Recognition of Human Signatures

Andrzej Pacut†,‡ and Adam Czajka†
† Warsaw University of Technology
Nowowiejska 15/19, 00–665 Warsaw, Poland
‡ Scientific and Academic Computer Network NASK
Bartycka 18, Warsaw, Poland
pacut@ia.pw.edu.pl, http://www.ia.pw.edu.pl/~pacut

Abstract

We used a digitizing tablet to collect handwritten signatures, with five quantities recorded, namely horizontal and vertical pen tip position, pen tip pressure, and pen azimuth and altitude angles. We divided the signature features into visible ones, namely those related to an "image on the paper" and hidden ones, i.e. those using time-related observations. Cluster analysis was applied to segment the feature space into sub-regions of "similar" signatures. The classification function was approximated with the use of neural networks, namely a two-layer sigmoidal perceptron and the RCE network which is a variety of radial-basis network. Both signature classification and signature verification problems are considered.

1 Introduction

The word *signature* has many meanings. According to American Heritage Dictionary [1], it is among many other, *one's name as written by oneself*, or *the act of signing one's name*, or also *a distinctive mark, characteristic, or sound indicating identity*.

The second definition refers to the entire act of signing thus stressing important fact that also *way* the signature is made is a part of this signature. This may include not only the horizontal (x) and vertical (y) coordinates of the pen, but also the speed of the pen tip, the pressure applied when signing, the way the pen is held, etc. The signature in this meaning can be treated as a multidimensional output s = $s(t),\ t \in T$, of a dynamic system observed during time T of making the signature. This system characterizes a person making the signature. The first definition is rather limited, when compared to the second one, since it refers to a *written* name, i.e. a two-dimensional image which includes no time-related information. It is equivalent to the set $S = \{(x, y):\ x = x(t),\ y = y(t),\ t \in T\}$ on a plane. Returning to the dictionary definitions of *signature*, note that while the third definition extends the meaning of "signature" beyond the handwritten signatures, it may also be used for handwritten signatures to abstract their *identifying property* from particular realizations, as understood in the first and the second meanings. Note that this very meaning of signature is employed when saying "this is my signature". Note also that the same person may have several different signatures (full name, initials, surname, etc.).

The first author acknowledges the support of Scientific and Academic Computer Network NASK

It is important to differentiate here between the three definitions of *signature* presented above. To stress this important distinction, we will say *token* for each particular instance of signatures (the second meaning), while by somebody's *signature* we mean an abstract entity of all possible tokens that identifies a person (the third meaning). Finally, each signature image (on a paper or another carrier, in particular obtained by scanning) will be called the x-y token or the scanned token, to stress the way it can be reproduced.

We approach here the problems of identity classification and identity verification. In the *classification task* we assume that the classes equivalent to signatures are known in advance. Each class is represented by one or several tokens. In other words, for a given new token one has to *estimate* the class it belongs to. In the *verification problem* each new token comes with a hypothetical class and a *hypothesis is tested* whether the token belongs to this class. While traditionally x-y tokens are used in both tasks, more dependable systems require information about the very signing process, like the velocities of pen movement, pen tip pressure during the signing, etc. In this paper we will use a digitizing tablet to record the time evolution of both the horizontal and vertical pen coordinates, the pen tip pressure, and two angles related to pen's position with respect to the paper plane, namely its azimuth and altitude. These data make it possible to select features necessary for signature identification and verification. The feature set is reduced to remove dependent features. The resulting feature space is partitioned into sub-regions, and the decision function is

Andrzej Pacut, Adam Czajka, "Recognition of Human Signatures", *IEEE-INNS-ENNS International Joint Conference on Neural Networks IJCNN'2001, Washington, D.C., USA*, Vol 2, pp. 1560-1564, IEEE, 2001

approximated separately in these subregions to increase approximation accuracy. The partitioning is made with the use of cluster analysis, different for identification and verification purposes.

2 Existing approaches

The existing approaches to signature verification can be classified on the basis of the assumed level of available information. Two basic classes of approaches can be distinguished here, namely *static problems* that assume no time-related information and the *dynamic problems* with time-related information available and the data in the form of p-dimensional functions of time, with p ranging from 2 to 5.

The *static approach* can be applied to a wide range of practical problems, and requires no special hardware. Dimauro et al. [4] present a multi-expert signature verification system for bank check processing. First, a pre-preprocessing is performed to localize a signature token, make it rotation invariant, and segment it into compact parts. Verification is made on the base of 6 features calculated for the entire token and 16 features related to the segments. Three algorithms are combined in a majority voting system, resulting in the false rejection rates (FRRs) about 22% and the false acceptance rates (FARs) about 3.9%. Lee and Pan [6] presented a methodology to trace the x-y tokens a human would do and to employ the resulting dynamic information. First, a 1 pixel-width signature skeleton is built and represented by a sequence of strokes, and critical points like the minima, maxima, and end points are found for each stroke. The critical points are normalized with the use of their covariance matrix to make them translation-, rotation-, and scaling-invariant.

The *dynamic approaches* can be classified by the dimension of the observation vectors, yet the methodology in principle would not depend on the dimension. A **2-D** observation space is obtained if x and y coordinates are observed in time. Under these assumptions, Brault and Plamondon [2] make a model of forgery dynamics that employs a nerve-muscular system. The task of imitating a signature is modeled here by a number of consecutive subtasks. In each subtask a signature element is imitated, defined as a triple: curvilinear stroke – angular stroke – curvilinear stroke. Each subtask consists of spatial target perception, preparation of strokes and stroke execution. Continuity of movement is ensured by overlap of subtasks. Lee et al. [5] presented an on-line signature verification system based on about 40 dynamic and static features out of which 10–15 features are chosen individually for reference tokens. Majority rules are applied to decisions based on a relative distance between the tested and the reference tokens in the feature space, and times of signing. A **3-D** observation space is obtained by additionally recording the pen tip pressure. For such data, Crane and Ostrem [3] based their procedure on 44 features that proved best in ERR (equal error rate) minimization. For verification purposes, a weighted Euclidean feature distance is applied, along with an arbitrary fixed threshold of the maximum distance between the reference and the presented tokens. Personalized feature sets resulted in better ERR rate (0.5%) than that obtained for a common feature set (1.75%) at the cost of a longer enrollment procedure due to the use of at least 100 tokes per signature. Two commercially available systems also seem to employ 3D observations, namely Signplus [8] and Cybersign [7]. The most extensive set of observations is **5-D**, where additionally two pen angles are recorded. Such data were employed by Wessels and Omlin [11] who proposed a hybrid verification system. Each token is first normalized to be scale- and rotation-invariant. Verification methodology based on Hidden Markov Models and Kohonen Self Organizing Map lead to the FAR of order of 13% for the FRR set to 0%.

3 Experimental design and data collection

Although the act of signing depends on the signer's personality, each signature instance (i.e. token) depends on particular circumstances. Consequently, the reference tokens should be taken over a possibly long period of time, and in different situations. It may even happen that certain circumstances make a token inconsistent with previous tokens. In such cases, it is convenient to introduce a second class (second signature) for the same person. In our experiments we took no more than two signature tokens per day from one person. The accompanying circumstances thus may include both tiredness at the end of the day and the morning energy, excitement before an exam and a the calm after it, happiness after a success and sadness of a defeat. We were prepared to obtain tokens that may have been classified to different signatures.

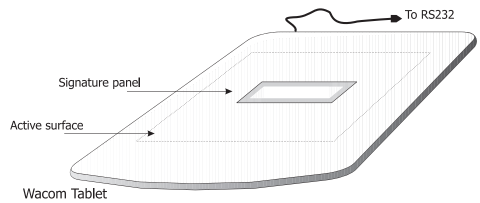

Figure 1: The signature panel

The signatures were recorded with the use of a Wacom digitizing tablet with a special paper template to make it a *signature panel*, Fig. 1. In the center of the template a frame was placed, whose dimensions were similar to a stan-

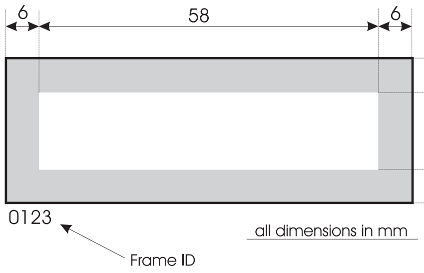

Figure 2: Signature frame

dard window used for confirming credit card transactions (Fig. 2). A white rectangle inside a bigger gray frame is to suggest the place for the signature. A black frame additionally encloses the gray one to localize the signatures. On the other hand, no *strict* rules were enforced, and some signatures extended beyond white or even gray spaces. The tablet enables the recording of pen tip coordinates, pen tip pressure, and pen orientation, namely its altitude and azimuth angles with respect to the tablet plane. The measurements were registered every 0.01 sec. To enter a signature token into our data base, the person first chose a signature version from among three possibilities, namely name + surname, only surname, or just the initials. Each version represented a different class, hence it was treated as a variant signature of the same person.

The data base we collected consists of 359 tokens made by 37 persons. These tokens were grouped later into 48 signatures (classes). The data were grouped into a *learning set* of 299 tokens (the reference tokens) and a test set of 48 tokens. The remaining 12 signature tokens were intensionally prepared to forge 9 signatures. To create the forgeries, several volunteers were asked to choose a signature he or she can forge, and to train their forgeries on the basis of x-y tokens for as long they wished before making the actual data base entry. Sometimes this training took as long as a few days.

4 From observations to features

4.1 Notation

Denote by $s(k)$ a single point in the observation space $\mathcal{S} \subset \mathbb{R}^5$ as registered by the tablet at the moment k, namely

$$s(k) = \begin{bmatrix} s_1(k) & \cdots & s_5(k) \end{bmatrix}^T$$

where s_1 denotes the pen tip x coordinate, s_2 the pen tip y coordinate, s_3 the pen tip pressure, s_4 and s_5 the pen orientation, namely the altitude and the azimuth angles, both related to the tablet plane. A single token can be then presented as a matrix

$$S = \begin{bmatrix} s^T(1) \\ \vdots \\ s^T(N) \end{bmatrix} = \begin{bmatrix} s_1(1) & \ldots & s_5(1) \\ \vdots & & \vdots \\ s_1(N) & \ldots & s_5(N) \end{bmatrix}$$

where N denotes the number of time steps for the token. Since the time can be labeled by the row numbers of S, the token can be visualized as a segment of a parametric curve in the observation space $\mathcal{S}$. Since a single token is represented by S, the signature is represented by a collection of matrices.

4.2 Feature selection

We started the design of the system from selecting a set of signature features. Since we tried to minimize the set of features, only the following features were considered

Token's length understood as the number of token's points N.

Average values of observation components $\overline{s_j} = \frac{1}{N}\sum_{k=1}^{N} s_j(k)$, $j = 1, \ldots, 5$, identical to the column averages of S.

Standard deviations of observation components $\sigma_j = \sqrt{\overline{s_j^2} - \overline{s_j}^2}$, $j = 1, \ldots, 5$

Trend coefficients, namely the slopes α_j of trend lines of each observation component. The trend lines are given by $s_j^0(k) = \overline{s_j} - \alpha_j k$ for $j = 1, \ldots, 5$, where

$$\alpha_j = \frac{\rho_j \sigma_j}{\sigma_k}$$

$$\rho_j = \frac{\overline{k\, s_j} - \overline{k}\,\overline{s_j}}{\sigma_j \sigma_k}$$

$$\overline{k\, s_j} = \frac{1}{N}\sum_{k=1}^{N} k\, s_j(k)$$

and $\sigma_k^2 = \overline{k^2} - \overline{k}^2$, $\overline{k} = \frac{1}{2}N(N+1)$, $\overline{k^2} = \frac{1}{6}N(N+1)(2N+1)$. We found that the trend coefficients were practically equal to zero for pen tip pressure and the angles, hence we always set $\alpha_3 = \alpha_4 = \alpha_5 = 0$ and only α_1 and α_2 were used as features. In the remaining feature calculations we always use "de-trended" coordinates $s_j^* = s_j - s_j^0$, Fig. 3.

Eigenvectors and eigenvalues of the inertia matrix. The inertia matrix $\mathbf{\Gamma}$ of two first coordinates s_1, s_2 is identical to the sample covariance matrix of these observation components, namely

$$\mathbf{\Gamma} = \begin{bmatrix} \overline{s_1^{*2}} & \overline{s_1^* s_2^*} \\ \overline{s_1^* s_2^*} & \overline{s_2^{*2}} \end{bmatrix} \tag{1}$$

The square roots of the eigenvalues $\lambda_1 \geq \lambda_2$ of Γ are used as features.

Third central moments We also use the third central moments of s_1^*, s_2^* and s_3^* observations, namely

$$\beta_{p,q,r} = \frac{1}{N} \sum_{k=1}^{N} (s_1^*(k))^p (s_2^*(k))^q (s_3^*(k))^r \qquad (2)$$

where $p + q + r = 3$, $p, q, r = 0, 1, 2, 3$, which are related to the symmetry of the curve in 3-dimensional observation space. To keep a similar scale ratio for all features, we will use the cube roots of the third central moments as the features.

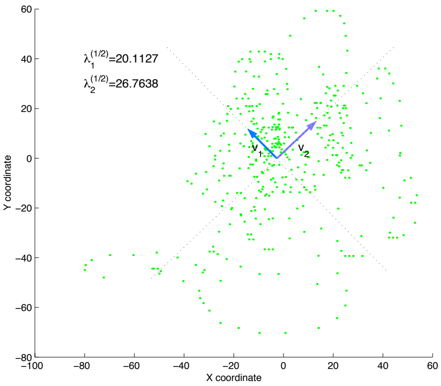

Figure 3: A single x-y token with the trends removed. Two orthogonal eigenvectors are also presented.

Note that of all the features introduced above, Tab. 1, only the mean values $\overline{s}_1$, $\overline{s}_2$ and the standard deviations s_1 and s_2 can be based on the observations of scanned tokens, i.e, when no time-related observations are given. The remaining features cannot be perceived from the x-y signature and will thus be called *hidden features*.

4.3 Minimal set of features

The initial family of features, Tab. 1, was found quite sufficient for the relatively small size of our signature database. This family can be enlarged by including more detailed properties of signatures, and in particular, taking into account the local properties of observation dynamics. On the other hand, growing dimensionality of the feature space makes it more difficult to train the classifying network. It is then important to construct a family of features that is not redundant, in such a way that the features are not dependent. The initial family of features *is* redundant, since some features are linearly dependent, as it may be seen from the sample correlation matrix. Linear regression analysis provides algorithms to remove highly correlated elements from a regression set and any of such methods can be adopted. The final set of features was assumed to have the absolute values of sample correlation coefficients not exceeding an arbitrary *correlation threshold* set at 0.4. Since the choice of "strongly correlated" features is not unique, we were inclined to remove those features which were more complex to calculate. For instance, the correlation coefficients between the eigenvalues of the inertia matrix and the standard deviation of the corresponding observation components are greater than 0.9, and we were apt to reject eigenvalues rather than standard deviations. This procedure lowered the size of the feature set to 12, Table 1. We did not use more sophisticated methods to detect nonlinear dependencies, and this set was used for both classification and verification purposes.

Table 1: Initial and final sets of signature features

Feature	Initial set	Final set
Token's length	N	N
Averages	$\overline{s_1}, \ldots, \overline{s_5}$	$\overline{s_1}, \ldots, \overline{s_5}$
Standard deviations	σ_1, σ_2	σ_1, σ_2
Trend coefficients	α_1, α_2	α_2
Cube roots of third central moments	$\sqrt[3]{\beta_{300}}$, $\sqrt[3]{\beta_{210}}$, $\ldots$, $\sqrt[3]{\beta_{003}}$	$\sqrt[3]{\beta_{300}}$, $\sqrt[3]{\beta_{120}}$, $\sqrt[3]{\beta_{012}}$

5 From features to classifying function

Mapped to the feature space, the distribution of tokens may be quite uneven. Some tokens may be close to each other even if they represent different signatures, and some other may be quite apart. Since the classification function must be approximated on the basis of feature vectors, this uneven distribution may lead to poor approximation quality. It is proposed here to approach this problem by performing a cluster analysis in the feature space. In this way the feature space can be divided into sub-regions of "similar" features, with the definition of similarity properly defined. The classification function can be approximated separately in each sub-region to assure smaller approximation error and thus better classification and verification. We tested several clustering methods and found that the best results were given by an algorithm based on a *hierarchical cluster tree*, with the distance (similarity) matrix based on the Mahalanobis distance $d(x, y) = (x - y)^T C^{-1}(x - y)$ where C denotes the sample covariance matrix of the features. The tokens are then organized into a binary hierarchical cluster tree which is consecutively divided intro clusters.

Both classification and verification were carried out with the use of neural networks. A two layer sigmoidal perceptron network was built with outputs corresponding to particular signatures. The network was trained to approximate the class membership function. After the training, the classification decisions were based on the winning output neuron. A special case of radial-basis function network, namely the RCE network [10] was also tested.

6 Results

6.1 Signature classification

For classification purposes, all signatures were first divided into six clusters. Neural classifiers were built separately for each cluster data. The results of classification strongly depend on the number of reference tokens representing each signature in the network training, Tab. 2. In the actual identification and verification experiments we used from 4 up to 8 reference tokens per signature. The correct classification ratio was equal to 95.8%.

Table 2: Classification quality vs. number of reference tokens. Nonlinear perceptron was trained for 1,2,3 and 4 reference tokens per one signature. Total number of $K = 48$ signatures, $2K + 1$ hidden neurons, K output neurons.

# of reference tokens per signature	classification error
1	36.4 %
2	25.0 %
3	16.7 %
4	12.5 %

6.2 Signature verification

Here all signatures were divided on the base of visible features into four groups. Verification was performed with the use of hidden features only. We employed here both the winning neuron output value $O^{(1)}$ and the next highest value $O^{(2)}$. For the hypothesis *the token presented to the network's input is genuine* not to be rejected, the network two highest outputs had to fulfill the relations

$$\begin{aligned} O^{(1)} &\geq m_1 - \frac{\varsigma_1\, t}{\sqrt{N-1}} \\ O^{(2)} &\leq m_2 + \frac{\varsigma_2\, t}{\sqrt{N-1}} \end{aligned} \tag{3}$$

where m_k denote the average values of the two outputs for the network stimulated by all reference tokens for this signature, ς_k are the corresponding sample standard deviations and N denotes the number of reference tokens for this signature. By t we denoted the upper α-quantile of Student's t distribution with $N-1$ degrees of freedom $\mathbf{t}_{N-1}$, namely $P\{|\mathbf{t}_N| \geq t\} = \alpha$, calculated for the given significance level α. This approach lead still to FAR=0% and FRR=22% showing that the standard deviation was greatly underestimated due to the small size of out database. We decided to replace the $t_\alpha / \sqrt{N}$ factor by a single number γ derived experimentally. The best results were obtained for $\gamma = 15$, namely FAR=0% and FRR=11.11%.

6.3 RCE network

The RCE network was tested in classification tasks, and showed better generalization abilities than the sigmoidal feedforward network, with correct classifications rate approaching 100%. On the other hand, this network failed in verification tasks, resulting in the FAR=8.33% and FRR=11.11%, much worse than for the perceptron.

References

[1] American Heritage Dictionary, Third Ed., ver. 3.6a, SoftKey International, Inc., 1994

[2] J.-J. Brault and R. Plamondon, "A Complexity Measure of Handwritten Curves: Modeling of Dynamic Signature Forgery", *IEEE Trans. On Systems, Man and Cybernetics*, vol. 23, No. 2, March/April 1993

[3] H. Crane and J.S. Ostrem, "Automatic Signature Verification Using a Three-Axis Force-Sensitive Pen", *IEEE Trans. On Systems, Man and Cybernetics*, vol. SMC-13, No. 3, May/June 1983

[4] G. Dimauro, S. Impedevo, G. Pirlo, and A. Salzo, "A Multi-Expert Signature Verification System for Bankcheck Processing", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, No. 5 (1992) pp. 827–844

[5] L.L. Lee, T. Berger, and E. Aviczer: "Reliable On-Line Human Signature Verification Systems", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 18, no. 6, June 1996

[6] S. Lee and J.C. Pan, "Offline Tracing and Representation of Signatures", *IEEE Trans. On Systems, Man and Cybernetics*, vol. 22, No. 4, July/August 1992

[7] Internet: http://www.cybersign.com

[8] Internet: http://www.signplus.com

[9] M. Mohamed and P. Gader, "Handwritten Word Recognition Using Segmentation-Free Hidden Markov Modeling and Segmentation-Based Dynamic Programming Techniques", *IEEE Trans. On Pattern Analysis And Machine Intelligence*, vol. 18, No.5, May 1996

[10] D.L. Reilly, *The RCE Neural Network*, Nestor, Inc., 1989

[11] T. Wessels, C.W. Omlin, "A Hybrid System for Signature Verification", *Transactions of IJCNN*, Como Italy, 2000