Mathematical Problems of Computer Science 39, 31--39, 2013.

Handwritten Signature Verification Using DRT

Vahe S. Khachaturyan

Institute for Informatics and Automation Problems of NAS of RA
e-mail: vahe@7smarts.com

Abstract

The purpose of this research is the development of mathematical and algorithmic support, which will improve the accuracy of signature verification. The algorithms compute the distances whilecomparing signatures based on DRT and HMM. For acceptance or rejection of the test signature a sliding threshold is used for all the authors, and depending on the author athresholdmethod is used, based on the distances between the test signature and the signatures of control, taking them as signs of the problem of two-class classification, using standard methods of imageclassification.

Keywords: Signature Verification, discrete Radon transform, hidden Markov model.

1. Introduction

The purpose of our research is to develop a system that automatically classifies handwritten signature images as authentic or fraudulent, with as little misclassifications as possible. At the same time, the processing requirements must be feasible so as to make the adoption of such an automated system economically viable.

Our work is inspired by, amongst others, the potential financial benefits that the automatic clearing of checks will have for the banking industry. Despite theincreasing number of electronic alternatives to paper checks, fraud perpetrated at financial institutions in the United States has become a national epidemic. The National Check Fraud Center Report of 2000 [1] states that: "...check fraud and counterfeiting are among the fastest-growing crimes affecting the United States' financial system, producing estimated annual losses exceeding $10 billion with the number continuing to rise at an alarming rate each year."

Since commercial banks pay little attention to verifying signatures on checks —mainly due to the number of checks that are processed daily - a system capable of screening casual forgeries should already prove beneficial. In fact, most forged checks contain forgeries of this type.

We developed a system that automatically authenticates documents based on the owner's handwritten signature. It should be noted that our system assumes that the signatures have already been extracted from the documents. Methods for extracting signature data from check backgrounds can be found in the following papers [2, 3, 4]. Our system will assist commercial banks in the process of screening checks and is not intended to replace the manual screening of checks entirely. Those checks the signatures of which do not sufficiently match a model of the owner’s genuine signature, are provisionally rejected. Generally, these rejected checks will constitute a small percentage of the total number of checks processed daily, and only these checks are selected for manual screening.

Since the introduction of computers, modern society has become increasingly dependent on the electronic storage and transmission of information. In many transactions, the electronic verification of a person’s identity proved beneficial and this inspired the development of a wide range of automatic identification systems.

Plamondon and Srihari [5] note that automatic signature verification systems occupy a very specific niche among other automatic identification systems: “On the one hand, they differ from systems based on the possession of something (key, card, etc.) or the knowledge of something (passwords, personal information, etc.), asthey rely on a specific, well learned gesture. On the other hand, they also differ from systems based on the biometric properties of an individual (finger prints, voice prints, retinal prints, etc.), asthe signature is still the most socially and legally accepted means of personal identification.”

Although handwritten signatures are by no means the most reliable means of personal identification, the signature verification systems are inexpensive and nonintrusive. Handwritten signatures provide a direct link between the writer’s identity and transaction, and are therefore perfect for endorsing transactions.

2. Image Processing

Each signature is scanned into a binary image at a resolution of 300 dots per inch, after which a median filtering is applied to remove speckle noise. On the average, a signature image has a width of 400 to 600 pixels and a height of 200 to 400 pixels. The image dimensions are not normalized.

Subsequently, the DRT of each signature is calculated. Each column of the DRT represents a projection or shadow of the signature at a certain angle. After these projections are processed and normalized, they represent a set of feature vectors (observation sequence) for the signature in question.

The DRT of an image is calculated as follows. Assume that each signature image consists of Ψ pixels in total, and that the intensity of the ith pixel is denoted by I_i, $i = 1, \ldots, \Psi$. The DRT is calculated using β not overlapping beams per angle and Θ angles in total. The cumulative intensity of the pixels that lie within the jth beam is denoted by R_j , $j = 1, \ldots, \beta\Theta$. This is called the jth beam sum. In its discrete form, the Radon transform can therefore be expressed as follows:

$$R_j = \sum_{i=1}^{\Psi} \omega_{ij} I_i, \; j = 1,2, \ldots ,\beta\theta, \qquad (1)$$

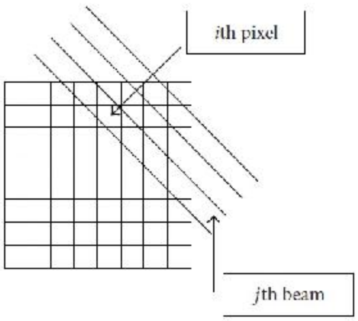

Fig 1. Discrete model for the Radon transform with$\omega_{ij} \approx 0.9$.

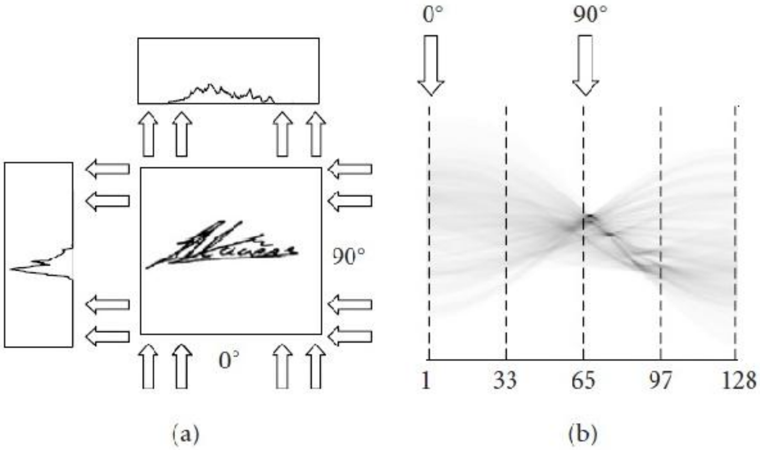

Fig 2. (a) A signature and its projections are calculated at angles of 0◦ and 90◦. (b) The DRT is displayed as a gray-scale image. This image has Θ = 128 columns, where each column represents a projection.

Here ω_{ij} indicates the contribution of the *i*th pixel to the *j*th beam sum (see Figure 1). The value of ω_{ij} is found through two-dimensional interpolation. Each projection therefore contains beam sums that are calculated at a given angle.

The accuracy of the DRT is determined by Θ (the number of angles), β (the number of beams per angle), and the accuracy of the interpolation method.

Note that the continuous form of the Radon transform can be inverted through analytical means. The DRT therefore contains almost the same information as the original image and can be efficiently calculated with an algorithm by Bracewell [8].

Our system calculates the DRT at Θ angles. These angles are equally distributed between 0◦ and 180◦. A typical signature and its DRT are shown in Figure 2. The dimension of each projection is subsequently altered from β to d. This is done by first decimating all the zero-valued components from each projection. These decimated vectors are then shrunk or expanded to a length of d through interpolation. Although almost all the information in the original signature image is contained in the projections at angles that range from 0◦ to 180◦, the projections at angles that range from 180◦ to 360◦ are also included in the observation sequence. These additional projections are added to the observation sequence in order to ensure that the sequence fits the topology of our HMM (see Section 3.2). Since these projections are simply reflections of the projections already calculated, no additional calculations are necessary. An observation sequence therefore consists of T = 2Θ feature vectors, that is, $X_1^T = x_1, x_2, \ldots, x_T$ .Each vector is subsequently normalized by the variance of the intensity of the entire set of T feature vectors. Each signature pattern is therefore represented by an observation sequence that consists of T observations, where each observation is a feature vector of dimension d. The experimental results and computational requirements for various values of d and Θ are discussed in Section 5, respectively.

The DRT, as a feature extraction technique, has several advantages. Although the DRT is not a shift invariant representation of a signature image, the shift and scale invariance is ensured by the subsequent image processing. Each signature is a static image and contains no dynamic information. Since the feature vectors are obtained by calculating projections at different angles, a simulated time evolution is created from one feature vector to the next, where the angle is the dynamic variable. This enables us to construct an HMM for each signature (see Section 3). The DRT is calculated at angles that range from 0◦ to 360◦ and each observation sequence is then modeled by an HMM the states of which are organized in a ring (see Section 3.2). This ensures that each set of feature vectors is rotation invariant. Our system is also robust with respect to moderate levels of noise. These advantages are now discussed in more detail.

Noise

We explained earlier in this section that the zero-valued components of each projection are decimated before the remaining non-zero components are shrunk or expanded through interpolation. In this way, a feature vector with the required dimension is obtained. The decimation of the zero-valued components ensures that moderate levels of noise (which are represented by a few additional small-valued components within certain projections) are "attached" to the other nonzero components before the decimated vector is shrunk or expanded. Since the dimension of the feature vectors are high compared to the number of these additional components, the incorporation of these components has little effect on the overall performance of the system.

Shift invariance

Although the DRT is not a shift invariant representation of a signature image, the shift invariance is ensured by the subsequent image processing. The zero-valued components of each projection are decimated and the corresponding feature vector is constructed from the remaining components only.

Rotation invariance

The DRT is calculated at angles that range from 0° to 360° and each set of feature vectors is then modeled by an HMM the states of which are organized in a ring (see Section 3.2). Each signature is therefore represented by a set of feature vectors that is rotation invariant.

Scale invariance

For each projection, the scale invariance has to be achieved in the direction perpendicular to the direction in which the image is scanned, that is, perpendicular to the beams, and in the direction parallel to the beams. The scale invariance perpendicular to the beams is ensured by shrinking or expanding each decimated projection to the required dimension. The scale invariance parallel to the beams is achieved by normalizing the intensity of each feature vector. This is achieved by dividing each feature vector by the variance of the intensity of the entire set of feature vectors.

3. Signature Modelling

We use a first-order continuous observation HMM to model each writer's signature. For a tutorial on HMMs, the reader is referred to a paper by Rabiner [9] and the book by Deller et al.

Notation

We use the following notation for an HMM λ.

(1) We denote the *N* individual states as

$$S = \quad s_1, s_2, \ldots, s_N \qquad (2)$$

and the state at time t as q_t .

(2) The initial state distribution is denoted by $\pi = \{\pi_i\}$, where

$$\pi_i = P \;\; q_1 = s_i \;, i = 1, \ldots , N. \qquad (3)$$

(3) The state transition probability distribution is denoted by $A = \{a_{i,j}\}$, where

$$a_{ij} = P \;\; q_{t+1} = s_j \;\; q_t = s_i), \;\; i = 1, \ldots, N, \;\; j = 1, \ldots , N. \qquad (4)$$

(4) The probability density function (pdf),

which quantifies the similarity between a feature vector x and the state s_j, is denoted by

$$f \;\; x \;\; s_j, \lambda \;\; , \;\; j = 1, \ldots , N. \qquad (5)$$

HMM topology

We use an HMM, the states of which are organized in a ring (see Figure 3). Our model is equivalent to a left-to-right model, but a transition from the last state to the first oneis allowed. Since the HMM is constructed in such a way that it is equally likely to enter the model at any state, and the feature vectors are obtained from all the projections, that is, the projections calculated at angles ranging from 0◦ to 360◦, the ring topology of our HMM guarantees that the signatures are rotation invariant. Each state in the HMM represents one or more feature vectors that occupy similar positions in a d-dimensional feature space. This implies that the HMM groups certain projections (columns of the DRT) together. It is important to note that this segmentation process only takes place after some further image processing has been conducted on the original projections.

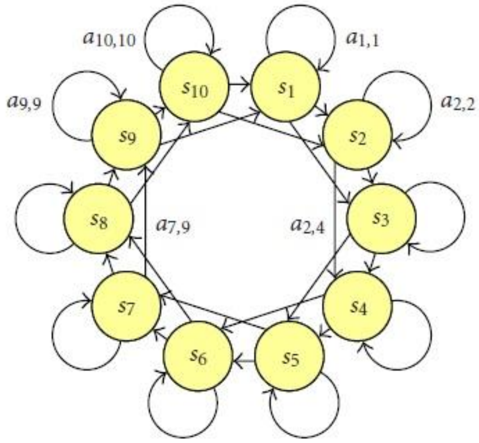

Fig 3. An example of an HMM with a ring topology. This model has ten states with one state skip.

Training

Each model is trained using the Viterbi reestimation technique. The dissimilarity between an observation sequence X and a model λ can therefore be calculated as follows (see [9]):

$$d\ X,\lambda\ = -\ln\ f\ X|\lambda\ , \qquad (6)$$

In real-world scenarios, each writer can only submit a small number of training samples when he or she is enrolled into the system. Since our algorithm uses feature vectors with a high dimension, the reestimated covariance matrix of the pdf for each state is not reliable and may even be singular. A Mahalanobis distance measure therefore cannot be found. Consequently, these covariance matrices are not reestimated and are initially set to *0.5I*, where *I*is the identity matrix. Only the mean vectors are reestimated, which implies that the dissimilarity values are based on theEuclidean distance measure.

We assume that training signatures, genuine test signatures, and forgeries are available for only a limited number of writers, that is, for the writers in our database. No forgeries are used in the training process since our system aims to detect only skilled and casual forgeries, and these types of forgeries are not available when our system is implemented. The genuine test signatures and forgeries are used to determine the error rates for our system (see Section 5). Assuming that there are W writers in our database, the training signatures for each writer are used to construct an HMM, resulting in W models, that is $\{\lambda_1, \lambda_2, \dots, \lambda_W\}$.

When the training set for the writer w is denoted by $X_1^{(w)}, X_2^{(w)}, \dots, X_{N_w}^{(w)}$, where N_w is the number of samples in the training set, the dissimilarity between every training sample and the model is used to determine the following statistics for the writer's signature:

$$\mu_w = \frac{1}{N_w} \sum_{i=1}^{N_w} d\ \ X_i^{\ w}\ , \lambda_w\ \ ,$$

$$\sigma_w^2 = \frac{1}{N_w} \sum_{i=1}^{N_w}\ d\ \ X_i^{\ w}\ , \lambda_w\ \ - \mu_w\ \ ^{2}. \qquad (7)$$

4. Verification

When a system aims to detect only random forgeries, the subsets of the other writer’s training sets can be used to model “typical” forgeries. This is called “an impostor validation” and can be achieved through strategies like test normalization. These techniques enable one to construct verifiers that detect random forgeries very accurately (see [6, 7]). Since we aim to detect only skilled and casual forgeries, and since the models for these forgeries are generally unobtainable, we are not able to utilize any of these impostor validation techniques. We also do not use any subset of genuine signatures for validation purposes.

Our verifier is constructed as follows. When a claim is made that the test pattern $X_{Test}^{(w)}$ belongs to the writer *w*, the pattern is first matched with the model λ_w through Viterbi alignment. This match is quantified by $f(X_{Test}^{w}\ |\lambda_w)$. The dissimilarity between the test pattern and the model is then calculated as follows (see [9]):

$$d\ \ X_{Test}^{w}\ \lambda_w\ \ = -\ln\ f(X_{Test}^{w}\ |\lambda_w)\ \ . \qquad (8)$$

In order to use a global threshold for all writers, Dolfing [6] suggests that every dissimilarity value in (8) is normalized, using the statistics of the claimed writer’s signature, that is, (7):

$$d_{Mah}\ \ X_{Test}^{w}\ \lambda_w\ \ = \frac{d\ \ X_{Test}^{w}\ \lambda_w\ \ -\mu_w}{\sigma_w}, \qquad (9)$$

where$d_{Mah}\ \ X_{Test}^{w}\ \lambda_w$ denotes the normalized dissimilarity between the test pattern and the model of the claimed writer’s signature. This normalization is based on the assumption that the dissimilarity value in (8) is based on the Mahalanobis distance measure.

For mean vectors the dissimilarity value in (8) is based on the Euclidean distance measure. When this is the case, we found that significantly better results are obtained when the standard deviation of the dissimilarities of the training set, that is, σ_w in (9), is replaced by the mean μ_w, that is,

$$d_{Eucl}\ \ X_{Test}^{w}\ \lambda_w\ \ = \frac{d\ \ X_{Test}^{w}\ \lambda_w\ \ -\mu_w}{\mu_w}. \qquad (10)$$

A sliding threshold τ, where τ □ (−∞,∞), is used to determine the error rates for the test patterns. When $d_{Eucl}\ \ X_{Test}^{w}\ \lambda_w\ \ < \tau$, that is,

$$d\ \ X_{Test}^{w}\ \lambda_w\ \ < \mu_w\ \ 1+\tau\ \ , \qquad (11)$$

the claim is accepted, otherwise, the claim is rejected. When τ = 0, all the test patterns, for which $d\ \ X_{Test}^{w}\ \lambda_w\ \ \geq \mu_w$, are rejected. This almost always results in an FRR close to 100% and an FAR close to 0%.When $\tau \rightarrow \infty$, all the test patterns, for which $d\ \ X_{Test}^{w}\ \lambda_w$ is finite, are accepted. This always results in an FRR of 0% and an FAR of 100%.

This technique is simple to apply in a program coding and further performance can be improved by using a better classification algorithm.

5. Experiments

Our data set contains 924 signatures from 22 writers. Ten training signatures were obtained from each writer during theinitial enrollment session. Thirty-two test signatures, that consist of 20 genuine signatures, 6 skilled forgeries, and 6 casual forgeries, were subsequently obtained over a period of two weeks. The 20 genuine test signatures consist of two sets of 10 signatures each. These signatures were supplied by the same writers one week and two weeks after the enrollment session. The forgeries were obtained from 6 forgers. The casual forgeries were obtained first. Only the name of the writer was supplied to the forgers and they did not have access to the writer's signatures. The skilled forgeries were then obtained from the same group of forgers. They were provided with several samples of each writer's genuine signature and were allowed ample opportunity to practice. Each forger submitted 1 casual forgery and 1 skilled forgery for each writer. The writers were instructed to produce each signature within an appropriate rectangular region on a white sheet of paper. The signatures were then digitized with a flatbed scanner at a resolution of 300 dots per inch. The genuine signatures were produced with different pens and the forgeries were produced with the same pens that were used for producing the genuine signatures. These signatures are free of excessive noise, smears, and scratches.

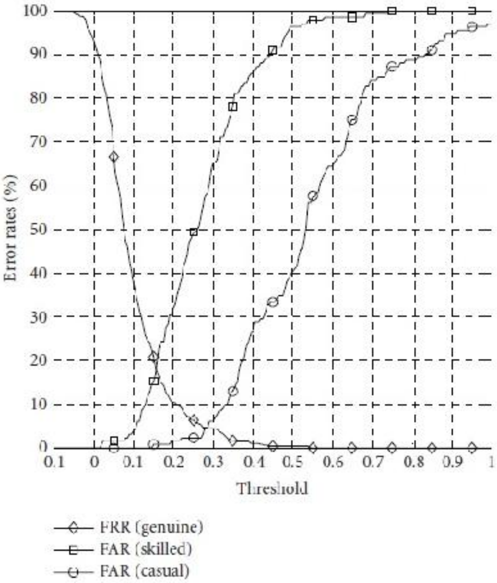

Fig. 4. The Stellenbosch data set. Graphs for the FRR and the FAR when d = 512, Θ = 128, N = 64, and $\ell = 1$.

Results

Let ℓ denote the number of allotted forward links in our HMM. Figure 4 shows the FRR and FAR as functions of our threshold parameter τ □ [−0.1, 1], when d = 512, Θ = 128, N = 64, and $\ell = 1$. The FRR and FAR for a test set that contains only skilled forgeries, and the FAR for a test set that contains only casual forgeries are plotted on the same system of axes. When, for example, a threshold of τ = 0.16 is selected, equation (11) implies that all the test patterns, for which $d\ X_{Test}^{w}\ \lambda_w\ \geq 1.16\mu_w$, are rejected;the other patterns are accepted. When only the skilled forgeries are considered, this threshold selection will ensure an EER of approximately 18%. When only the casual forgeries are considered, our algorithm achieves an EER of 4.5%.

It is clear that when the dimension of the feature vectors is decreased from d = 512 to d = 256 or even to d = 128, the performance of the system is not significantly compromised. The performance of our system is generally enhanced when the number of feature vectors, that is, $T=2\theta$, or the number of states in the HMM, that is, N, is increased. The best results are obtained when only one forward link is allowed in the HMM, that is, when $\ell = 1$.

6. Conclusions

The DRT is a stable and robust method of feature extraction. The DRT creates a simulated time evolution from one feature vector to the next one and enables us to model a signature with an HMM. Our system is not sensitive to moderate levels of noise, and the feature vectors are extracted in such a way that rotation, shift, and scale invariance is ensured.

Our system does not outperform all these systems. These systems do, however, utilize either a technique or features that are fundamentally very different from ours. This implies that it is very likely that a combination of their systems and that of ours will result in a superior merged system, making their approaches complementary to ours. We also expect a significant improvement in our results when local features are incorporated into our algorithm. This is currently being investigated.

References

[1] National Check Fraud Center, *National Check Fraud Center Report*, 2000.

[2] S. Djeziri, F. Nouboud, and R. Plamondon, “Extraction of signatures from cheque background based on a filiformity criterion,” *IEEE Trans. Image Processing*, vol. 7, no. 10, pp. 1425– 1438, 1998.

[3] A. L. Koerich and L. L. Lee, “Automatic extraction of filledin information from bankchecks based on prior knowledge about layout structure,” in *Advances in Document Image Analysis: First Brazilian Symposium*, Lecture Notes in Computer Science, vol. 1339, pp. 322–333, 1997.

[4] J. E. B. Santos, F. Bortolozzi, and R. Sabourin, “A simple methodology to bankcheck segmentation,” in *Advances in Document Image Analysis: First Brazilian Symposium*, Lecture Notes in Computer Science, vol. 1339, pp. 334–343, 1997.

[5] R. Plamondon and S. N. Srihari, “On-line and off-line handwriting recognition: a comprehensive survey,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63–84, 2000.

[6] R. Sabourin, G. Genest, and F. Prˆeteux, “Off-line signature verification by local granulometric size distributions,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 976–988, 1997.

[7] A. El-Yacoubi, E. J. R. Justino, R. Sabourin, and F. Bortolozzi, “Off-line signature verification using HMMs and cross-validation,” in *IEEE International Workshop on Neural Networks for Signal Processing*, pp. 859–868, 2000.

[8] R. N. Bracewell, *Two-Dimensional Imaging*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1995.

[9] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

Submitted 10.12.2012, accepted 15.02.2013.

Ձեռագիր ստորագրություններ իստուգումը DRT-ի օգտագործմամբ

Վ. Խաչատուրյան

Ամփոփում

Տվյալ աշխատանքի նպատակն է մաթեմատիկական և ալգորիթմային ապահովման մշակումը, որը թույլ կտա բարձրացնել ստորագրության ստուգման ճշգրտությունը: Ալգորիթմները հաշվարկում են հեռավորությունները ստորագրությունների համեմատման ժամանակ՝ օգտագործելով DRT-ն և HMM-ը: Փորձարկվող ստորագրության ընդունման կամ մերժման համար օգտագործվում է սահող շեմքի մեթոդը բոլոր հեղինակների համար և կախված հեղինակից՝ շեմքի մեթոդը՝ օգտագործելով փորձարկվող ստորագրության և ստուգողական ստորագրությունների միջև եղած հեռավորությունները՝ դասակարգելով դրանք երկու դասերի՝ պատկերների դասակարգման ստանդարտ մեթոդների միջոցով:

Проверка рукописных подписей с использованием DRT

В. Хачатурян

Аннотация

Целью данного исследования является разработка математического и алгоритмического обеспечения, которая позволит повысить точность проверки подписи. Алгоритмы вычисляют расстояния при сравнении подписей с помощью DRT и HMM. Для принятия или отказа тестовой подписи используется метод скользящего порога для всех авторов и зависящий от автора – метод порога, с использованием расстояний между тестовой подписью и контрольными подписями, классифицируя их по двум классам, с использованием стандартных методов классификации изображений.