On-Line Signature Verification by Dynamic Time-Warping.

Ronny Martens, Luc Claesen IMEC vzw Kapeldreef 75, 3001 Heverlee, Belgium. E-mail: martens@imec.be

Abstract.

In this paper, we discuss an on-line signature verification system based on dynamic time-warping (DTW). The DTW-algorithm originates from the field of speech recognition, and has been applied successfully in the signature verification area more than once. However, until now, few adaptations have been made in order to take the specific characteristics of signature verification into account. According to us, one of the most important differences is the availability of a rather large number of reference patterns, making it possible to determine which parts of a reference signature are important and which are not. By disconnecting the DTW-stage and the feature extraction process we are able to deal efficiently with this extra amount of We demonstrate the benefits of our information. approach by building and evaluating a complete system.

1. Introduction.

As a result of the growing automation, there is an increasing need for efficient and reliable identity verification by machines in our community. For decades, automatic identity verification has been covering almost exclusively the non-biometric part of the field. Non-biometric verification means that a person identifies himself by a key, a pass-word, a PIN-code,... These techniques are, however, far from perfect. The "secrets" that are shared with the machine can be lost or stolen, and the person who is using them to prove his identity must do a considerable effort. To overcome these disadvantages is the objective of biometric identity verification. An overview of the different methods is given in [1].

Here we focus on identity verification by on-line signature verification. This means that we use a special input instrument to collect data such as pen-tip positions, speeds, accelerations or forces, while a person is signing. Several different methods have already been used to classify a certain signature as "original" or "forgery". An overview can be found in [2] and [3]. The complete verification process is performed in different steps. The most important ones are data-acquisition and preprocessing, feature-extraction and classification. This subdivision in three stages forms the backbone of this paper. Chapter 2 deals with data-acquisition and preprocessing. In chapter 3 we reveal our feature-extraction approach. The actual classification is done by the methods described in chapter 4. Chapter 5 evaluates the performance of the complete system. Conclusions are drawn in chapter 6.

2. Data-acquisition and pre-processing.

On-line signature verification requires the use of a special input instrument to collect the signals to be classified. The instrument that is used in this research (figure 1 and [4]) looks like a conventional pen, but it contains several sensors. These sensors serve two purposes:

- Measuring the forces on the pen-tip in 3 directions.
- Measuring the pen inclination angles.



Figure 1: The SmartPen and some of its components (force sensor and controller ASIC).

As we are not interested in rotations of the pen around its own axis, 2 angle sensors instead of 3 are sufficient. All 5 signals are low-pass filtered with cut-off-frequency 40 Hz. The resulting signals are sampled at 100 Hz. As already mentioned, the orientation of the pen in the signer's hand is not characteristic, but it does influence the other signals, as those are measured relative to the pen. We eliminate the effect of these pen-rotations by redefining the co-ordinate system that is used. The new reference axis's are chosen as the ones with extreme energy contents.

3. Feature-extraction.

When comparing a test and a reference signature in order to extract the features used for classification, we are faced to the presence of non-linear timing differences between the 2 patterns involved. A common way to deal with this problem is to use the DTW-algorithm. First we summarise the basics of the DTW-approach, afterwards we discuss how DTW is integrated in the featureextraction process.

3.1. The DTW-approach.

The goal of the DTW-algorithm - as presented in [5] is to find an optimal time-alignment between two patterns R (Reference) and T (Test), both sampled with the same, constant sampling-rate. We define a timealignment (warping-path) p as:

(1)
$$p = c(0), c(1), c(2), ..., c(K)$$

(2)
$$p_k = c(k) = (i(k), j(k))$$

I and j refer to the i-th/j-th sample of R respectively T. As the conditions for p to be a valid warping-path are not crucial in this discussion, we do not focus on them. The interested reader is referred to [5]. Assume:

(3)
$$d(p_k) = d((i(k), j(k))) = ||R_i - T_j||$$

(4)
$$D(p) = \frac{\sum_{k=1}^{K} w(k).d(p_k)}{\sum_{k=1}^{K} w(k)}$$

The goal of the DTW-algorithm is to find the path P that minimises D(p). This path is, from the DTW point of view, the optimal time-alignment between T and R.

3.2. DTW and feature extraction.

One of the first successful attempts to use the DTWalgorithm in signature verification has been reported by Sato and Kogure [6]. Their approach is typical for how DTW has been used in signature verification until now [7, 8]. Sato uses the DTW-algorithm to extract two features x_{Form} (5) and x_{Motion}^{1} (6) out of a reference/test couple of signatures.

(5)
$$x_{Form} = D(P) = \min_{p} D(p)$$

(6)
$$x_{\text{Motion}} = \sum_{(i, j) \in P} \left\| i - j \right\|$$

So far [6, 7, 8], x_{Form} and x_{Motion} have been constructed using w(k) (see (4)) independent of i(k). As a result, neither x_{Form} nor x_{Motion} does reflect any information about the relative importance of the different parts of the reference signature in time or frequency domain. Because both parameters are global, it's not possible to construct a classifier out of them that does so.

The improved results presented in [9] illustrate that it can be advantageous to use this type of stability-information. The author uses x_{Form} and x_{Motion} as computed in [6]. Both features are however extracted for couples of corresponding segments in R and T, and not for R and T as a whole. The resulting local parameters are then combined to a global one. Each local parameter contributes to the global one in correspondence to its stability over the complete set of reference signatures created for one person.

In the next sections we describe an alternative method to determine form and motion features allowing the construction of a classifier that makes it possible to use the available stability information more explicitly.

3.3. Form information-features.

"Form information" (for example x_{Form} defined in (5)) is extracted from the signals R and T without taking any absolute timing aspects into account. As a result, in order to derive form information, one should not look directly at the optimal warping-path P, but at the aligned signals that are constructed using it. A problem with these signals is that the number of samples they contain is not necessarily the same for different T's. We solve this difficulty by constructing a signal F out of P, R and T.

(7) $F_i = \underset{(i, j) \in P}{\text{mean}}(T_j)$

The length of F is equal to the length N of R. We will use these N scalars F_i as local form classification parameters.

¹ Our definitions of x_{Form} and x_{Motion} differ from those in [6] as we do not apply a linear rescaling from the signals' time-axis's prior to DTW. This rescaling had almost no effect on our verification results.

3.4. Motion information-features.

"Motion information" (for example x_{Motion} defined in (6)) reveals the magnitude of the actual timingdifferences between the signals T and R. It is extracted using only the optimal warping-path P. A very annoying property of this warping-path is that a time-correction at sample k is reflected in the time-correction for all samples l ($l \ge k$). This is why we do not use the original path itself, but we extract from it a signal that's strongly connected to its derivative. We define our local motion parameter $M_i(n)$ as:

(8)
$$M_i(n) = P_{i+n} - P_i$$

(9)
$$P_i = j \Leftrightarrow (i(k), j(k)) \in P$$

Both the form parameters F_i and the motion parameters $M_i(n)$ contain local information about the signing process. Due to this property we are able to incorporate stability information in the classification procedure. In the next chapter we will reveal how this is realised in practice.

4. Classification.

The final stage in the verification procedure is the actual classification. The choice of the discriminant method that's best suited for a particular application is extremely important as it should allow to include all available statistical knowledge about the classes involved in the problem. In this paper, we use Mahalanobis decision making. The Mahalanobis distance metric (D_M) is optimal on 2 conditions:

- The feature vector X has a multivariate normal distribution over the class of original signatures for a certain person.
- The probability density of X for the class of forgeries can be considered constant in the area where the probability density of X for the class of originals is non-zero.

Both conditions are reasonably well satisfied. The Mahalanobis distance from a pattern X to a class of patterns T, characterised by its mean feature vector \overline{T} and its covariance matrix Σ is defined by:

(10) $D_M(X) = (X - \overline{T})^T \cdot \Sigma^{-1} \cdot (X - \overline{T})$

We can construct a good estimate \overline{T}_E for \overline{T} (11), but as the number of references we have is much smaller than the dimension of our feature vector, we must use a very rudimentary approach Δ for Σ (12). (11) $\overline{T}_{E} = \underset{\text{originals}}{\text{mean}}(T)$

(12)
$$\Delta = \operatorname{diag}(\operatorname{mean}_{\text{originals}}((T - \overline{T}_{E}).(T - \overline{T}_{E})^{T}))$$

The resulting classifier D_S is constructed as:

(13)
$$D_{s}(X) = (X - \overline{T}_{E})^{T} \cdot \Delta^{-1} \cdot (X - \overline{T}_{E})$$

We call D_s the simplified Mahalanobis discriminant function. In the next sections, we will focus on the use of the form and motion parameters as discussed in the preceding chapters to do the actual classification using D_s .

4.1. Classification using form information.

A generally accepted performance index for identity verification systems is the Equal Error Rate (EER). This is the error-percentage that occurs when False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal for a certain database of signatures. The database used here contains 360 signatures. from 18 different persons, collected over a period of 3 months. For each participant 15 signatures are used as originals for the construction of the classifier. The remaining 5 signatures are used as test originals. As forgeries for a certain person, we use the original signatures produced by the other signers.

Using $D_S(X)$ with X=F as a classifier we achieve an EER of 6.7%. Using X=x_{Form} we have an EER of 8.9%. The main reason for this rather small improvement is that Δ_{Form} is a very poor approximation of Σ_{Form} . This is illustrated in figure 2.



Figure 2: Correlations between classification features.

Figure 2 shows the average value of a correlation coefficient that is "Distance from diagonal" away from the main diagonal of its correlation matrix. The correlations computed using X=F are shown by the dashed line. These results are computed using the 18 sets of 20 original signatures in our database. As expected, the correlation coefficients are becoming smaller when farther removed from the diagonal. We are however still far from the ideal case, where the correlation matrix is a unity-matrix, making $D_{S}(X)$ and $D_M(X)$ equal. There exists however a linear transformation of the parameters F_i that produces a feature vector with corresponding Σ diagonal. This transformation is called the Karhunen-Loève Transform (KLT). Unfortunately we can not compute the exact KLT since this would require knowledge of Σ_{Form} . This is why KLT-approximations like the Discrete Cosine Transform (DCT) have been developed. Here we will demonstrate how another transformation, the Gabortransform [10], can be used successfully for the same purpose. We construct a new feature vector G out of F, by shifting a Gaussian window (±300 msec.) over F and by computing the FFT of the different signals generated. Figure 2 displays in full line the average correlation between the different components of G. It can be seen that the correlations are much smaller than the ones computed using X=F. This fact, combined with the good locality of our parameters in both time and frequency domain that is necessary to take into account the stability information for a certain signer, allows us to reduce the EER from 6.7% to 3.3%.

So far, we have been using the complete information contents of the 5 signals used for verification. It is however generally accepted [2] that the relevant information in human handwriting is concentrated in a much smaller frequency range. This is illustrated in figure 3.



Figure 3: Relation between frequency range in use (0..x Hz) and EER.

Figure 3 reveals that optimal classification is achieved when we limit ourselves to using only those Gabor transform-coefficients describing the signal contents from the 0 Hz. to ± 30 Hz. Acting like this, the EER is further reduced to 1.4%.

4.2. Classification using motion information.

Here we use $D_S(X)$ with X=M(n). Results are optimal if M(n) is constructed by using a time-shift of 100 msec. (n=10). The major advantage of the use of local stability information for classification using motion information is clearly illustrated by the reduction of the EER from 13.3% using X=x_{Motion} to 2.0% using X=M(10).

5. Global system performance.

Our complete verification system integrates the use of form and motion information by using the feature vector $X_1 = \begin{bmatrix} G \\ M(10) \end{bmatrix}$ as an input for its (simplified Mahalanobis) classifier. Since the dimension of X_1 is dependent on the number of samples N in the reference pattern R, $D_s(X_1)$ is divided by this dimension before being compared to the classification offset. Acting like this - and keeping in mind that only random forgeries are used to test the system - we achieve perfect classification (EER = 0%).

Figure 4 compares classification using X_1 to classification using $X_2 = \begin{bmatrix} X_{Form} \\ X_{Motion} \end{bmatrix}$. The classification starting from X₂ is performed by using the kernel approach [11] instead of by Mahalanobis decision making, because for X₂ the kernel approach results in a much better classification performance. Using the kernel approach for classification by X₂ we achieve an EER of 0.3%. The use of the Mahalanobis discriminant function on the contrary results in an EER of 1.6%. To make a comparison between classification using X1 and X2 possible, the decision threshold has been defined as the logarithm of the average probability density per variable for both types of classification. Clearly, the separation between original signatures and forgeries is more evident using stability information as described in this paper than without using this type of information.



Figure 4: Comparison of verification results with and without using stability information.

6. Conclusion.

We have built a signature verification system that combines the benefits of DTW and the use of local stability information for a certain reference signature. The crucial point for its success has to be found in the decoupling of the DTW-algorithm and the actual feature extraction step. By using the feature extraction procedures explained, we are not confronted with the problem of finding the real segments in a signature. Consequently, we do not need special heuristics to deal with an unequal number of segments in both signatures that have to be compared.

The major drawback of our approach is that we use each of the 5 signals available individually, instead of using a higher dimensional approach. Acting like this we lose the link between the signals. Furthermore, better transformations than the Gabor-transform to decorrelate the form classification features are possible. A last weakness is that the Mahalanobis-classification we use is definitely not optimal. Optimal classification requires the use of probability density information of all classes (originals as well as forgeries) involved.

7. Acknowledgement.

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

8. Bibliography.

- [1] B. Miller, "Vital signs of identity", IEEE spectrum, pp. 22-30, feb. 1994.
- [2] R. Plamondon, G. Lorette, "Automatic Signature Verification and Writer Identification - State Of The Art", Pattern Recogn., vol. 22, no. 2, pp. 107-131, 1989.

- [3] F. Leclerc, R. Plamondon, "Automatic Signature Verification: The State Of The Art -- 1989-1993"; Int. J. Pattern Recogn. Artif. Intell., vol. 8, pp. 643-660, 1994.
- [4] L. Claesen, D. Beullens, R. Martens, R. Mertens, S. De Schrijver, W. de Jong, "SmartPen: An Application of Integrated Microsystem and Embedded Hardware/Software CoDesign", ED&TC'96 User Forum, pp. 201-205, 1996.
- [5] H. Sakoe, S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition", IEEE Trans. on acoustics, speech and signal processing, vol. 26, no. 1, pp. 43-49, 1978.
- [6] Y. Sato, K. Kogure, "On-line Signature Verification Based on Shape, Motion, and Writing Pressure", Proc. 6th Int. Conf. on Pattern Recognition, pp. 823-826, 1982.
- [7] M. Yasuhara, M. Oka, "Signature verification experiment based on nonlinear time alignment, A feasibility study", IEEE Trans. Syst. Man Cybern. SMC-7, pp. 212-216, 1977.
- [8] M. Parizeau, R. Plamondon, "A comparative analysis of regional correlation, dynamic time warping and skeletal tree matching for signatures", IEEE Trans. Pattern Anal. Mach. Intell.,vol. 12, no. 7, pp. 710-717, 1990.
- [9] B. Wirtz, "Stroke-Based Time-Warping for Signature Verification", Proc. of the 3rd Int. Conf. on Document Analysis and Recognition, pp. 179-182, 1995.
- [10] S. H. Nawab, T. Quartieri, Advanced Topics in Signal Processing", Prentice-Hall, Englewood Cliffs, N. J., 1988.
- [11] M. Nadler, E. P. Smith, "Pattern Recognition Engineering", John Wiley and Sons, 1993.