Off-Line Signature Verification based on Ordered Grid Features: An Evaluation

Konstantina Barkoula, George Economou
Physics Department
University of Patras
Patras, Greece
email: kbarkoula@gmail.com, economou@upatras.gr

Elias N. Zois, Evangelos Zervas
Electronics Engineering Department
Technological and Educational Institution of Athens
Egaleo, Greece
e-mail: {ezois, ezervas}@teiath.gr

***Abstract*— A novel offline signature modeling is introduced and evaluated which attempts to advance a grid based feature extraction method uniting it with the use of an ordered powerset. Specifically, this work represents the pixel distribution of the signature trace by modeling specific predetermined paths having Chebyshev distance of two, as being members of alphabet subsets-events. In addition, it is proposed here that these events, partitioned in groups, are further explored and processed within an ordered set context. As a proof of concept, this study progresses by counting the events' first order appearance (in respect to inclusion) at a specific powerset, along with their corresponding distribution. These are considered to be the features which will be employed in a signature verification problem. The verification strategy relies on a support vector machine based classifier and the equal error rate figure. Using the new scheme verification results were derived for both the GPDS300 and a proprietary data set, while the proposed technique proved quite efficient in the handling of skilled forgeries as well.**

Grid Features, Power Set, Ordering, Signature Verification

I. INTRODUCTION

Automated handwritten signature verification systems (ASVS) remain up to now an accepted way for humans to declare their identity in many application areas including civilian ones [1], [2], [3], [4]. ASVS are separated into two major categories based on the method that the signature is obtained. Both online and offline ASVS must cope with the evidence that the process of creating handwritten signatures, even when they originate from a well trained genuine writer, will carry natural variations, defined as intra-writer variability [5]. It is adopted that the online ASVS are generally more efficient when compared to offline. A commonly used figure of merit which is employed in order to characterize the efficiency of ASVS is the equal error rate (EER) which is calculated from the ROC or DET plots of both types of error rates.

The goal of an offline ASVS is to efficiently transform an image into a mathematical measurable space where it will be represented by means of its corresponding features [6]. Next, the features are feeding computational intelligence techniques and pattern recognition classifiers which will decide, after appropriate training and testing procedures, if a signature under query belongs to a claimed writer [7], [8]. According to the experimental protocol followed, there are two major approaches which have been applied to off-line ASVS; writer dependent (WD) and writer-independent (WI). The WD approach uses an atomic classifier for each writer. The WI approach uses a classifier to match each input questioned signature to one or more reference signatures, and a single classifier is trained for all writers [9], [10].

Feature extraction is considered to be one of the most challenging tasks when ASVS are designed. An important feature extraction philosophy which attracts increasing interest, exploits the signature using a coarse or fine detail grid which is imposed upon the image. Among others, examples of grid based feature extraction can be found in the work provided by references [10], [11], [12], [13], [14], [15], [16], [17], [18] and [19].

In another work provided by Tselios, Zois, Nassiopoulos and Economou [20], a grid based feature extraction method was developed which represents the signature trace by taking into account the histogram of specific pixel path transitions along predefined paths within pre-confined Chebyshev distances of two (F_{CB2} feature). The feature extraction concepts have been advanced by describing these paths in a way in which they can be viewed as symbols transmitted by a discrete space random source. The combination of the produced F_{CB2} symbols defines the message or event that the random source sends out when a certain sequence of signature pixels is accounted. They are treated according to the event concept, reported in standard set and information theory and they are complemented along with their corresponding probabilistic moments [21]. In this work and in order to further increase our signature discriminating capability the potential messages-events of the F_{CB2} paths are organized in sub-groups of independent tetrads. Each tetrad is organized according to its ordered powerset with respect to inclusion [22]. The outcome of this procedure provides an attempt to model the handwriting process in concordance with basic elements of information and coding theory.

The distributions of the now ordered transition paths in the new feature space are used to code the signature image. In the case study presented here a WD verification scheme is followed which comprises of the training and testing phase. Verification results have been drawn with the use of two databases, the GPDS300 and a proprietary one by means of the false acceptance, false rejection and the equal error rate (EER) figure of merit. The rest of this work is organized as follows: Section 2 provides the database details and the description of the feature extraction algorithm. Section 3 presents the experimental verification protocol which has been applied. Section 4 presents the comparative evaluation results while section 5 draws the conclusions.

II. Database And Feature Extraction Procedure

A. Database Description

The proposed feature extraction modeling has been studied with the use of two databases of 8-bit grey scale signatures: a Greek signers' database (CORPUS1) [20] and GPDS-300 (CORPUS2) [12]. CORPUS1 comprises of a domestic Greek collection of 105 genuine and 21 simulated forgery signature samples for each of the 69 signers of the database. Genuine samples were acquired in a one month time frame. CORPUS2 contains 24 genuine signatures and 30 simulated forgeries for each of the 300 signers of the database and is publicly available. During the experimental process, two schemes of randomly selected training and testing samples were used for comparison with the outcomes of contemporary research in the field. In the first scheme, 12 genuine and 12 simulated-forgery reference samples per writer are used, while in the second scheme 5 genuine and 5 simulated forgery reference samples are used. The remaining samples are used for testing.

B. Preprocessing

In order to produce the binary form of the acquired signatures the following preprocessing steps have been carried out: thresholding using Otsu's method [6], skeletonization, cropping and segmentation. This procedure is expected to reduce a number of side effects of the writing instruments variations. The result is the generation of the most informative window (MIW) of the image. The features are extracted either from the whole MIW of the signature or from segments of signature's MIW with the use of the equimass sampling grid method [14]. Equimass sampling grid segmentation provides strips of the signature with uniform size of signature pixels instead of the trivial distance grid segmentation which provides segments of equal area. The result is depicted in Fig. 1. In this work the feature vector is generated from the 'S2' scheme used in [20].

C. Alphabet Description

Fig. 2 depicts the alphabet which is defined as a set of symbols, emerging from the F_{CB2} description according to [12]. To be more specific, F_{CB2} alphabet is the set of transition paths of three consecutive pixels under the constraint of having the first and third pixels restrained to a Chebyshev distance equal to two.

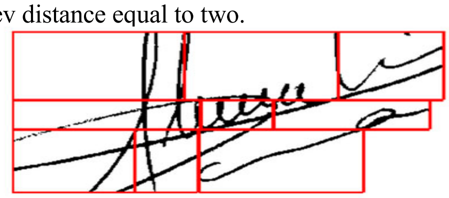

Figure 1. Signature image with equimass made segments

Since, in offline signatures, signature-pixel ordering is unknown, the ordered sequence of the pixels cannot be estimated. This note diminish the number of queried F_{CB2} transition paths, in a 5x5 pixel grid window, with center pixel each black pixel of signature's image, to the sixteen independent transition paths presented in Fig. 2. In this case study only the F_{CB2} paths have been taken into account. It is advantageous in our case to explicitly treat the notion of the signature pixels indexes (i,j) as a transformation of sequences produced by the source. As a consequence, the feature extraction grid can be identified as a discrete space – discrete alphabet source.

D. Ordered Event Modeling

Let the triad (Ω , B, P) indicate the probability space on which all the potential outcomes are identified. By definition Ω is the sample space upon which a discrete digital source transmits alphabet symbols. The source may transmit either single symbols or sets of them (events) from a 16 symbol alphabet as figure 2 illustrates. Let B a sigma field (the event space) that encloses all potential occurrences of symbols combinations from the F_{CB2} alphabet. That is, B is the largest possible σ -field [23] which is the collection of all subsets of Ω and is called the power set. Finally, let P be the corresponding distributions of the σ -field.

In order to evade the problem of 2^{16} space management Ω is grouped into T subsets $\{\Omega_t\}_{t=1,\dots,T}$ and we define the sub-s-fields B_t as the power sets for each Ω_t. In this work we choose to group the 16-F_{CB2}(i) components into ensembles of four tetrads (call it hereafter F_4-collection) thus resulting to an early set of 4×2^4=64 possible event combinations. From the complete set of all the possible ensembles of the F_4 collection only 87 orthogonal cases shall be enabled along with their corresponding probabilities. From a mathematical point of view the signature image is analyzed into four major subspaces where each of them is composed of 16 orthogonal dimensions. The term orthogonal denotes that each symbol in a sub-alphabet space of a F_4 tetrad cannot be derived as any combination of the same subspace F_4 symbols. This constraint provides each signature with 87 different F_4 orthogonal tetrad event sets, found through exhaustive search. Fig. 3 provides the F_{CB2} alphabet along with a F_4 orthogonal collection. As a proof of concept, the orthogonal F_4 collection #44, selected randomly is illustrated in figure 3.

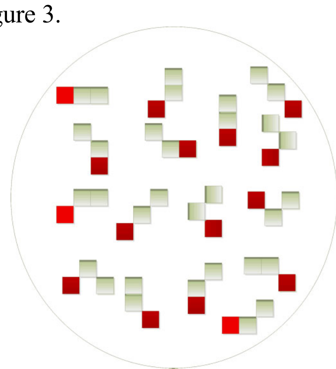

Figure 2. F_{CB2} alphabet set which forms the probability space Ω .

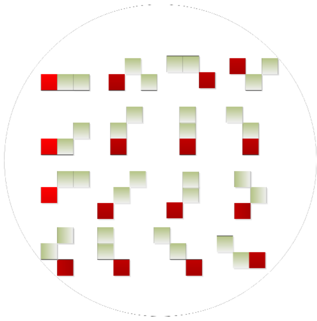

Figure 3. One F_4 collection of tetrads (#44). Each horizontal tetrad is considered to form a subspace in the original 16-dimnensional feature space and consequently generates a powerset of events

Finally, each one of the four F_4 power-sets of figure 3b is evaluated by ordering the elements of the powerset with respect to inclusion. Fig. 4 provides a graphical explanation of one powerset in line with the proposed modeling. In order to illustrate the method with clarity, figure 4 has been created which shows the powerset of the #44 F_4 collection with respect to inclusion. The indexes x, y, z, w are associated with one tetrad's elements of the F_4 collection. For each arrow in figure 4 there is a corresponding probability evaluated for every segmented image. Thus, the overall dimensionality of the feature vector for one F_4 collection is equal to 32 (4+12+12+4) for each image segment.

According to the exposed material, a discrete source, designated as S_n, can be defined by its transmitted set of symbols-events which are now members of an ordered F_4 collection. This novel modeling of the feature generation process is an evolution of the previous method as it was described in [20]. It attempts to model the distribution of the signature pixel paths as an information source and to associate events of ordered paths (arrows as seen in fig. 4) along with their corresponding first order probabilities.

E. Creation of the ordered feature vector

To make this work robust a short description is provided for generating the ordered feature components. According to the material exposed in sections II_C, II_D, each one of the preprocessed image segments is scanned top-down and left-right to identify its signature pixels. Let us denote with the labels One (O) and Two (T) a conjugated pair of 5×5 moving grids with the property that their topological centers are distant by a Euclidean distance of one. Then for each signature pixel the {O, T} grids are imposed. Next, detection of discrete events at both {O, T} grids is performed followed by the evaluation of the corresponding ordered probabilities, as described in fig. 4. In addition, fig. 5 presents in a graphical manner the generation of a feature component namely the {X, XY}. In this work the overall feature dimensionality is 128 due to the selection of the segmentation preprocessing steps.

III. CLASSIFICATION PROTOCOL

On the grounds of proofing the proposed concept and according to the discussion exposed in section II the training

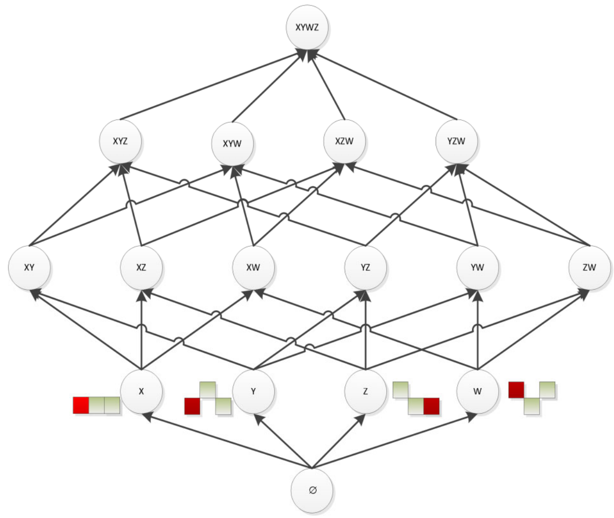

Figure 4. Power set for one subspace (the first horizontal line of fig. 3) of the #44 F_4 collection ordered with respect to inclusion

phase of the WD verification scheme follows: for each writer, *#nref* reference samples of genuine along with an equal number of simulated-forgery signature samples are randomly chosen in order to train the classifier. The "S2" image segmentation scheme combines the features calculated on the whole signature image as well as the relevant 2x2 equimass segmentation grid [20]. These features supply the classifier training section without assuming any additional processing. The classifier used is a hard-margin two class support vector machine (SVM) classifier using radial basis kernel. Selection of the training samples for the genuine class was accomplished using randomly chosen samples according to the hold-out validation method. The remaining genuine and simulated forgery signatures feature vectors, drawn using the same F_4 collection, feed the SVM classifier directly for testing. The SVM output apart from the binary class decision provides a score value which is equal to the distance of the tested sample from the SVM separating hyperplane. The operating parameters of the SVM have been determined through exhaustive search. It is noted that there is a wide area of rbf sigma values that the system has the reported results.

Evaluation of the verification efficiency of the system is accomplished with the use of a global threshold on the overall SVM output score distribution. This is achieved by providing the system's False Acceptance Rate (FAR: samples not belonging to genuine writers, yet assigned to them) and the False Rejection Rate (FRR: samples belonging to genuine writers, yet not classified) functions. With these two rates, the receiver operator characteristics (ROC) are drawn by means of their FAR/FRR plot. Then, classification performance is measured with the utilization of the system Equal Error Rate (EER: the point which FAR equals FRR).

IV. RESULTS

According to the discussion presented above, FAR, FRR and the relevant EER rates, are evaluated for (a) CORPUS 1 and and (b) CORPUS 2 with five and twelve reference

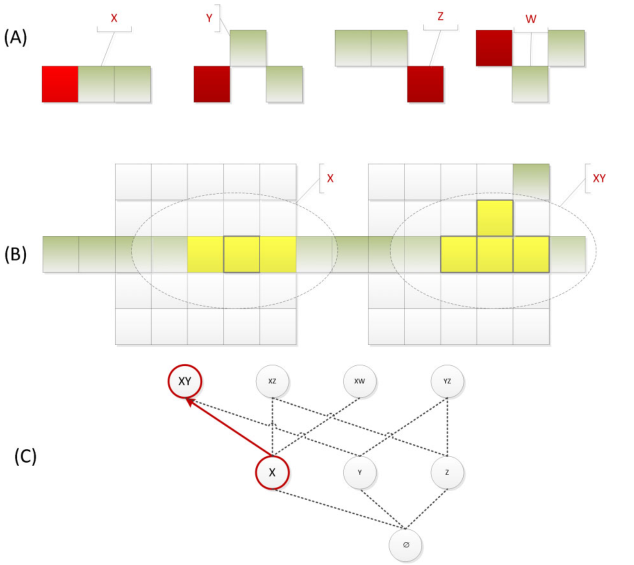

Figure 5. (A) One set of the #44 F_4 collection as depicted in fig. 3b. (B) left and right grids labeled as One (O) and Two (T) respectively imposed on a signature trace (mark with green shadowed pixels) and corresponding events activated. For illustration purposes the topological grids have a distance of 7 instead of 1 that is followed at the actual feature extraction method. (C) Ordered event detection is designated between the red circles and feature component update along red line.

samples for both genuine and forger class. The corresponding results are presented in Table I by means of the mean FAR, FRR and EER values. The letters G and F in Table I designate the genuine and skilled forgery samples respectively. In addition, the ROC curves are presented for both databases in fig. 6 along with their corresponding EER defined as the cross section of the ROC curves and the diagonal.

Our results are compared to recently published relevant figures. The reported results for CORPUS 1 are compared with the results relevant to those reported in [12] for feature level simulated forgery verification tests using 'S2' scheme using (a) nref=5 and (b) the mean value of nref=10 and nref=15 tests for comparison with our test using nref=12. The comparison results are presented in Table II. Concerning CORPUS 2, we present in Table III, the results of recently reported research work using nref=5 and nref=12, along with the results of the current approach.

V. Conclusions

In this work a handwritten model based on the powerset of an ordered event topology with respect to inclusion is considered as a tool for offline signature verification. A number of verification experiments based on an SVM classifier have been carried out in two signature databases namely the GPDS and a proprietary one. Primary verification results indicate that the proposed feature extraction method has an appealing aspect; As a comment on the efficiency of the method one can state that in the case of the Corpus 1 a substantial improvement is observed while in the case of Corpus 2 the results are comparable with those of the literature. Since the approach described in this case study is preliminary it is anticipated that further exhaustive research will unveil important conclusions with respect to the modeling of handwriting. However a number of various other models and experimental setups including i.e. the dissimilarity framework [10] need to be examined in order to verify the effectiveness of the proposed approach.

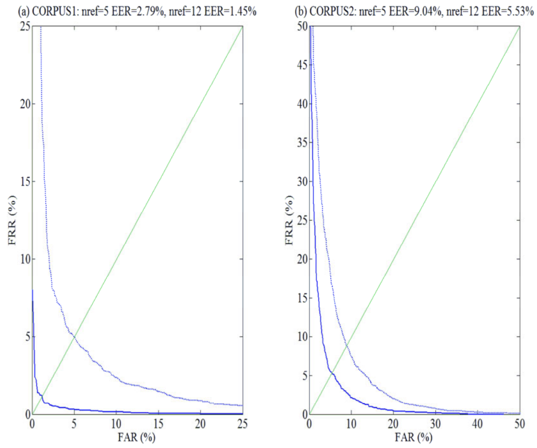

Figure 6. ROC curves with the corresponding EER for corpuses 1, 2.

TABLE I. VERIFICATION EFFICIENCY (%)

Experimental Set	FAR	FRR	EER
CORPUS 1, #nref=5 (GF)	2.18	3.29	2.79
CORPUS 2, #nref=5 (GF)	13.03	5.23	9.04
CORPUS 1, #nref=12 (GF)	1.13	1.60	1.45
CORPUS 2, #nref=12 (GF)	7.73	3.45	5.53

TABLE II. COMPARING EER WITH APPROACH [20]

Experimental Set	EER (%)
[20] #nref=5 (GF)	9.16
Proposed #nref=5 (GF)	2.79
[20] #nref=12 (GF)	4.65
Proposed #nref=12 (GF)	1.45

TABLE III. COMPARING EER WITH VARIOUS APPROACHES (%)

Method	EER		EER
[20] #nref=5 (GF)	12.32	***Proposed #nref=5***	9.04
[12] GPDS-100 nref=5 (GF)	12.02		
[19] #nref=13 (only G)	4.21		
[20] for nref=12 (GF)	6.2	***Proposed #nref=12***	5.53
[12] # ref = {10G, 15F}	8.26		
[13] #refn=12 (GF)	13.76		
[24] #nref=12 (GF)	15.11		
[25] # nref=12 (only G)	15.4		

References

[1] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 63-84, 2000.

[2] F. Leclerc and R. Plamondon, "Automatic Signature verification: the state of the art-1989-1993", International Journal of Pattern Recognition and Artificial Intelligence, vol. 8, pp. 643-660, 1994.

[3] D. Impedovo, and G. Pirlo, "Automatic signature verification: The state of the art, "IEEE Transactions on Systems Man and Cybernetics" vol. 38, pp. 609-635, 2008.

[4] L. Batista, D. Rivard, R. Sabourin, E. Granger, and P. Maupin, "State of the art in off-line signature verification". In Verma, B., Blumenstein, M. (eds.) Pattern Recognition Technologies and Applications: Recent Advances, pp. 39–62, 2008.

[5] M. C. Fairhurst, "Signature verification revisited: promoting practical exploitation of biometric technology", Electron. Commun. Eng. J., vol. 9, pp. 273–280, 1997.

[6] R. C. Gonzalez and R. E. Woods, "Digital Image processing", Addison Wesley, Reading, 1992.

[7] S. Theodoridis and K. Koutroumbas, "Pattern Recognition", Academic Press, 2009.

[8] R. O. Duda and P. E. Hart, Pattern classification. New York: John Wiley and Sons, 2001.

[9] D. Bertolini, L. S. Oliveira, E. Justino, and R. Sabourin, "Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers". Pattern Recognition, vol. 43, pp. 387–396, 2010.

[10] D. Rivard, E. Granger and R. Sabourin, "Multi feature extraction and selection in writer independent off-line signature verification", International Journal on Document Analysis and Recognition, vol. 16, pp. 83-103, 2013.

[11] V. K. Madasu, and B. Lovell. "An Automatic Off-line Signature Verification and Forgery Detection System", in Verma, B., Blumenstein, M. (eds.) Pattern Recognition Technologies and Applications: Recent Advances, pp. 63–89, 2008.

[12] J. F. Vargas, M. A. Ferrer, C. M. Travieso and J. B. Alonso, "Off-line signature verification based on grey level information using texture features", Pattern Recognition, vol. 44, pp. 375-385, 2011.

[13] R. Kumar, J. D. Sharma, and B. Chanda, "Writer independent off-line signature verification using surroundedness feature", Pattern Recognition Letters, vol. 33, pp. 301-308, 2012.

[14] D. Impedovo, G. Pirlo, L. Sarcinella, E. Stasolla, and C. A. Trullo, "Analysis of Stability in Static Signatures using Cosine Similarity", in: Proc of International Conference on Frontiers in Handwriting Recognition, pp. 231-235, 2012.

[15] B. H. Shekar, and R. K. Bharathi, "LOG-Grid based off-line signature verification", in Fourth International Conference on signal and image processing, 2012. S. Mohan, S., Kumar, S.S. (eds), LNEE 222, pp. 321-330, 2013.

[16] J. P. Swanepoel, and J. Coester. "Off-line signature verification using flexible grid features and classifiers fusion", in International Conference on Frontiers in Handwriting Recognition, pp. 297-302. , 2012.

[17] M. K. Kalera, S. Shrihari and A. Xu, "Offine line signature verification using distance statistics", International Journal of Pattern Recognition and Artificial Intelligence, vol. 18, pp. 1339-1360, 2005.

[18] A. Gilperez, F. Alonso-Fernandez, S. Pecharroman, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Off-line signature verification using contour features", in International Conference on Frontiers in Handwriting Recognition ICFHR 2008.

[19] M. Parodi, J. C. Gomez, and A. Belaid, "A circular grid-based rotation invariant feature extraction approach for off-line signature verification", in 11th International Conference on Document Analysis and Recognition, pp. 1289-1293. 2011.

[20] K. Tselios, E. N. Zois, E. Siores, A. Nassiopoulos, and G. Economou, "Grid-based feature distributions for off-line signature verification". IET Biometrics, vol. 1, pp. 72-81, 2012.

[21] T. M. Cover, and A. Y. Thomas, "Elements of information theory", 2nd ed. John Wiley and Sons (2006).

[22] P. R. Halmos, "Naive set theory". The University Series in Undergraduate Mathematics. van Nostrand Company, Princeton, 1960.

[23] Hazewinkel, M.: Encyclopedia of Mathematics, Springer, 2001.

[24] V. Nguyen, Y. Kawazoe, T. Wakabayashi, U. Pal, and M. Blumenstein, "Performance Analysis of the Gradient Feature and the Modified Direction Feature for Off-line Signature Verification", in Proc. 2010 Int Conf on Frontiers in Handwriting Recognition, pp. 303-307, 2010.

[25] M. B. Yilmaz, B. Yanikoglu, C. Tirkaz, and A. Kholmatov, "Offline signature verification using classifier combination of HOG and LBP features", in Proc. Int Joint Conf on Biometrics, pp. 1-7, 2011.