UTILIZING BAUM-WELCH FOR ON-LINE SIGNATURE VERIFICATION

R. MARTENS, L. CLAESEN
IMEC V.Z.W.,
Kapeldreef 75, B-3001 Heverlee, Belgium
E-mail: {martens, claesen}@imec.be

This paper describes a fully operational on-line signature verification system. From a hardware point of view, its heart is the SMARTpen^{TM}, a special input-device allowing the recording of force and angle signals. The most important software aspect that we focus on here is the exploitation of the Baum-Welch procedure in the feature extraction process. This algorithm provides a mathematical basis for classifying a signature taking into account the relative importance of both the different signals under observation, and the distinct phenomena that are present in these. The usefulness of the approach is illustrated by presenting the results of a full-scale field test.

1 Introduction

The most typical characteristic of our current society is undoubtedly its high degree of automation. At the moment a very large amount of jobs, previously done exclusively by humans, can be executed without quality loss by machines. The weakness of this automation process is however in many cases situated in the problem of autonomously identifying the user.

Until very recently most of the identity verification systems have been using non-biometric identity verification techniques. These techniques do not base the actual classification on a unique characteristic of a certain person, but on some kind of secret - a key, a pass-word - shared between the subject being verified and the object that is performing the verification task. In return for the simplicity of the concept, a rather high price has to be paid. First of all, it is relatively easy to gain insight in the secrets that are used for verification, or to get possession of the objects employed for this purpose. This means that the degree of security that is achieved by a non-biometric identity verification system is in practice rather low. Secondly, the person who is to be verified has to do a non-trivial effort, like remembering a PIN-code or carrying a key.

As the number of occasions where a certain person has to prove his identity is rapidly growing, these disadvantages are getting more and more coercive. For this reason, and thanks to the technological evolution that makes more complex verification techniques feasible, there is in the last few years a steadily growing interest for biometric identity verification. These alternative identity verification systems use some kind of physiological (fingerprint, iris-pattern) or behavioral characteristic (handwriting, typing rhythm) to identify a certain person.

An excellent overview of the different biometric identity verification techniques is given in [1]. In this paper we concentrate on on-line signature verification. This means that we classify a certain person as a genuine or forger by observing the characteristics of his/her way of signing. An overview of the work that has recently been done in this area can be found in [2,3].

A description of the different steps that have to be taken when building a complete on-line signature verification system is given in for instance [4]. In general, we distinguish between three major stages: data-acquisition and pre-processing, feature-extraction, and classification. This separation is the backbone of our paper. Sec. 2 focuses on data-acquisition. Sec. 3 and 4 deal with parameter-extraction respectively classification. We evaluate the performance of the global system in Sec. 5. Conclusions are drawn in Sec. 6.

2 Data-acquisition and pre-processing

Two different trends can be observed in data-acquisition for on-line signature verification. A majority of the researchers [5,6,7] opt for a classic tablet for the registration of the positions of the pen-tip over time. A second group [8,9] hopes to end up in a better cost-benefit situation by building a special input-device that allows the observation of signals like forces, accelerations etc. In this work the second approach is chosen for. Our input-device, the SMARTpen™[a] (Fig. 1 and [10]), captures 5 different signals:

- Forces on the pen-tip in three directions.
- Angles of the pen-shaft relative to the writing surface in two dimensions.

We do not have any information about the rotation of the pen around its own axis. In practice, this is not a problem since this rotation is not characteristic for the signing process. The actual pen-rotation does however influence the other signals that are registered, as these are measured relative to the pen. Using a new reference co-ordinate system eliminates this effect. The main axes of this system are defined as the ones in which the force signals have extreme energy contents. The resulting signals are low-pass filtered with cut-off frequency about 30 Hz. It is generally accepted [3] that higher frequency regions contain no signer-specific information.

[a] SMARTpen™is a registered trademark is of LCI-SMARTpen N.V.

Figure 1: The SMARTpen^{TM}.

3 Feature-extraction

The ability to deal with non-linearities in the time-domain is an important property of a good on-line signature verification system. This is illustrated by the fact that many recent publications describe verification techniques that allow handling them [6,11,12].

The traditional approach for removing timing-differences in the signature verification area is to use Dynamic Time Warping (DTW) [13]. This technique originates from the speech recognition field [14]. In the last decade, Hidden Markov Models (HMMs) have been applied very successfully in the speech area for the same purpose. Recently, the HMM-approach migrated to the signature verification field, with good results as well [15,12].

However, while the concept of a state is very clear in the speech recognition area [16], it is not in the signature verification field. A first step towards solving this problem is the use of the Bakis type model[b], where the concept of a state is clearly weakened. We make one more step in this softening process, as will become clear in the next sections. We start by discussing the general characteristics of our approach in Sec. 3.1, and focus on its peculiarities in Sec. 3.2.

[b]In a Bakis type model the number of states is proportional to the average duration of the observation sequences.

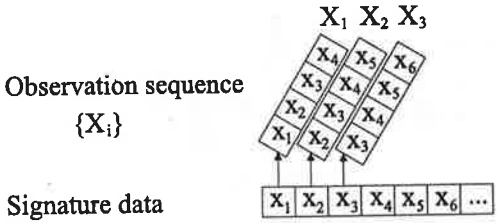

Figure 2: Construction of observation sequence out of signature data.

3.1 General characteristics

In previous work[17] we have illustrated the importance of taking into account the relative weight of phenomena situated in different frequency regions during the classification process. As a consequence, an observation has to contain information about several succeeding time-samples. In this context, a useful observation sequence for a certain signal might look like in Fig. 2. In Eq. (1), the i-th observation out of the complete sequence is denoted by X_i . The length of this vector is denoted by $n(X_i)$.

$$X_i = \left(x_{\,i} \; x_{i+1} \cdots x_{n(X_i)} \right) \tag{1}$$

The use of a classical HMM for signal classification requires the ability to compute the probability of occurrence for every state at each time instant. As we have only a very limited number of genuine signatures at our disposal, we have to make good assumptions about the characteristics of the observed features. Otherwise, it will become impossible to compute the probability density function (PDF) for an observation X in state i ($p_i(X)$) with a reasonable accuracy. If we succeed in decorrelating the individual features of the observations, we can approach the multivariate PDF $p_i(X)$ by a product of much easier to approximate univariate ones (Eq. (2)).

$$p_i(X) = \prod_{j=1}^{n(X)} p_{i,j}(DWT(X)_j) \tag{2}$$

As is clear from Eq. (2), we use the discrete wavelet transform (DWT)[11] for this purpose. The individual $p_{i,j}(x)$s from Eq. (2) can be approximated accurately by simple univariate Gaussian distributions.

Next to the construction of the PDFs, we have to compute transition probabilities from one state to another. Our approach is based on the concept

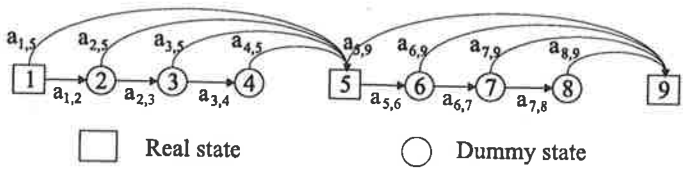

Figure 3: Part of the signature model, containing real and dummy states.

of a left-to-right HMM. This means that when a certain state is left, it can never be reached again. Furthermore, we will incorporate explicit time duration modeling into our model. We assume the duration to stay in a certain state is a Gaussian distributed variable. The fact that all our parameters have a Gaussian distribution will have an important effect on the actual classification process, as discussed in Sec. 4.

3.2 Special features

When building a HMM with the characteristics mentioned higher, a state is associated with every observation. This is problematic, as we do not have the same strong state-concept as in for instance speech. For the situation of Fig. 2, for instance, only one observation out of four should be matched to a state, the other ones representing transitions from one state to the next.

We solve this problem by introducing the concept of a dummy state. Fig. 3 shows a part of our model, consisting of a sequence of real and dummy states. For a dummy state i, we have $p_i(X) = 1$. As a result, observations matched to dummy states contribute only for timing purposes. The probability density function $p_i(X)$ for the observations in real states is defined by Eq. (2). The $a_{i,j}$-values of the model are chosen in order to make sure that exactly one observation is matched to a real state, while the duration between succeeding real states has a Gaussian distribution. By increasing the number of real states, relative to the number of dummy states, it becomes clear that this approach is just an evolution of the Bakis type model.

At this moment we can define the features that are used as an input to the actual classification process that is described in Sec. 4. After computing the most probable sequence of state transitions for a certain signature, we have two types of parameters at our disposal:

- Motion information parameters, describing the time-interval between two succeeding real states. The vector that contains all motion information parameters is denoted as ξ_{Motion}.

- Form information parameters, used to construct a vector by concatenating all the observations that are matched to real states. This vector is indicated by ξ_{Form}. As should be obvious from the previous discussion, we apply the DWT on these observations, before using them to construct ξ_{Form}.

We define the resulting feature vector ξ in Eq. (3).

$$\xi = (\,\xi_{Form}\;\xi_{Motion}) \tag{3}$$

The number of components in ξ is denoted by $n(\xi)$.

4 Classification

As has already been mentioned in Sec. 3, the distribution of the components of both ξ_{Motion} and ξ_{Form} can be reasonably well approximated in a Gaussian way. It is also intuitively clear that the degree of correlation between ξ_{Motion} and ξ_{Form} is small. This is also true for the correlation coefficients between different components of ξ_{Motion}. Thanks to the fact that the components of ξ_{Form} have been computed using the DWT, these are also decorrelated [17]. Both conditions (Gaussian distribution, absence of relevant correlations) are enough to conclude that χ^2 (Eq. (4)) has a chi-square distribution [18].

$$\chi^2 = \sum_{i=1}^{n(\xi)} \frac{(\xi_i - \mu_i(\xi))^2}{\sigma_i(\xi)} \tag{4}$$

In Eq. (4) $\mu(\xi)$ is the expected value for the feature vector ξ corresponding to our model. $\sigma(\xi)$ is a vector containing the variances on the components of ξ. To make a final decision, we simply compare the value of χ^2 to a threshold, that is defined by $n(\xi)$ and the user requirements.

Normally, a biometric identity verification system is characterized by two numbers. The False Acceptance Rate (FAR) is the percentage of forgeries that is considered as genuines. The False Rejection Rate (FRR) is the percentage of genuine signatures that is misclassified as originals. We can not predict the relation between FAR and classification threshold, because this relation depends on the quality of the forgeries in our database. However, we can estimate the value of the threshold for a required FRR[c].

---

[c] In order to describe the performance of the system independently of the classification threshold, we can use a third system parameter: the Equal Error Rate (EER). This is the percentage of misclassifications that occurs when the decision threshold is chosen in order to make FAR and FRR equal.

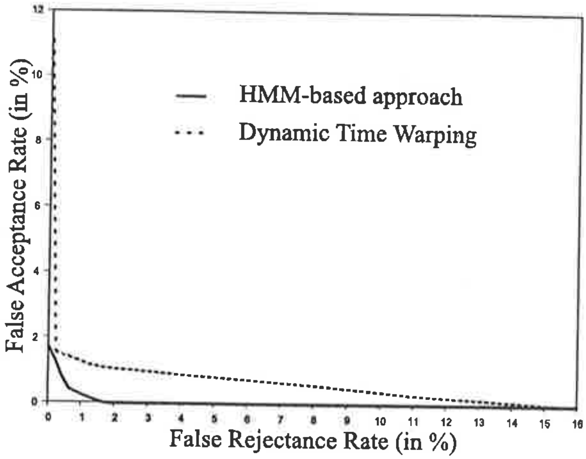

Figure 4: The benefits of the HMM-based verification approach.

5 Performance evaluation

The use of the Baum-Welch algorithm to model a class of genuine signatures, as has been described higher, is closely related to the use of the DTW-algorithm for feature extraction[17]. This last approach also allows dealing with the relative importance of phenomena situated in different frequency or time-intervals for a certain class of signatures. The only additional feature of the approach that is described here is that we can take into account the relative importance of the different signals that are observed. This importance is even allowed to change dynamically in the time and frequency domain. Fig. 4 compares both systems. It is clear that our HMM-based system performs relevantly better than the one using DTW-based feature extraction. The EER[d] is reduced from over 1% to about 0.5%.

6 Conclusion

We have built a fully operational biometric identity verification system based on on-line signature analysis. The hardware-heart of the system is the SMART-

[d] The EER is very dependent on the database that has been used. The database used here has been constructed over a period of three months. 57 people have been involved. 18 of them provided us with 20 genuine signatures each. 10 of these have been used as references. The 10 remaining ones are part of the set of test originals. The set of test forgeries for person i, consists of one random signature for each of the 56 persons j ($j \neq i$).

pen^{TM}. This instrument allows observing forces on the pen-tip and angles of the pen-shaft while a person is writing. The main topic of this paper is to answer the question on how to extract features that satisfy three conditions:

- They should be local in both time and frequency-domain.
- They reflect the relative importance of the different types of signals in a mathematically correct manner.
- They allow dealing with non-linear phenomena.

A first step in solving this problem is the use of the discrete wavelet transform to construct observation sequences where the individual features are decorrelated. Afterwards, we can build a model for the class of genuine signatures using the Baum-Welch algorithm. The approach that is presented is in fact an evolution of the Bakis type HMM.

Using this model, we extract form and motion parameters for a certain signature that should be verified by simply computing the optimal sequence of state transitions.

The last step in the verification process is the classification. The construction of a good classifier is in general a difficult problem, but in this case it is simplified a lot because the form and motion parameters we use have a Gaussian distribution and are decorrelated. As a result, we can easily combine them into a single parameter that has a chi-square distribution for the class of genuine signatures under study. If the class of reference signals we have at our disposal is really representative for the class of genuine signatures this even allows to predict the FRR, by examining the classification offset.

We have illustrated that the technique offers better classification results than a technique based on the combined use of the discrete wavelet transform and dynamic time warping. The price that has to be paid for this is a drastic increase of the complexity of the training process. The actual verification procedure does not suffer very much from this drawback, as the major complexity of the process really has to be situated in this training stage.

Acknowledgments

This research was supported by a scholarship from the Flemish Institute for the promotion of the scientific-technological research in industry (IWT).

References

1. B. Miller, IEEE Spectrum **31**, 22 (1994)

2. R. Plamondon, G. Lorette, Pattern Recognition **22**, 107 (1989)
3. F. Leclerc, R. Plamondon, Int. Journal of Pattern Rec. and Art. Intelligence **8**, 643 (1994)
4. R. Plamondon, Int. Journal of Pattern Rec. and Art. Intelligence **8**, 795 (1994)
5. J. Bromley, J.W. Bentz, L. Bottou, I. Guyon, Y. Lecun, C. Moore, E. Säckinger, R. Shah, Int. Journal of Pattern Rec. and Art. Intelligence **7**, 669 (1993)
6. B. Wirtz, in *Proc. 3rd Int. Conf. on Doc. Analysis and Recognition* 179 (1995)
7. K. Huang, H. Yan, Optical Engineering **34**, 3480 (1995)
8. R. Baron and R. Plamondon, IEEE Trans. on Instrumentation and Measurement **38**, 1132 (1989)
9. N. M. Herbst, C. N. Liu, IBM Journal of Research and Development **21**, 245 (1977)
10. L. Claesen, D. Beullens, R. Martens, R. Mertens, S. De Schrijver, W. De Jong in *ED&TC'96 User Forum* 201 (1996)
11. R. Martens, L. Claesen, in *Advances in Document Image Analysis*, eds Nabeel A. Murshed, Flávio Bortolozzi (Springer-Verlag, Berlin Heidelberg New York, 1997).
12. R. S. Kashi, J. Hu, W. L. Nelson, in *Proc. 4th Int. Conf. on Doc. Analysis and Recognition* 253 (1997)
13. Y. Sato, K. Kogure, in *Proc. of the 6th Int. Conf. on Pattern Recognition* 823 (1983)
14. H. Sakoe, S. Chiba, IEEE Transactions on Acoustics, Speech and Signal Processing **26**, 43 (1978)
15. L. Yang, B. K. Widjaja, R. Prasad, Pattern Recognition **28**, 161 (1995)
16. J. D. Markel, A. H. Gray in *Linear Prediction of Speech* (Springer Verlag, New York, 1976).
17. R. Martens, L. Claesen, in *Proc. 13th Int. Conf. on Pattern Recognition* **3** 38 (1996)
18. A. Papoulis in *Probability and Statistics* (Prentence-Hall International Editions, 1990).