Locally Weighted Enhanced DTW Based Online Signature Verification

Vishwaas Pratap Singh, Prabhnoor Singh, Yash Kamlaskar & Ramesh K Bhukya, *IEEE Member*
Department of Electronics and Communication Engineering
Indian Institute of Information Technology Allahabad, Prayagraj
Uttar Pradesh - 211012, India
{iec2019019, iec2019061, iec2019020115, rkbhukya}@iiita.ac.in

***Abstract*—In this paper, we present enhance *Dynamic Time Warping* (DTW) based method on the online *Signature Verification* (SV) by employing the code-vectors generated from *Vector Quantization* (VQ) to match the aligned pairs in the warping path. The DTW based method is used to compute the distance score between the test signature and the genuine enrolled signatures, for the decision making. In order to improve the results, we conducted the evaluation using MCYT-100 database and obtained an *Equal Error Rate* (EER) of 1.55% and SVC-2004 database provides the EER of 2.73%. In this work, we exploit the characteristics of the warping path for online SV and obtained enhanced efficacy of the system.**

***Index Terms*—Signature Verification, Dynamic Time Warping, Vector Quantization, Bio-metrics**

I. Introduction

In the rising era of authentication using biometrics, verification through signature requires less equipment for data acquisition and is one of the most widely used method. Banks, intelligence agencies, and other high-profile organisations utilise *signature verification* (SV) to authenticate an individual's identification. As a behavior feature, handwritten signatures are widely used in financial and administrative institutions. There are distance and model-based techniques in terms of classifier structures. The *Dynamic Time Warping* (DTW) technique is used to compare the test signature to a set of reference signatures in this document. Following that, a set of three distances between the test and reference signatures is calculated and normalised, yielding a three-dimensional feature vector. After that, the feature vector is loaded into an *Support Vector Machine* (SVM) that has been trained on both real and fake signatures. The output of the SVM is then utilised to determine whether or not the signature is genuine.

To make it easier to compute the similarity of the test signature with the enrolled reference signatures, a set of length normalisation methods based on three re-sampling approaches (namely spatial based, temporal based, and mean based) are used. Even though we are using SVC2004 database but we are comparing the results with MCYT-100 database for the enhanced DTW scheme A method for matching signatures is given based on the concept of the Longest Common Sub-Sequence, in which a sequence of turning angle-based features is derived at each sample point of the online trace.

The main objective here is to classify the signature as forged or genuine, most of the existing methods use model based techniques. Model based approaches describe the distribution of the data by the use of generative-based classifiers like *Hidden Markov Models* (HMM) and *Gaussian Mixture Models* (GMM) or discriminative ones such as *Multi Layer Perceptrons* (MLP). For low resolution data or in case of offline SV this model based approaches do quite well, but when we deal with high resolution data, and can iterate through a image over its finest details then distance based model work more effectively. The distance scores are used in developing of the decision rule in classic DTW algorithms for online signing. The generated scores correspond to the cumulative distance values/distortions along the warping path. The warping path is created by constraining the alignments between the sample points of the two signatures being compared. However, relying solely on the DTW score may not always be enough to distinguish fake signatures from authentic ones. When the signature patterns of genuine and fake have values that are quite similar to each other, this is very likely.The traditional DTW algorithms however, do not consider this aspect in the scoring of the test signature. This is the research gap from literature that we address in our proposal.

We use DTW to calculate a score that measures the proportion of aligned pairings that produce a low distortion value in relation to the length of the warping path. We explore using a code-book of adequate size, built from a Vector Quantization (VQ) model, to compute this proportion. We then combine this score with the DTW score (using popular score combination methodologies) to verify the validity of a test signature. The topic of signature recognition has been the focus of VQ application in the literature. The quantized feature vector of a claimed user's test signature is compared to those of the enrolled reference signatures. The motivation for quantizing the feature vector was to improve the biometric system's privacy by making it impossible for hackers to duplicate biometric data from the extracted characteristics communicated.

The remaining paper is organized as follows: Section II describes the methodology used for the paper. Section III describes the experimental results and discussion done on the MCYT-100 database. Finally, the conclusion and scope of future work is in Section IV.

II. Methodology

Initially we acquired the online MCYT-100 SV database. During the training phase, we applied the feature extraction from the given database to process the training vectors. The created training vectors are used to build the model by using VQ model with the help of code-books. During testing phase, the feature extractor module is used to process the test signatures through the same process as did during the training phase. The enrolled training feature vectors are stored in the system. Once the test signatures claims against the enrolled reference signatures, the DTW method is used to match the corresponding code vectors of the test and enrolled reference signatures based on the warping path of the DTW distance score. Finally, the decision making is based on the threshold (threshold value is pre-decided based on the offline experiments) and to verify the signature's validity.

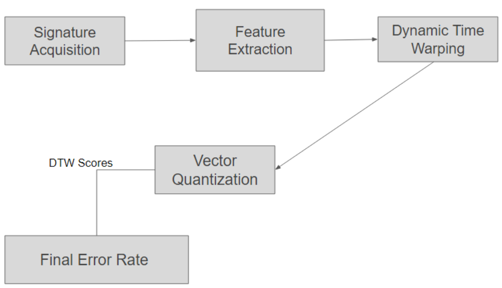

Fig. 1. Generic block diagram of the online signature verification system.

A. Feature Extraction Phase

During pre-processing, all the strokes are combined into one long stroke. The number of features are extracted at each sample point of the signatures. The extracted features are include the change in x- and y- co-ordinates, provides two variables features. The variable features are spatial (extracted from signature's shape) and temporal (speed of the writing). In this work, we consider the changes in x- and y- co-ordinates and the pressure between consecutive places as well characteristics based on angle.

$$\Delta x_i = x_{i+1} - x_i$$

$$\Delta y_i = y_{i+1} - y_i$$

$$\Delta p_i = p_{i+1} - p_i$$

The 5 features namely the x- coordinate, the y- coordinate, pressure, azimuth and elevation of different points in a given signature are retrieved from the MCYT database provided as output features. The Figure 2 gives the illustration about the changes in the signature during the enrollment phase and their respective variations of the features of the given signature can be seen from the Figure 3.

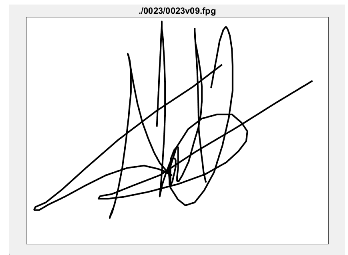

Fig. 2. The signature considered from MCYT database for the illustration

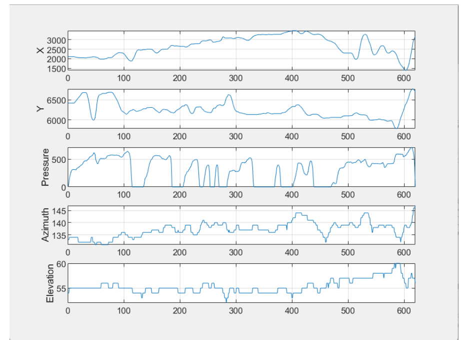

Fig. 3. The features are extracted from the MCYT-100 database using their respective pre-decided features

B. DTW Based Matching

In this work, we consider two signatures for matching, T (test signature) and S_ρ (reference signature),each represented by a feature vector sequence of length $n_T - 1$ and $n_\rho - 1$ respectively. The measure of dissimilarity $d(l, m)$ between the l^{th} point of T and the m^{th} point of S_ρ is represented by the $(l, m)^{th}$ member of the $(n_T - 1) * (n_\rho - 1)$ cost matrix (denoted by C).The optimum warping path (W_ρ) in C is determined by a contiguous collection of matrix components that define a mapping between T and S_ρ. To find the DTW distance between T and S_ρ are as follows:

$$\psi(l, m) = d(l, m) + min(\psi(l-1, m), \psi(l-1, m-1), \psi(l, m-1))$$

The DTW score on the length of each warping path is normalized using DTW score as :

$$d_1^p = \frac{\psi(n_T - 1, n_\rho - 1)}{l_{W_\rho}}$$

Analysis of the distortions d along the warping path reveals particular cells with low values, corresponding to areas of the trace between the signatures being matched, that an impostor is likely to have trouble forging. High values, on the other hand, refer to regions that differ between the signatures being compared.

C. Vector Quantization

VQ is an efficient and effective coding technique for quantizing signature vectors. The primary step in VQ is the representation of code vectors. We create a model for each user based on VQ by concatenating all of the feature vector sequences corresponding to the N genuine signatures registered in the system to create a lengthy training sequence. The final set of code-book of size 2 for the two-level quantizer is achieved after convergence. These two code-vectors are then split again, yielding four seed vectors for a four-level quantizer. The process of updating the code-vectors is repeated until a quantizer with a code book of size M is obtained. The final code-vectors of user u are denoted by $c_u^1, c_u^2, c_u^3, ..., c_u^M$.

D. Enhanced DTW based system

We assign the appropriate feature vectors $f_T^{\psi_i}$ (from signature T) and $f_\rho^{\phi_i}$ (from signature S_ρ) to their nearest code vector acquired from the code book generated using the VQ procedure for each aligned pair (ψ^i, ϕ^i) along the warping path W_ρ. If the indices of assigned code-vectors are $l_T^{\psi_i}$ and $l_\rho^{\phi_i}$, we can write:

$$l_T^{\psi_i} = arg \min_{1 \leq k \leq M} |f_T^{\psi_i} - c_u^k|$$

$$l_p^{\phi_i} = arg \min_{1 \leq k \leq M} |f_\rho^{\phi_i} - c_u^k|$$

We issue a vote of 1 to each cell (ψ_i, ϕ_i) based on the code-vector indices $l_T^{\psi_i}$ and $f_\rho^{\phi_i}$ after they assign feature vectors $f_T^{\psi_i}$ and $f_\rho^{\phi_i}$ to the code-vectors. To this end, we create the following indicator variable:

$$I(\psi_i, \phi_i) = \begin{cases} 1 & for \quad l_T^{\psi_i} \neq l_\rho^{\phi_i} \\ 0 & for \quad otherwise \end{cases}$$

The total number of votes received by cells along the warping path can now be normalized with the length as,

$$d_2^\rho = \frac{\sum_{i=1}^{l_{W\rho}} I(\psi_i, \phi_i)}{l_{W_\rho}}$$

The average degree of mismatch that occurs between a pair of aligned locations (ψ_i, ϕ_i) along W with the VQ code-book of size M can be read as this metric.

Consider comparing a test genuine signature to one from the enrolled group. For a large number of cells along the warping path W, the distortion $d(\psi_i, \phi_i)$ between feature vectors $f_T^{\psi_i}$ and $f_\rho^{\phi_i}$ is low due to the high degree of similarity between parts of the trace. With an adequate code-vector of size M, the feature vectors are likely to fall into the same cluster. The percentage of cells d_2^ρ along W whose characteristics are allocated to distinct code-vector indices can then be used to determine how divergent the two signatures' traces are. This number will be low if the signature matches are genuine to genuine.

For verification, the resulting feature d_2 is coupled with the normalised DTW distance d_1. We acquire N_f values of d_2 and d_2 from N_f enrolled signatures of a user u. The distances are averaged as follows:

$$d_1^{mean} = \frac{1}{N_f} \sum_{i=1}^{N_f} d_1^i$$

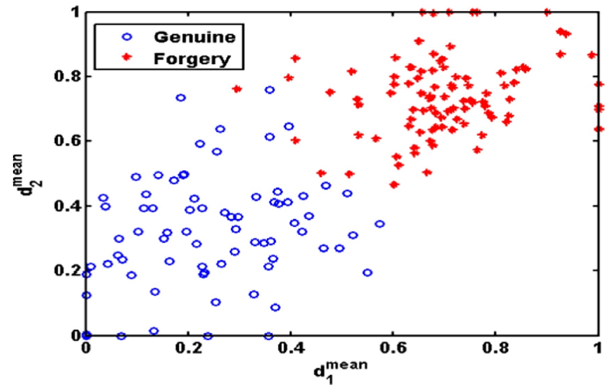

Fig. 4. Two dimensional distance plot of a user's genuine and skilled forgery signatures.

$$d_2^{mean} = \frac{1}{N_f} \sum_{i=1}^{N_f} d_2^i$$

We depict the trend of values of d_1 and d_2 for genuine and skilled forgery signatures of a user as a 2-D plot in order to assess the trend of values of d_1 and d_2 for genuine and skilled forgery signatures of a user (Fig. 4). The signatures of a user from the SVC-2004 database were used in this case. The scores are complimentary, and using only d_1 and d_2 results in a substantial overlap in the distance values of genuine and forged signatures. As a result, we combine them with common score combination strategies. We look at four of these methods in this paper: the sum rule, weighted sum rule, weighted product, and weighted minimum strategies. Table 1 shows a list of them. If the fused value d_T is less than the defined threshold, the verification system accepts the sample signature T .

TABLE I
VARIOUS SCORE COMBINATION TECHNIQUES FOR DETERMINING DISTANCE SCORE

Sum Rule	$d_1^{mean} + d_2^{mean}$
Weighted Sum	$(\alpha - 1)d_1^{mean} + (2 - \alpha)d_2^{mean}$
Weighted Minimum	$min(\alpha d_1^{mean}, d_2^{mean})$
Weighted Product	$d_1^{mean} * d_2^{mean(\alpha-1)}$

III. Experimental Results and Discussion

A. Experimental Set-up

The SV experiments are carried out using the SVC 2004 and MCYT-100 databases, both of which are freely available. The following is a brief overview of each of them: The signature data for the SVC 2004 database is taken from the WACOM Intuos tablet, containing 20 authentic signatures and 20 professional forgeries for each participant. In total, 40 people contributed to the data over the course of two sessions separated by a week. The spatial co-ordinates, pressure, azimuth angle, inclination angle, time stamp, and pen up/down status are all captured.

The MCYT-100 is a larger data set that includes signature samples from 100 people. The WACOM pen tablet, model-Intuos A6 USB, was used to collect 25 genuine and 25 fake signatures for each participant. The information in each trace of the signature is identical to that in the SVC 2004 database. Five distinct users with access to the static photos of the genuine signature contributed the skilled forgeries.

For most experiments on the SVC 2004 database, we randomly select a set of five genuine signature of an user for enrollment. The chosen signatures are from the first session as described. The remaining 15 genuine signature and 20 skilled forgeries are used to evaluate the efficacy of our proposal. This process is repeated over a set of five repetitions. A similar protocol is followed for the MCYT-100 database as well. Here, we consider five genuine signatures per user for enrollment. The remaining 20 genuine signatures and 25 skilled forgeries are used for testing the efficacy of our proposal.

B. Results on MCYT-100 database

On the signatures from the bigger MCYT-100 database, we now test the effectiveness of the suggested methodology. Table 2 shows the EERs derived utilising the Table 1 combination procedures, along with their accompanying parameters. For the parameters (2, 32, 13), the Weighted Product rule has the lowest EER of 1.55 percent, which is a 53.32 percent improvement over the baseline DTW system's EER of 3.32 percent. Across all of the schemes, there is a similar level of performance.

TABLE II
EERs DEVISED USING TECHNIQUES MENTIONED IN TABLE 1

Scheme	EER in %	(α,M,W)
Baseline DTW	3.32	(2,-,-)
Sum Rule	1.81	(1.5,32,13)
Weighted Sum	1.72	(1.2,32,13)
Weighted minimum	2.07	(2,32,13)
Weighted product	1.55	(2,32,13)

The MCYT-100 database's attributes account for its substantially superior verification performance (as compared to the SVC-2004 database). Unlike the SVC-2004 database, where people were invited to 'create' signatures, this database was compiled using real signatures of people. As a result, there are less intra-class variances among a user's legitimate signatures. As a result, the cost matrix and associated warping routes have a comparable structure, allowing the combination of the DTW score d_1^{mean} with the scores d_2^{mean} of the aligned pairings by the code-book vectors to distinguish genuine from sophisticated forged signatures.

We tabulate the EERs for varied numbers of reference signature samples enrolled in the system in Table 3 to further highlight the efficacy of the proposal. We can see that even with a small number of samples, a user's signature can be adequately validated.

TABLE III
EERs VARYING ON NUMBER OF SIGNATURES

# of reference signatures	Baseline system	Proposed system
1	7.81	4.32
2	5.65	3.97
5	3.32	1.81
10	2.82	1.6

C. Discussion

Table 4 compiles a list of related works on online SV. It's worth noting that each of these systems uses distinct experimental setups and features, as well as various classifiers, making a straight one-to-one comparison of the results. The systems that use the DTW algorithm are marked with a 'Yes' in the third column. Our concept, in particular, yields promising outcomes in relation to these systems. Using Locally Weighted Enhanced DTW we observed that our results slightly improves as compared to the existing publications.

TABLE IV
PROPOSED METHOD'S COMPARISON WITH EXISTING METHODS

Method	EER %	DTW Based
Kholmatov and Yanikoglu	6.96	Yes
Fierrez-Aguilar et al.	10.91	Yes
Faundez-Zanuy	5.42(MCYT)	Yes
Muramatsu and Matsumoto	10.15	Yes
Pascual-Gaspar et al.	3.38	Yes
Yanikoglu and Kholmatov	7.22(MCYT)	Yes
Barkoula et al.	6.23	Yes
Fierrez et al.	6.90	-
Van et al.	4.83	-
Guru and Prakash	6.12(MCYT)	-
Ibrahim et. al	1.09(MCYT)	-
Gruber et al.	6.84	-
Enhanced DTW	**2.73** **1.55**(MCYT)	Yes

IV. CONCLUSION AND FUTURE WORK

This is the first attempt to examine the distortion values of the cost matrix's warping path.We investigated following fields:

- A score to describe the trend of distortion values along the warping paths of genuine and forgery signatures.
- For decision-making, combining the calculated score with the normalised DTW score.
- Incorporating vector quantization to tackle distortions formed in the base DTW method.

For further improvements, we can explore DTW variations employed in various study disciplines. As an alternative to VQ, we will try to implement other machine learning and deep learning techniques. The current approach places little emphasis on feature extraction, this gap can be filled by investigating the DTW-VQ strategy's with combination of several discriminative local features.

References

[1] K. Barkoula, G. Economou, S. Fotopoulos Online signature verification based on signatures turning angle representation using longest common subsequence matching Int. J. Doc. Anal. Recognit. (IJDAR), 16 (3) (2013), pp. 261-272

[2] M. Faundez-Zanuy On-line signature recognition based on VQ-DTW Pattern Recognit., 40 (3) (2007), pp. 981-992

[3] J. Fierrez-Aguilar, S. Krawczyk, J. Ortega-Garcia, A.K. Jain Fusion of local and regional approaches for on-line signature verification Proceedings of the 2005 International Conference on Advances in Biometric Person Authentication (2005), pp. 188-196

[4] J. Fierrez, J. Ortega-Garcia, D. Ramos, J. Gonzalez-Rodriguez HMM-based on-line signature verification: Feature extraction and signature modeling Pattern Recognit. Lett., 28 (16) (2007), pp. 2325-2334

[5] C. Gruber, T. Gruber, S. Krinninger, B. Sick Online signature verification with Support Vector Machines based on LCSS kernel functions Syst. Man Cybern. Part B Cybern. IEEE Trans., 40 (4) (2010), pp. 1088-1100

[6] D. Guru, H. Prakash Online signature verification and recognition: An approach based on symbolic representation Pattern Anal. Mach. Intell. IEEE Trans., 31 (6) (2009), pp. 1059-1073

[7] M.T. Ibrahim, M.A. Khan, K.S. Alimgeer, M.K. Khan, I.A. Taj, Guan L. Velocity and pressure-based partitions of horizontal and vertical trajectories for on-line signature verification Pattern Recognit., 43 (8) (2010), pp. 2817-2832

[8] A. Kholmatov, B. Yanikoglu Identity authentication using improved online signature verification method Pattern Recognit. Lett., 26 (15) (2005), pp. 2400-2408

[9] D. Muramatsu, T. Matsumoto Effectiveness of pen pressure, azimuth, and altitude features for online signature verification Proceedings of the International Conference on Advances in Biometrics, Springer-Verlag (2007), pp. 503-512

[10] J. Pascual-Gaspar, V. Cardeoso-Payo, C. Vivaracho-Pascual Practical on-line signature verification M. Tistarelli, M. Nixon (Eds.), Advances in Biometrics, Lecture Notes in Computer Science, vol. 5558, Springer Berlin Heidelberg (2009), pp. 1180-1189

[11] R. Szeliski Computer Vision: Algorithms and Applications (first), Springer-Verlag New York, Inc., New York, NY, USA (2010)

[12] B. Yanikoglu, A. Kholmatov Online signature verification using Fourier descriptors EURASIP J. Adv. Signal Process, 2009 (2009), pp. 12:1-12:1

[13] Sharma, Abhishek, and Suresh Sundaram. "An enhanced contextual DTW based system for online signature verification using vector quantization." Pattern Recognition Letters 84 (2016): 22-28.

[14] B. Yanikoglu, A. Kholmatov Online signature verification using Fourier descriptors EURASIP J. Adv. Signal Process, 2009 (2009), pp. 12:1-12:1

[15] C. Vivaracho-Pascual, M. Faundez-Zanuy, J.M. Pascual An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances Pattern Recognit., 42 (1) (2009), pp. 183-193