Dynamic Alignment Distance Based Online Signature Verification

Amaç Herdağdelen, Ethem Alpaydın

Department of Computer Engineering, Boğaziçi University
Istanbul, TR-34342, Turkey
amac@boun.edu.tr, alpaydin@boun.edu.tr

Abstract. A self-contained application that verifies the signatures of enrolled users is developed. The verifier uses dissimilarity values based on dynamic alignment distances. Three different threshold calculation techniques are investigated: The first one makes use of the information provided by forged signatures and the other two methods use the information gathered from genuine signatures. The methods are evaluated on two different datasets: One with skilled forgery and the other with random forgery. Best results are obtained by using both genuine and forgery information to calculate the threshold. The lowest reported total error value for the dataset with skilled forgeries is 7.83 percent.

1 Introduction

Automatic handwritten signature verification is an active research area in the field of pattern recognition, because it has a potential of use in many areas concerned with security and access control. Normally, signature verification is done by manual visual inspection. But in the majority of applications where a handwritten signature is used, no or little verification takes place at all [1]. Thus, automating this verification process in a reliable manner will assist to reduce fraud caused by forged signatures.

In this paper, an on-line signature verification method, in which a special hardware such as a digitizing tablet or a touch sensitive pad is used to obtain the movement of the pen, are investigated. The reliability of a signature verification technique can be measured by two factors that have varying relative weights depending on the area of application: False rejection rate (FRR) and false acceptance rate (FAR). FRR is also called type I error. It is the rate of falsely rejecting a genuine signature as forgery. FAR is also called type II error. It is the rate of falsely accepting a forged signature as genuine.

This paper is organized as follows: In Section 2, we discuss the methods for data acquisition, preprocessing and verifier. We give simulation results in Section 3 and conclude in Section 4.

2 Methodology

Our goal was to develop a self-contained system which consists of modules for enrollment of a user, preprocessing the signatures, and comparing (and verifying or rejecting) newly acquired signatures to the stored reference signatures.

2.1 Preprocessing and Feature Extraction

There is a considerable amount of irrelevant variance in the signature samples of the same individual. Five genuine samples of the same individual are given in Figure 1. In order to remove this variance, the raw signature samples were first preprocessed and

ISBN : 975-441-213-8

normalized using the following steps: Gaussian smoothing, size normalization and translation, rotation, and slant normalization. The effects of these operations can be seen in Figure 2. From this normalized representation, ten global features representing the global aspects of the signature are extracted. The features are: Sum of pressures (SoP), number of pen-ups (NoP), average velocity (AV), aspect ratio (AR), total displacement (TD), total signing time (TST), number of horizontal velocity sign changes (NoHVS), number of vertical velocity sign changes (NoVVS), number of horizontal acceleration sign changes (NoHAS), and number of vertical acceleration sign changes (NoVAS). A forward feature selection algorithm is employed to choose the feature subset that gives the best results in terms of total error rate.

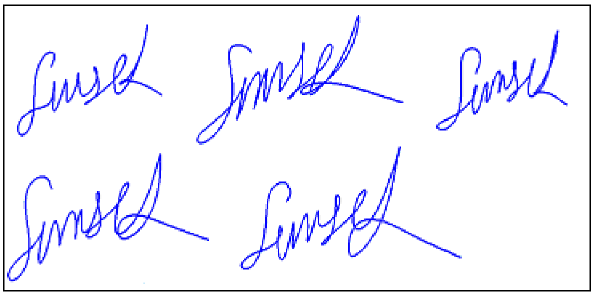
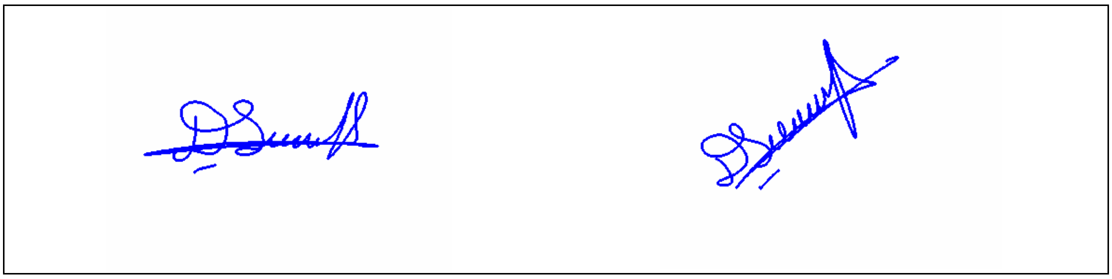
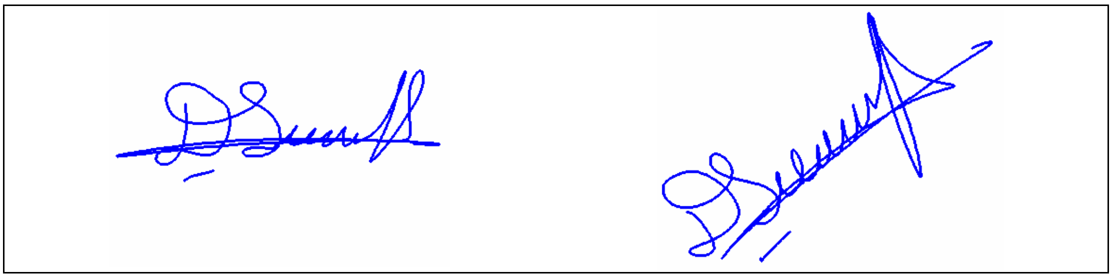


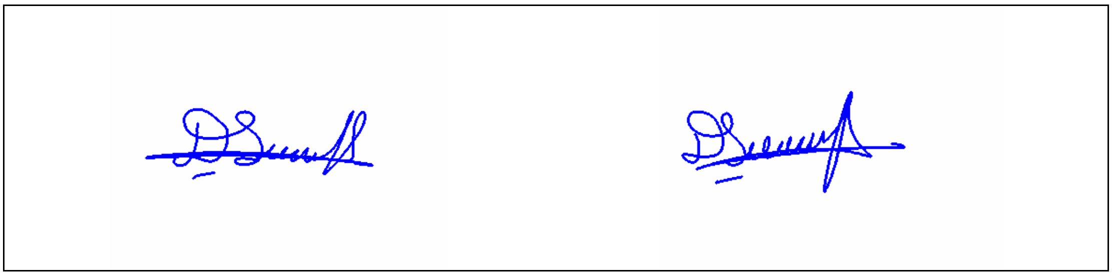
Figure 1. Original signatures of the same individual



a. Original signatures



b. Size normalized signatures



c. Rotated signatures

Figure 2. Effects of normalization techniques

ISBN : 975-441-213-8

2.2 Dynamic Alignment Distance Verifier

Dynamic alignment is a technique widely used to compare strings (vectors in this case) of different lengths. It finds an alignment between the strings such that the sum of differences between each aligned sample points is minimal.

An alignment is a set of pairings between the points of a reference signature *RS* and an unlabeled signature *US*. Alignment of the points should follow the temporal order. Thus, assume a point (x_m, y_m) of the signature *RS* is aligned with a point (x_k, y_k) of signature *US*, then any alignment of points (x_n, y_n) of *S* and (x_l, y_l) is invalid if one of the two following conditions holds:

$$n > m \text{ and } l < k$$
$$n < m \text{ and } l > k$$

Since the two signatures are of different lengths, some points can be skipped in a signature. But having this ability will lead to an arbitrary number of missing points: An empty set of pairings would always minimize the distance with a value of 0. In order to avoid this situation, skipping a point in any of the two signatures introduces a penalty in the overall distance. A missing point in the reference signature is called a missing point and a missing point in the unlabeled signature is called a spurious point. The penalty added to the overall distance is called the missing penalty (MP) and the penalty added for the spurious points is called the spurious penalty (SP). DADV is evaluated with different values for both SP and MP.

The dynamic alignment distance between two signatures (*RS* and *US*) is calculated dynamically as follows:

$$D(i,j) = \min \begin{cases} D(i-1, j-1) + SP + MP & \\ D(i-1, j) + MP & 1 \le i \le T_{RS} \\ D(i, j-1) + SP & 1 \le j \le T_{US} \\ D(i-1, j-1) + e & \end{cases}$$

where *e* is defined as the Euclidean distance between the i^{th} point of *RS* and j^{th} point of *US*:

$$e = \sqrt{(x_i^{RS} - x_j^{US})^2 + (y_i^{RS} - y_j^{US})^2}$$

Finally $D(T_{RS}, T_{US})$ gives us the dynamic alignment distance between *RS* and *US*.

Dynamic alignment dissimilarity value is calculated as follows. The minimum distance between *US* and the all reference signatures is the dissimilarity value of *US*. A reference signature is not compared to itself, thus, its dissimilarity value is the minimum distance between itself and the other reference signatures. We investigated three different techniques to calculate the dissimilarity value that differ in the lengths of the normalized signatures.

Two different approaches for threshold calculation are investigated. In the first one, two Gaussian distributions are estimated based on the information gathered during training and an individual threshold is calculated based on these distributions. In this approach the information provided by the forged signatures are taken into consideration. The training phase for individual threshold selection case is the same with the EVD. For a specific writer, all of the signature samples are processed and dissimilarity measures of all them are calculated. The reference and the genuine training sets are used for estimating the parameters (m_g^w, s_g^w) of the genuine Gaussian model, and the forged training set is used for estimating the parameters (m_f^w, s_f^w) of the forgery Gaussian model. Thus, a dynamic alignment dissimilarity value DA_g (of writer *w*) of a genuine signature is assumed to be:

$$DA_g^w \sim N(m_g^w, s_g^w)$$

while the dissimilarity value DA_f (of writer w) of a forged signature is assumed to be:

$$DA_f^w \sim N(m_f^w, s_f^w)$$

The individual threshold is the value that makes the probability distribution functions of those two distributions equal.

The other approach for calculating a threshold value is based on the fact that in real life, skilled forgeries of a new writer will not be readily available. Because of this constraint, one can choose to calculate a threshold value based solely on the genuine signatures. We seek alternative methods that use information gathered from only genuine signatures. In the first one, the minimum of the dissimilarity values of the reference signatures is chosen as the individual threshold. In the second method, the maximum of the dissimilarity values of the reference signatures is chosen as the threshold value. These two methods are evaluated with several MP and SP values as the individual threshold calculation is done.

Also a final parameter called threshold ratio is used to adjust the sensitivity of verifier in the last two threshold calculation configurations. An effective threshold is calculated by multiplying the threshold ratio with the calculated threshold.

Whenever an unlabeled signature, *US,* is presented, DADV calculates the dissimilarity value of *US*, and it compares the dissimilarity value to the calculated effective threshold value. If the dissimilarity value is higher then the threshold, *US* is rejected, otherwise it is accepted.

2.3 Evaluation

During the evaluation phase, the model is evaluated with the different configurations mentioned in Section 2.2. The model is run on over all the data and overall number of false rejections and false acceptations are used in calculating overall random-FRR, random-FAR, skilled-FRR and skilled-FAR, as it will be described in Section 3. We define the error rate for a specific configuration as the sum of FRR and FAR for that configuration. As will be seen in Section 3, different choices of different sensitivities allow constructing informative error trade-off graphs.

3 RESULTS

3.1 Dataset

In this study, two datasets, one of which is released for Signature Verification Contest (SVC2004 organization) are used [8]. Both the original dataset released for SVC2004 and the other modified dataset consist of 40 sets. In both of the datasets, each set contains 20 genuine signatures from one signature contributor and 20 forged signatures from five other contributors. In the original dataset, the forgeries are skilled whereas in the modified dataset, they are simply random forgeries which are genuine signatures of other contributors used as forgeries. The signatures are mostly in English or Chinese. Throughout this paper, the dataset with the skilled forgeries will be called skilled dataset and the dataset with random forgeries will be called random dataset.

The division method of a dataset into two sets as reference and training signatures is as follows: Six randomly selected signatures from the 20 genuine signatures are labeled as reference signatures. From the remaining 14 genuine signatures, eight of them are labeled as training and six of them are labeled as test signatures randomly. The 20 forged signatures are divided into two sets, each of 10 signatures, as training and test sets. A signature that is to be verified is denoted by *US* (unlabeled signature).

3.2 Dynamic Alignment Distance Verifier Results

Our aim is to obtain the error trade-off graphs of dynamic alignment distance verifier with different threshold selection mechanisms. Adjusting the sensitivity of verifier is done by choosing different values for threshold ratio that was defined in Section 2.2.

In order to evaluate the verifier, three parameters need to be set: Spurious penalty (SP), missing penalty (MP), and the vector length (VL) which determines whether the resampling procedure is applied to the signatures before calculating the dissimilarity values and if so with which parameter value. So before obtaining the trade-off graphs of the verifier, the configuration of those three parameters which leads to the best result in terms of total error needs to be determined.

In addition to the variety in the possible values for the parameters, another variable factor is the threshold calculation method. In Section 2.2, we define three methods: Statistical model estimation for the dynamic dissimilarity values, maximum dissimilarity method and the minimum dissimilarity method. While the first one makes use of the forgery information to calculate a better estimate for the threshold value, the latter ones only depend on the information provided by the genuine signatures. In order to evaluate those three methods, three sets of experiments are carried out, one for each of the threshold calculation methods. In each set of experiments, different configurations for MP, SP, and VL are investigated and the configuration that leads to the best result in term of total error is chosen as the "winner". SP and MP can take values of 300, 500, 800, 1200 and VL can take values of -1, 50, 150, where a value of -1 denotes that the lengths of signatures are left intact and they were not resampled.

After obtaining the winner configuration, the error trade-off graph is obtained for that configuration by varying the value of threshold ratio.

Dynamic alignment distance verifier is evaluated on the two datasets: Random dataset and skilled dataset. So the procedures described above are replicated two times (once for each dataset).

3.2.1. Results for Random Dataset

Best results on the random dataset with Gaussian model estimation for threshold calculation are obtained by the following configuration: VL = 50, MP = 800, and SP = 1200. The FAR and FRR pairs obtained with this configuration and varied threshold ratio values are tabulated Table 1 and visualized in Figure 3.

Table 1. FAR-FRR pairs on random dataset (Gaussian threshold)

Threshold Ratio	FRR (percent)	FAR (percent)
0.50	90.83	0.00
0.75	22.50	0.00
0.90	4.58	0.00
1.00	0.00	0.00
1.10	0.00	0.25
1.25	0.00	1.25
1.50	0.00	9.00
2.00	0.00	42.25
2.25	0.00	60.25
2.50	0.00	71.25
2.75	0.00	81.00

ISBN : 975-441-213-8

Best results on the random dataset with maximum dissimilarity method for threshold calculation are obtained by the following configuration: VL = 50, MP = 1200, and SP = 500. The FAR and FRR pairs obtained with this configuration and varied threshold ratio values are tabulated in Table 2 and visualized in Figure 3.

Table 2. FAR-FRR pairs on random dataset (Maximum dissimilarity method)

Threshold Ratio	FRR (percent)	FAR (percent)
0.75	8.33	0.00
0.90	0.42	0.25
1.00	0.42	1.25
1.10	0.00	2.50

Best results on the random dataset with minimum dissimilarity method for threshold calculation are obtained by the following configuration: VL = 50, MP = 1200, and SP = 300. The FAR and FRR pairs obtained with this configuration and varied threshold ratio values are tabulated in Table 3 and visualized in Figure 3.

Table 3. FAR-FRR pairs on random dataset (Minimum dissimilarity method)

Threshold Ratio	FRR (percent)	FAR (percent)
2.00	0.83	0.00
2.25	0.83	2.25
2.50	0.00	6.25

As it can be observed in Figure 3, the pairings of the FRR and the FAR error values mostly lie on the two axes. So it is not possible to observe the trade-off curves.

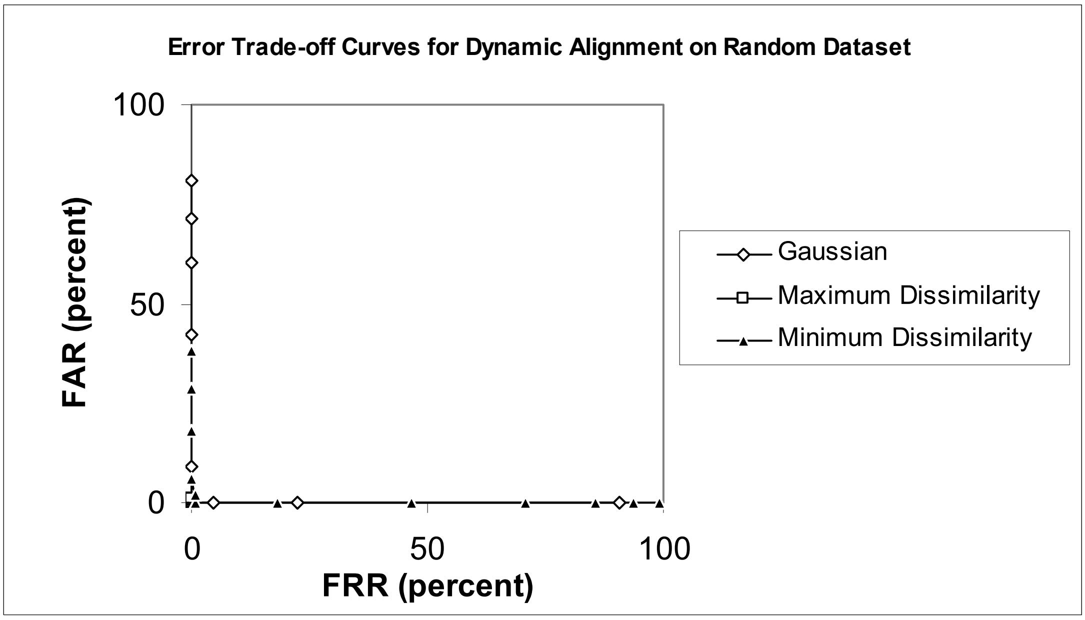


Figure 3. Error trade-off curves for dynamic alignment on random dataset

3.2.2. Results for Skilled Dataset

Best results on the random dataset with Gaussian model estimation for threshold calculation are obtained by the following configuration: VL = 50, MP = 1200 and, SP = 300. The FAR and FRR pairs obtained with this configuration and varied threshold ratio values are tabulated in Table 4 and visualized in Figure 4.

Table 4. FAR-FRR pairs on skilled dataset (Gaussian threshold)

Threshold Ratio	*FRR (percent)*	*FAR (percent)*
0.75	32.92	0.00
0.90	5.83	2.00
1.00	5.00	8.00
1.10	2.08	16.50
1.25	0.83	31.25
1.50	0.42	55.25
2.00	0.00	82.25

Best results on the random dataset with maximum dissimilarity method for threshold calculation are obtained by the following configuration: VL = -1, MP = 800, and SP = 800. The FAR and FRR pairs obtained with this configuration and varied threshold ratio values are tabulated in Table 5 and visualized in Figure 4.

Table 5. FAR-FRR pairs on skilled dataset (Maximum dissimilarity method)

Threshold Ratio	FRR	FAR
0.50	60.83	0.00
0.60	26.67	2.25
0.75	4.58	9.50
0.90	1.25	24.00
1.00	0.83	32.75
1.10	0.42	42.50
1.25	0.42	56.75
1.50	0.42	75.00
2.00	0.00	87.00

Best results on the random dataset with minimum dissimilarity method for threshold calculation are obtained by the following configuration: VL = -1, MP = 300, and SP = 500). The FAR and FRR pairs obtained with this configuration and varied threshold ratio values are tabulated in Table 6 and visualized in Figure 4.

Table 6. FAR-FRR pairs on skilled dataset (Minimum dissimilarity method)

Threshold Ratio	FRR (percent)	FAR (percent)
0.75	96.25	0.00
0.90	80.42	0.00
1.00	65.83	0.00
1.10	44.58	0.25
1.25	25.00	2.25
1.50	8.33	8.25
2.00	0.83	33.25
2.25	0.42	46.00
2.50	0.42	55.00
3.00	0.00	75.25
3.25	0.00	80.00

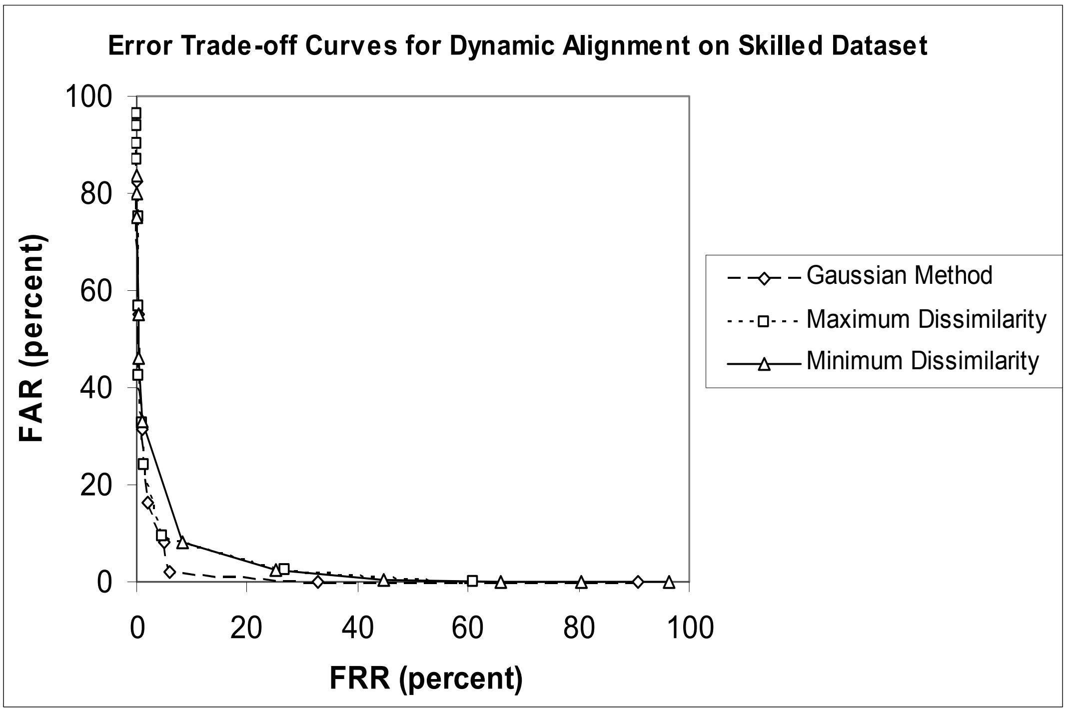


Figure 4. Error trade-off curves on skilled dataset

4 Conclusion

We evaluated the dynamic alignment verifier model on two datasets. The best result for dynamic alignment distance verifier on the random dataset was a total error value of 0 percent. Calculating the threshold by estimating two Gaussian models for the genuine and forged signatures led to this impressive result. But it should be noted that the other two threshold calculation methods, maximum dissimilarity value and minimum dissimilarity value method, led to total error values of 0.67 percent and 0.83 percent respectively. Thus, the lack of information provided by the forged data does not seem to be vital for obtaining a near-0 error performance on random dataset.

Dynamic alignment distance verifier gave higher error rates for the skilled dataset: 7.83 percent for Gaussian threshold calculation, 14.08 percent for maximum dissimilarity method, and 16.58 percent for minimum dissimilarity method. In the skilled dataset, the advantage of having skilled forgery during training phase becomes obvious.

In this study, we aligned only the spatial coordinates of the two signatures in dynamic alignment verifier. Instead, one can extract more information like pressure, speed, acceleration, horizontal speed, and vertical speed for each sample point and align the points in this multi-dimensional space. This method would allow the verifier to make use of the correlations between the various characteristics of the signature.

Also the performance of the verifier on skilled forgeries, when it is trained on random forgeries, is not evaluated. Since this case will be likely the one that will be encountered in real life applications, future study on this issue is needed.

Acknowledgements

This work was supported by Boğaziçi University research grant 03K120250.

References

1. Griess, F. D., "Online Signature Verification", Project Report, http://www.cse.msu.edu/publications/tech/TR/MSU-CSE-00-15.ps, Michigan State University, May 2000.
2. Gupta, G., A. McCabe, "A Review of Dynamic Handwritten Signature Verification", http://www.cs.jcu.edu.au/~alan/Work/HSV-Lit_rev.html, James Cook University, September 1997.
3. Leclerc, F. and R. Plamondon, "Automatic Signature Verification: The State of the Art -1989 – 1993", *Proceedings of the International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 8, No. 3, pp. 643–660, 1993.
4. Wessels, T., C. W. Omlin, "A Hybrid System for Signature Verification", *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Vol. 5, pp. 509-514, 2000.
5. Kashi, R. S., J. Hu, W. L. Nelson, W. Turin, "On-Line Handwritten Signature Verification Using Hidden Markov Model Features", *Proceedings of International Conference Document Analysis and Recognition (ICDAR)*, Ulm-GERMANY, August 18-20 1997, Vol. 1, pp. 253-257, 1997.
6. Jain, Anil K., F. D. Griess, and S. D. Connell, "On-Line Signature Verification", preprint submitted to Elsevier Science, 2000
7. Nelson W., W. Turin, and T. Hastie, "Statistical Methods For On-Line Signature Verification", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 8, No. 3, pp. 749–770, 1994.
8. The First International Signature Verification Competition (SVC 2004), http://www.cs.ust.hk/svc2004/.
9. Vuori, V. J. Laaksonen, E. Oja, and J. Kangas. "Experiments with Adaptation Strategies for a Prototype-based Recognition System for Isolated Handwritten Characters", *International Journal on Document Analysis and Recognition*, Vol. 3, No. 3, pp. 150-159, 2001.
10. Alimoğlu, F., "Combining Multiple Classifiers for Pen-based Handwritten Digit Recognition", M.Sc. Thesis, Boğaziçi University, 1996.