the client models.

Table 1. Equal Error Rate (EER) as a function of the number of Gaussian mixtures in the client models, protocol with time variability

number of mixtures	8	16	32	64	128
signature	7.3	6.4	7.4	8.6	11.0
speech	12.0	12.2	14.1	14.8	15.0
sum fusion	4.7	4.7	5.8	6.8	8.0

Figure 3 illustrates the DET curve of the speech system, the signature system and the z-norm fusion of both systems for the protocol *with* (left part) and *without* (right part) time variability, and using random forgeries. We can conclude from this figure that the speech modelisation performs better than the signature for single session experiments. However, when multi-session accesses are considered, signature performs better than speech. Signature and speech modalities suffer from time-variability but in different degrees. The speech modality seems to be more sensitive to time variability than the signature modality. The z-norm fusion brings a clear amelioration of the results for both protocols.

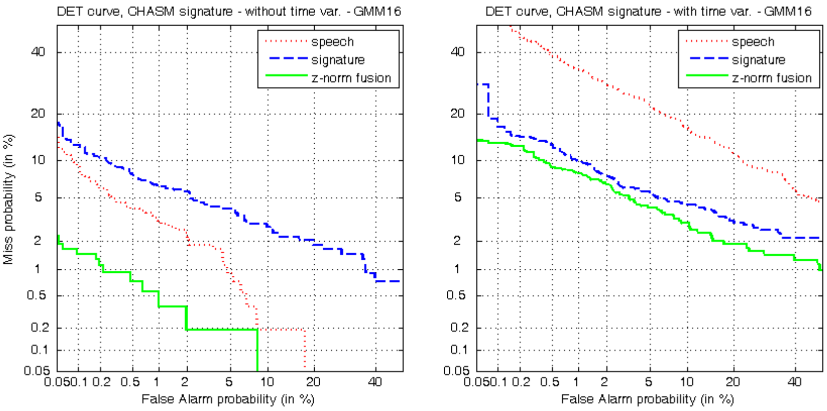

Fig. 3. DET curve - fusion of the signature and speech GMM systems, protocol without time variability (left) and with time variability (right), random forgeries

Table 2 summarizes the results in terms of ERR for the different protocols. The following conclusions can be drawn. For the skilled forgeries protocol, a drop of 3.3% of absolute performance is observed (9.4 - 6.1) when testing on signatures acquired in different sessions as the enrollment. For the speech modality, the impact is even more important with an absolute decrease of about 15% of the EER (19.5 - 3.7). Such a drop in the performance can be due to several factors. First, the modelling technique we used may not be robust enough against time variabilities. Using a MAP adaptation of the world model to build client models would be a potential amelioration in this regards. Second, it is probable that users show a larger intra-variability for the speech than for the signature modality. Third, the speech modelisation may suffer from variabilities of the acquisition conditions: different position of the headset-mounted microphone, environmental noise, etc., while the signature acquisition is more stable.

Another conclusion is that skilled forgeries decreases systematically and significantly the performance in comparison to random forgeries. For the protocol *with time variability*, a drop of almost 100% relative performance is observed for the signature modality and about 50% for the speech modality. We have to note again here that the forger do not try to imitate the voice of the user but actually say the genuine verbal content.

The sum fusion, although very straightforward, brings systematically a clear improvement of the results. These results are in favor of the CHASM methodology. Interestingly, the z-norm fusion is better than the sum fusion for the protocol without time variability and is worse in the case of the protocol with time variability. A visual analysis of the score distribution of both modalities, before z-norm and after z-norm, lead us to a potential intuitive interpretation of this behavior. The application of the z-norm is, by nature, aligning the score distributions of both modalities. While this is good to fuse scores that lies in different ranges, the z-norm is also giving equal importance to each modalities. This is of course not favorable in the case of systems showing very different individual performances.

Table 2. Protocol with and without time variability, 16 Gaussian mixtures for the client GMMs, 64 mixtures for the world

time variability	without (%EER)		with (%EER)	
forgeries	random	skilled	random	skilled
signature	4.0	6.1	5.3	9.4
speech	2.0	3.7	14.0	19.5
sum fusion	**1.7**	**3.1**	**3.5**	**6.9**
z-norm fusion	**0.6**	**1.3**	**4.1**	**8.7**

5 Conclusions and Future Work

A baseline verification system using GMMs for modelling CHASM signatures has been presented. Results obtained with this system show that the use of both modalities outperforms these modalities used alone. Results also show that there is a clear impact of time variability and skilled forgeries on the performances. In our future work, we plan to investigate the use of more robust modelling techniques against time variability and forgeries. In this direction, we have identified potential modelling techniques such as MAP adaptation of the world GMMs, user-dependent model order, HMMs, time-dependent score fusion, fusion at the feature level followed by joint modelling, etc. As soon as an extended set of CHASM signature data will be available, experiments will be conducted according to a development/evaluation set framework. Another part of our future work will be to investigate CHASM handwriting to build verification systems.

Acknowledgments

This work was mainly supported by the Swiss National Science Foundation program "Interactive Multimodal Information Management (IM2)", as part of NCCR and by the EU BioSecure NoE project. We warmly thank Asmaa El Hannani for her precious feedbacks when we were experimenting with GMMs based systems.

References

1. A.P. Dempster, N.M. Laird, and Rubin D.B. Maximum likelihood from incomplete data via the em algorithm. *Journal of Royal Statistical Society*, 39(1):1–38, 1977.
2. B. Dumas et al. Myidea - multimodal biometrics database, description of acquisition protocols. In *In proc. of Third COST 275 Workshop (COST 275)*, pages 59–62, October 27 - 28 2005. Hatfield (UK).
3. M. Fuentes et al. Identity verification by fusion of biometric data: On-line signature and speech. In *Proc. COST 275 Workshop on The Advent of Biometrics on the Internet*, pages 83–86, November 2002. Rome, Italy.
4. S. Garcia-Salicetti et al. Biomet: a multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In *4th AVBPA*. Springer-Verlag, 2003.
5. J. Hennebert, B. Dumas, C .Pugin, F. Evéquoz, A. Humm, D. Petrovska-Delacrétaz. MyIDea home page. http://diuf.unifr.ch/go/myidea, 2005.
6. A. Humm, J. Hennebert, and R. Ingold. Combined handwriting and speech modalities for user authentication. Technical Report 06-05, University of Fribourg, Department of Informatics, 2006.
7. A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern Recognition*, 38:2270–2285, 2005.
8. S. Krawczyk and A. K. Jain. Securing electronic medical records using biometric authentication. In *Audio- and Video-based Biometric Person Authentication (AVBPA)*, pages 1110–1119, Rye Brook, NY, 2005.
9. F. Leclerc and R. Plamondon. Automatic signature verification: the state of the art–1989-1993. *Int'l J. Pattern Recog. and Artif. Intel.*, 8(3):643–660, 1994.
10. L. Lee, T. Berger, and E. Aviczer. Reliable on-line human signature verification systems. *IEEE Trans. Pattern Anal. and Mach. Intel.*, 18(6):643–647, June 1996.
11. B. Ly-Van, R. Blouet, S. Renouard, S. Garcia-Salicetti, B. Dorizzi, and G. Chollet. Signature with text-dependent and text-independent speech for robust identity verification. In *Proc. MMUA*, pages 13–18, December 2003.
12. B. Ly Van, S. Garcia-Salicetti, and B. Dorizzi. Fusion of hmm's likelihood and viterbi path for on-line signature verification. In *Biometrics Authentication Workshop*, May 15th 2004. Prague.
13. A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. The det curve in assesment of detection task performance. In *Eurospeech 1997*, pages 1895–1898, Rhodes, Greece, 1997.
14. L. Rabiner and B.-H. Juang. *Fundamentals Of Speech Recognition.* Prentice Hall, 1993.
15. D. Reynolds. Automatic speaker recognition using gaussian mixture speaker models. *The Lincoln Laboratory Journal*, 8(2):173–191, 1995.
16. D. Reynolds. An overview of automatic speaker recognition technology. In *Proc. IEEE ICASSP*, volume 4, pages 4072–4075, 2002.
17. J. Richiardi and A. Drygajlo. Gaussian mixture models for on-line signature verification. In *Proc. 2003 ACM SIGMM workshop on Biometrics methods and applications*, pages 115–122, 2003.
18. C. Vielhauer. *Biometric User Authentication for IT Security.* Springer, 2006.