Learning Strategies and Classification Methods for Off-line Signature Verification†

Sargur N. Srihari, Aihua Xu and Meenakshi K. Kalera
Center of Excellence of Document Analysis and Recognition
University at Buffalo, the State University of New York
Buffalo, NY 14260, USA
{srihari, aihuaxu, mkkalera}@cedar.buffalo.edu

Abstract

Learning strategies and classification methods for verification of signatures from scanned documents are proposed and evaluated. Learning strategies considered are writer-independent– *those that learn from a set of signature samples(including forgeries) prior to enrollment of a writer, and* writer dependent– *those that learn only from a newly enrolled individual. Classification methods considered include two distance based methods (one based on a threshold, which is the standard method of signature verification and biometrics, and the other based on a distance probability distribution), a Nave Bayes (NB) classifier based on pairs of feature bit values and a support vector machine (SVM). Two scenarios are considered for the writer-dependent scenario: (i) without forgeries (one-class problem) and (ii) with forgery samples being available (two-class problem). The features used to characterize a signature capture local geometry, stroke and topology information in the form of a binary vector. In the one-class scenario distance methods are superior while in the two-class SVM based method outperforms the other methods.*

1. Introduction

Automatic verification of signatures from scanned paper documents has many applications such as authentication of bank checks, questioned document examination, biometrics, etc. On-line, or dynamic, signature verification systems have been reported with high success rates [15]. However, off-line, or static research is relatively unexplored– which difference can be attributed to the lack of temporal information, the range of intra-personal variation in the scanned image, etc.

[1]This work was supported in part by the U.S. Department of Justice, National Institute of Justice grant 2002-LT-BX-K007

Methods have been described for both writer-dependent (WD) and writer-independent (WI) signature verification. WD models extract features from genuine signatures of a specific writer and are trained for that writer. The questioned signature is compared against the model for that writer. This is the standard approach to signature verification [14]. Based on a writer-independent approach to determining whether two handwritten documents– not just signatures– were written by the same person or not [19], a writer independent (WI) signature verification method was proposed in [9]. In the WI model the probability distributions of within-writer and between-writer similarities, over all writers, are computed in the training phase. These distributions are used to determine the likelihood of whether a questioned signature is authentic.

2. Learning strategies

Signature verification is a problem that can be approached using machine learning techniques. A set of samples of signatures, D, can be prepared with the help of several individuals. The parameters derived from such a set can be used in determining whether an arbitrary pair of signatures, e.g., a questioned signature and a genuine signature, match or not. One can also learn from samples of a specific individual and use only these parameters (or model) in matching for that individual.

These two learning strategies are: writer-independent (WI) and writer-dependent (WD), as shown in Fig. 1. In WI learning D_g and D_f are the training data sets of genuine and forged signatures from several writers. A model S is trained from pairs of samples (genuine-genuine and genuine-forgery) from D_g and D_f. Given a questioned signature Q and a set K of genuine signatures for individual w , S is used to determine whether Q is accepted as genuine. In WD learning, only the genuines for individual w, i e., the set K, is used to determine the model S, which is then

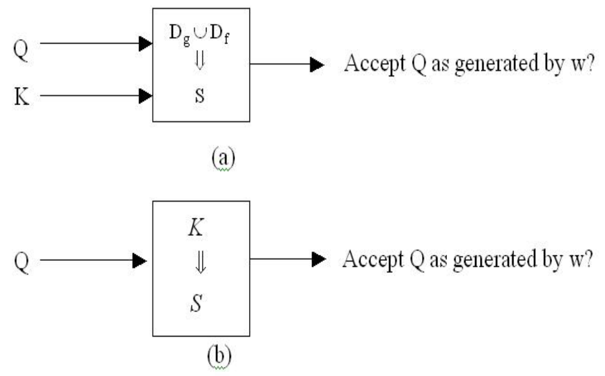

Figure 1. Verification Models: (a) writer independent: a questioned (Q) is matched against a set of genuines K using a model S derived from genuines and forgeries of other writers, and (b) writer dependent: a model for an individual is determined using only K.

used to determine whether Q is accepted as genuine.

3. Test-Bed

A database of off-line signatures was prepared as a test bed. Each of 55 individuals contributed 24 signatures– thereby creating 1320 genuine signatures. Some were asked to forge three other writers' signatures, eight times per subject, thus creating 1320 forgeries. One example of each of 55 genuines are shown in Figure 2. Ten examples of genuines of one subject (subject no. 21) and ten forgeries of that subject are shown in Fig. 3

3.1. Image Preprocessing

Each signature was scanned at 300 dpi gray-scale and binarized using a gray-scale histogram. Salt-and-pepper noise was removed by median filtering. Slant normalization was performed by extracting the axis of least inertia and rotating the curve until this axis coincides with the horizontal axis [7]. Given an $M \times N$ image, $G = (u_k, v_k) = (x_{(i,j)}, y_{(i,j)} | x_{(i,j)} \neq 0, y_{(i,j)} \neq 0)$. Let S be the size of G, and let $\overline{u} = \frac{1}{S}\sum_k u_k$ and $\overline{v} = \frac{1}{S}\sum_k v_k$ be the coordinates of the center of mass of the signature. The orientation of the axis of least inertia is given by the orientation of the least eigenvector of the 2x2 matrix $I = \begin{pmatrix} \overline{u^2} & \overline{uv} \\ \overline{uv} & \overline{v^2} \end{pmatrix}$ where $\overline{u^2} = \frac{1}{S}\sum_k (u_k - \overline{u})^2$, $\overline{v^2} = \frac{1}{S}\sum_k (v_k - \overline{v})^2$ and $\overline{uv} = \frac{1}{S}\sum_k (u_k - \overline{u})(v_k - \overline{v})$ are the second order moments

Figure 2. Genuine signature samples.

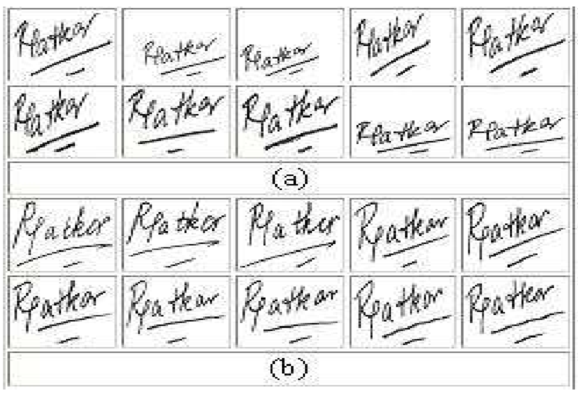

Figure 3. Samples for one writer: (a) genuines and (b) forgeries.

Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004)
0-7695-2187-8/04 $20.00 © 2004 **IEEE**

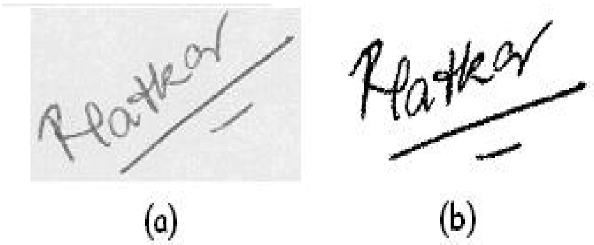

Figure 4. Pre-processing: (a) original (b) final.

of the signature [12]. The result of binarization and slant normalization of a gray-scale image is shown in Fig. 4.

3.2. Feature Extraction

Features for static signature verification can be one of three types [5, 10]: (i) *global*: extracted from every pixel that lie within a rectangle circumscribing the signature, including image gradient analysis [16] , series expansions [11], etc., (ii) *statistical*: derived from the distribution of pixels of a signature, e.g., statistics of high gray-level pixels to identify pseudo-dynamic characteristics [1], (iii) *geometrical and topological*: e.g., local correspondence of stroke segments to trace signatures [6], feature tracks and stroke positions [5], etc. A combination of all three types of features were used in a writer independent (WI) signature verification system [9]– which were previously used in character recognition [20], word recognition [21] and writer identification [19]. These features, known as gradient, structural and concavity (or GSC) features were used here.

The average size of all reference signature images was chosen as the reference size to which all signatures were resized. The image is divided into a 4x8 grid from which a 1024-bit GSC feature vector is determined (Fig. 5). Gradient (G) features measure the magnitude of change in a 3 x 3 neighborhood of each pixel. For each grid cell, by counting statistics of 12 gradients, there are 4x8x12 = 384 gradient features. Structural (S) features capture certain patterns, i.e., mini-strokes, embedded in the gradient map. A set of 12 rules is applied to each pixel– to capture lines, diagonal rising and corners– yielding 384 structural features. Concavity (C) features, which capture global and topological features, e.g., bays and holes, are 4x8x8 = 256 in number.

3.3. Distance Measure

A method of measuring the similarity or distance between two signatures in feature space is essential for classification. The correlation distance performed best for GSC

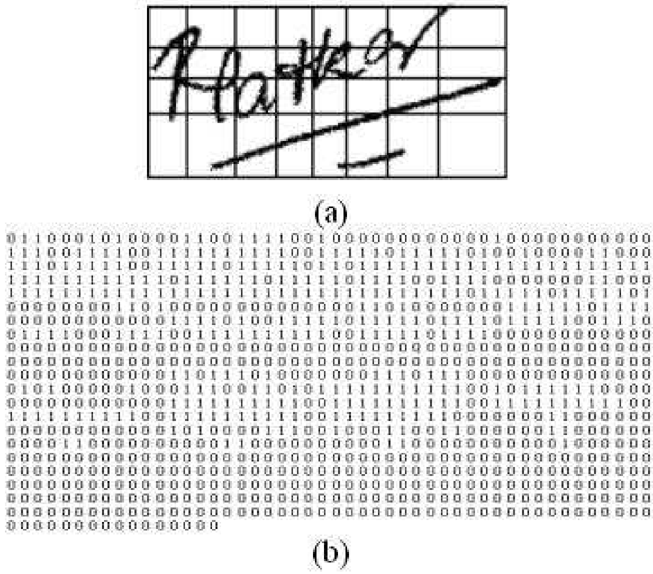

Figure 5. Features: (a) variable grid, and (b) feature vector.

binary features [22]which is defined for two binary vectors X and Y, as follows:

$$d(X,Y) = \frac{1}{2} - \frac{s_{11}s_{00} - s_{10}s_{01}}{2((s_{10}+s_{11})(s_{01}+s_{00})(s_{11}+s_{01})(s_{00}+s_{10}))^{1/2}}$$

where s_{ij} represent the number of corresponding bits of X and Y that have values i and j. Both the $WI-DS$ and $WD-DT$ methods described below use $d(X,Y)$ as the distance measure.

4. Writer-independent Verification

The objective is to determine whether pair (K,Q) belongs to the same individual, where Q is a questioned signature and K is a set of known signatures for that individual. Two WI classification methods, distance statistics [19]and naive Bayes, are presented below.

4.1. Distance Statistics (DS) Method

The verification approach of [19]is based on two distributions of distances $d(X,Y)$: genuine-genuine and genuine-forgery pairs. The distributions are denoted as P_g and P_f respectively. The means and variances of of $d(X,Y)$ where X and Y are both genuine and X is genuine and Y is a forgery are shown in Figure 6– where the the number of writers varies from 10 to 55. Here 16 genuines are 16 forgeries were randomly chosen from each subject. For each n, there are two values corresponding to the mean and variance of n x C_2^{16} pairs of same writer (or genuine-genuine pair) distances and n x 16^2 pairs of different writer

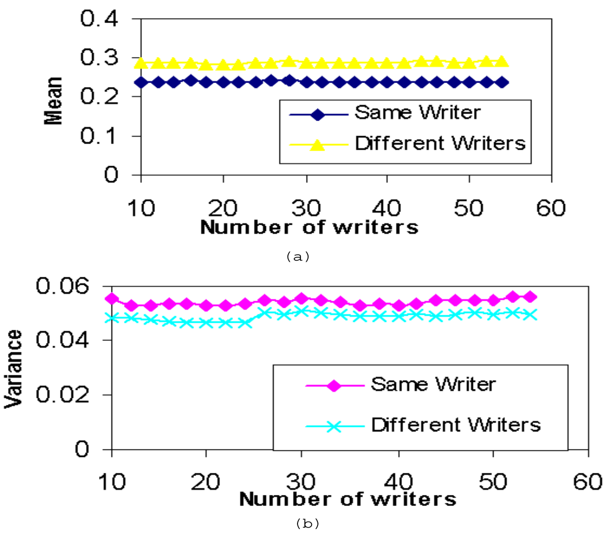

Figure 6. Statistics of genuine-genuine and genuine-forgery distances.

(or genuine-forgery pair) distances. The values are close to constant with μ_g = 0.24 and μ_f = 0.28 with corresponding variances $\sigma_g = 0.055$ and $\sigma_f = 0.05$. Given a questioned signature Q and a single known signature K the probabilities of $d(K, Q)$ are: $P(genuine|Q) = P_g(d(K, Q))$ and $P(forged|Q) = P_f(d(K, Q))$. Q is accepted as genuine if the genuine probability exceeds the forgery probability. Normal distributions are assumed for genuine-genuine distances and genuine-forgery distances.

Generalization to n genuines: When there are n genuine signatures available, i.e., $|K| > 1$, given a questioned signature Q, $P(genuine|Q) = \prod_{j=1}^{n} P_g(d(K_j, Q))$

$P(forged|Q) = \prod_{j=1}^{n} P_f(d(K_j, Q))$

4.2. Naïve Bayes (NB) Method

Rather than determining the distributions of distances between two feature vectors, each pair of corresponding bits in the questioned and known feature vectors can be treated as random variables. The pairs corresponding to different positions in the feature vector are considered to be independent and identically distributed-which is the naive Bayes (NB) assumption. Let feature vectors $X = \{x_1, x_2, ..., x_n\}$ and $Y = \{y_1, y_2, ..., y_n\}$. The probabilities of i^{th} same-value bits in a genuine-genuine pair and a genuine-forgery pair are computed using:

$$P_{s,x_i=y_i} = \frac{|(X,Y)|x_i=y_i, X,Y \in D_g|}{|(X,Y)|X,Y \in D_g|}$$

Table 1. Writer-independent methods with 1 and 16 training samples

Methods(n)	FRR(%)	FAR(%)	AER(%)
Distance Stats(1)	27.6	27.8	27.7
Naive Bayes(1)	27.2	26.0	26.6
Distance Stats(16)	21.3	22.1	21.7
Naive Bayes(16)	22.9	24.1	23.5

$$P_{s,x_i \neq y_i} = \frac{|(X,Y)|x_i \neq y_i, X,Y \in D_g|}{|(X,Y)|X,Y \in D_g|}$$

$$P_{d,x_i=y_i} = \frac{|(X,Y)|x_i=y_i, X \in D_g, Y \in D_f|}{|(X,Y)|X \in D_g, Y \in D_f|}$$

$$P_{d,x_i \neq y_i} = \frac{|(X,Y)|x_i \neq y_i, X \in D_g, Y \in D_f|}{|(X,Y)|X \in D_g, Y \in D_f|}$$

where , D_g and D_f are the training sets of genuine and forged signatures. Knowing the probabilities of the values of the bit pair for each feature, given (K, Q), the overall genuine-genuine and genuine-forgery probabilities are calculated as the product of the probabilities for all feature pairs, i.e.,

$$P(genuine|Q) = P_s(K, Q) = \prod_{i=1}^{1024} P_{s,k_i=q_i}^{k_i \otimes q_i} P_{s,k_i \neq q_i}^{k_i \oplus q_i}$$

$$P(forged|Q) = P_d(K, Q) = \prod_{i=1}^{1024} P_{d,k_i=q_i}^{k_i \otimes q_i} P_{d,k_i \neq q_i}^{k_i \oplus q_i}$$

They are compared to determine whether they are from the same writer or not. Generalization to n genuines is as follows: $P(genuine|Q) = \prod_{j=1}^{n} P_s(K_j, Q)$

$P(forged|Q) = \prod_{j=1}^{n} P_d(K_j, Q)$

4.3. Performance

The two writer independent methods were evaluated using the test bed. False reject rate (FRR), false accept rate (FAR) and average error rate (AER = (FRR + FAR) / 2) were determined. To calculate probabilities 16 genuines and 16 forgeries from each subject were randomly chosen as the training set and the rest are used as test set. FRR, FAR and AER of two methods are shown in Table 1. The probabilities in WI-DS and feature probabilities in WI-NB, the product of 16 distance probabilities in WI-DS or the product of feature p. The WI-DS and WI-NB were evaluated with 16 genuines in training compared to one in Section 4.1. Instead of original distance probabilities in WI-DS and feature probabilities in WI-NB, the product of 16 distance

Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004)
0-7695-2187-8/04 $20.00 © 2004 **IEEE**

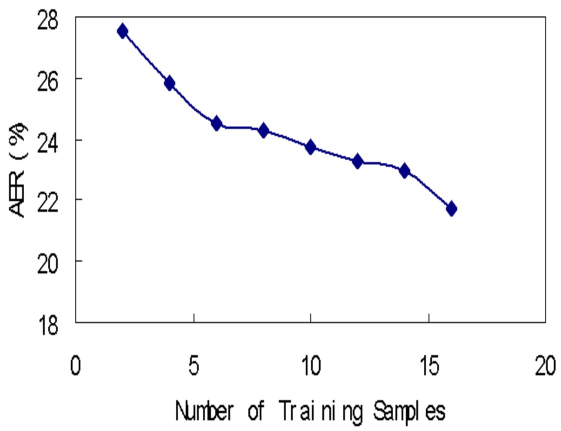

Figure 7. Average Error Rate of Writer Independent Distance Statistics method.

probabilities in WI-DS or the product of feature probabilities in WI-NB were used.

With both methods performance increases with more training genuines. For training, n genuines were randomly chosen. The test set consisted of 8 genuines from the rest and 24 forgeries. WI-DS has the best performance– Fig. 7 shows performance improvement of WI-DS with n.

5. Writer-dependent Verification

Assuming that there exist sufficient training genuines, a machine of S is learned only from the training data for a specific individual. Four classification methods were considered: distance threshold (which is the standard method used of biometrics), distance statistics, naive Bayes and SVM. Two sub-formulations are considered: one-class where forgeries for the individual are unavailable, and two-class where genuines and forgeries are available.

5.1. Training with Genuines only

Distance Threshold (DT): The DT method is the common signature verification model. The first step is to enroll genuines K as reference signatures. The distance $d(X, Y)$ is computed for each pair (X, Y) in K to determine the threshold $thres = max\{d(X, Y)|X, Y \in K\}$. Given a questioned signature Q, the average of $\{d(Q, Y)|Y \in K\}$, denoted as $dist$, is computed. If $dist < thres$, then Q is accepted as a genuine; and rejected otherwise.

Distance Statistics (DS): Here the genuine-genuine distance distribution is obtained only from K, i.e., the mean and variance of P_g are determined from $\{d(X, Y)|X, Y \in K\}$. The genuine-forgery distance distribution P_f is from D_g and D_f as in $WI - DS$.

Table 2. One -class writer-dependent methods (trained with genuines only).

Methods(n)	FRR(%)	FAR(%)	AER(%)
Distance Threshold	22.5	19.5	21.5
Distance Statistics	23.0	21.7	22.4
Naive Bayes	25.9	24.1	25.0
One-Class SVM	47.6	44.4	46.0

Naive Bayes (NB): Let $X = \{x_1, x_2, ..., x_n\}$ where $X \in K$. Two distributions are computed from K: Given a test signature $Q = \{q_1, q_2, ..., q_n\}$, the likelihood, $P(genuine|Q) = \prod_{i=1}^{n} P_{s,x_i=q_i}$. A common optimal threshold $thres$ for the likelihoods is trained for all writers. Q is accepted as a genuine if $P(genuine|Q) \geq thres$.

Support Vector Machine (SVM): While SVM classification is popular in applications [8, 13] , its use in signature verification is unknown. SVMs match the requirements of signature verification: high dimensionality and sparse instance spaces. The GSC feature space is high-dimensional (1024) and very sparse. One-class SVMs [18] attempt to distinguish genuines from the rest given only genuine data– by drawing the class boundary of the genuine data set in feature space.

Experimental results for four methods, with a training set size of 16, are shown in Table 2. Here the distance threshold performs best with SVM being very poor.

5.2. Training with Genuines and Forgeries

Forgeries were included in training in the following experiments. In $WD - DT$, since the threshold is determined only by genuines, it is unchanged. However in $WD - DS$, instead of gathering genuine-forgery distance distribution from all writers, such distribution is generated directly from the genuine-forgery pairs for the individual. Similarly in $WD - NB$, 0 and 1 distributions of each feature in forgeries are generated from the feature vectors in the forgery set.

In an SVM, hyperplanes are determined by support vectors instead of training samples. Due to unbalanced training datasets, instead of finding equal size maximal margin on both sizes of optimal hyperplanes, margins are dynamically adjusted according to sample sizes. With more positive samples than negative samples, different penalty parameters are used to balance weights. Given training vectors $x_i \in R^n, i = 1, ..., l$, in two classes, and a vector $y \in R^l$ such that $y_i \in \{1, -1\}$, the formation in [2] solves the clas-

Table 3. Two-class writer-dependent methods(trained with 16 genuines and 5 forgeries)

Methods(n)	**FRR(%)**	**FAR(%)**	**AER(%)**
Distance Statistics	17.6	20.7	19.2
Naive Bayes	9.95	13.0	11.45
SVM	8.5	10.1	9.3

sification problem for unbalanced data.

For each writer, 16 genuines were randomly selected from the genuine set and 5 forgeries selected from the forgery set by one forger. The other 8 genuines and 16 forgeries by other forgers constitute the test set. Table 3 presents the classification results showing that SVM outperforms others.

6. Conclusions

Several learning strategies for signature verification were evaluated using a high-dimensional feature space that captures both local geometric information as well as stroke information. In the writer-independent case, the newly introduced distance statistics method outperformed classical distance threshold and naive Bayes approaches. In the writer dependent case distance threhold performed best with distance statistics being close. The distance statistics method has the advantage that it can be used when few training examples, even one, are available, and it generates a match likelihood rather than a distance score.

Acknowledgement

The authors wish to thank Hansheng Lei, Siyuan Chen and Chen Huang for their assistance.

References

[1] M. Ammar, Y. Yoshido, and T. Fukumura. A new effective approach for off-line verification of signatures by using pressure features. *Proceedings of the 8th International Conference on Pattern Recognition*, pages 566–569, 1986.

[2] B. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 1992.

[3] C. C. Chang and J. L. C. Libsvm: a library for support vector machines. *http://www.csie.ntu.edu.tw/ cjlin/libsvmtools*, 2001.

[4] Y. Chen, X. Zhou, and T. S. Huang. One-class svm for learning in image retrieval. *IEEE Proceedings of International Conference on Image Processing*, pages 440–447, 1999.

[5] Fang, C. H. Leung, Y. Y. Tang, K. W. Tse, P. C. K. Kwok, and Y. K. Wong. Off-line signature verification by the tracking of feature and stroke positions. *Pattern Recognition*, 36:91–101, 2003.

[6] J. K. Guo, D. Doermann, and A. Rosenfeld. Local correspondence for detecting random forgeries. *Proceedings of the International Conference on Document Analysis and Recognition*, pages 319–323, 1997.

[7] B. Horn. Robot vision. *MIT Press*, 1986.

[8] T. Joachims. Text categorization with support vector machines: Learning with many relevant features. *Proceedings of the European Conference on Machine Learning*, pages 137–142, 1998.

[9] M. K. Kalera, B. Zhang, and S. N. Srihari. Off-line signature verification and identification using distance statistics. *Proceedings of the International Graphonomics Society Conference, Scottsdale, AZ,*, pages 228–232, November 2003.

[10] S. Lee and J. C. Pan. Off-line tracing and representation of signatures. *IEEE Transactions on Systems, Man and Cybernetics*, 22:755–771, 1992.

[11] C. C. Lin and R. Chellappa. Classification of partial 2-d shapes using fourier descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:696–690, 1997.

[12] M.E.Munich and P.Perona. Visual identification by signature tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):200–217, February 2003.

[13] E. Osuna, R. Freund, and F. Girosi. Support vector machines: training and applications. *AI Memo 1602, Massachusetts Institute of Technology*, 1997.

[14] R. Plamondon and G. Lorette. On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):63–84, January 2000.

[15] R. Plamondon and S. N. Srihari. On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):63–84, January 2000.

[16] R. Sabourin and R. Plamondon. Preprocessing of handwritten signatures from image gradient analysis. *Proceedings of the 8th International Conference on Pattern Recognition*, pages 576–579, 1986.

[17] B. Scholkopf, C. J. Burges, and A. J. Smola. Advances in kernel methods. *MIT Press*, 1998.

[18] B. Scholkopf, J. Platt, J. S. Taylor, A. J. Smola, and R. C. Williamson. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13:1443–1472, 2001.

[19] S. N. Srihari, S. Cha, H. Arora, and S. Lee. Individuality of handwriting. *Journal of Forensic Sciences*, 47(4):856–872, July 2002.

[20] G. Srikantan, S. Lam, and S. Srihari. Gradient-based contour encoding for character recognition. *Pattern Recognition*, 7:1147–1160, 1996.

[21] B. Zhang and S. N. Srihari. Analysis of handwriting individuality using handwritten words. *Proceedings of the Seventh International Conference on Document Analysis and Recognition, Edinburgh, Scotland (Los Alamitos, CA: IEEE Computer Society Press)*, 2003.

[22] B. Zhang and S. N. Srihari. Binary vector dissimilarity measures for handwriting identification. *Proceedings of SPIE, Document Recognition and Retrieval*, X:155–166, 2003.

0-7695-2187-8/04 $20.00 © 2004 **IEEE**