PERGAMON

Pattern Recognition 34 (2001) 47–61

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Fusion of correlated decisions for writer verification

E.N. Zois, V. Anastassopoulos*

Electronics Laboratory, Electronics and Computers Division, Physics Department, University of Patras, Patras 26500, Greece

Received 18 February 1999; received in revised form 12 August 1999; accepted 15 September 1999

Abstract

A fusion approach is proposed for improving the efficiency of writer verification systems. A short handwritten sentence is employed for this purpose. Each word of the sentence is used to tackle an individual verification problem. Then, the word-level (local) decisions are fused in order to obtain a more reliable global decision by means of the Neyman–Pearson approach. The correlation of the local decisions is extensively studied and incorporated in the fusion procedure by means of the Bahadur–Lazarsfeld expansion series. A database containing 4800 sentences is employed to validate the performance of the method. The improvement in verification performance is due to both the fusion procedure applied and the full discretion of the writer to choose his own secret word.

Keywords: Writer verification; Correlated decisions; Decision fusion; Handwritten text

1. Introduction

In financial transactions or other high security contacts, our identity is required in order to gain access to a number of facilities. A substantial number of person verification systems incorporates handwriting which is a behavioral biometric [1,2]. The acquisition of written patterns constitutes a non-invasive process with minimum effects on health or private rights [2,3]. Security systems based on handwriting can be categorized according to Fig. 1. Writer verification can be carried out using either signature or handwritten words [4,5]. Fusion of the information comprised in handwritten patterns for improving verification performance is proposed in this work.

A great deal of work has been reported in the literature for writer verification by means of signature analysis [3–5]. Signature analysis and verification is usually carried out by modeling the genuine and false samples using an adequate shape descriptor [6–10]. The verification efficiency of these systems is measured by means of type I (miss) and type II (false alarm) errors. Despite the fact that signature is considered an attribute which uniquely characterizes a person, it may lead to a number of drawbacks when it is employed in verification procedures. Handwriting variability as well as the capability of forgers to produce good quality specimens, diminish the reliability of the verification system. Accordingly, other handwritten patterns could be used on a complementary basis to enhance the overall system efficiency.

In this paper a method for improving the reliability of an automated handwritten signature verification system (AHSVS) using a short sentence is proposed. The employment of handwritten words is justified by the fact that handwritten text contains stable and significant information for the handwriting of a writer [5]. Additionally, the use of a sentence drawn up by the writer himself will further increase the reliability of the verification system. However, the number of words in the sentence must be small in order to avoid mistakes in memorizing its content. In our experiments a five-word sentence is used for writer verification. For this purpose, a large database was created employing 20 persons to record two different types of sentences, containing a total of 24,000 words. In the verification procedure a computationally simple feature was selected, since we are not

* Corresponding author. Tel.: + 3061-996147; fax: + 3061-997456.

E-mail address: vassilis@physics.upatras.gr (V. Anastassopoulos).

0031-3203/00/$20.00
PII: S0031-3203(99)00210-1

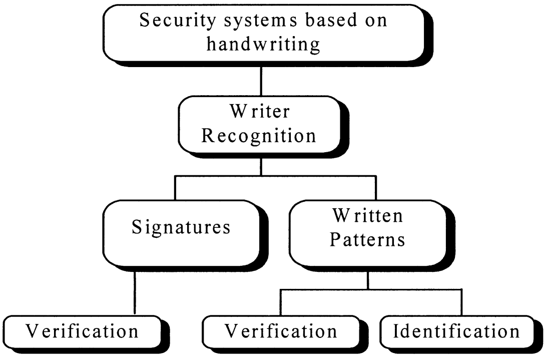

Fig. 1. Categories of security systems based on handwriting.

interested in the meaning of the words but only in the general characteristics of the curves involved. Accordingly, each word is represented in the feature space by means of a granulometric feature [9]. This feature is based on morphologically processing the projective profiles of the words. For improved discrimination performance other features [4–10] can also be used for word-level decision making.

The extracted features are used to tackle an individual verification problem for each word. Thus, five decisions are obtained from each sentence concerning the identity of a specific person (word-level decision). Each decision is obtained using single hypothesis testing with weighted distance measures. These five individual decisions are combined by means of a decision fusion algorithm (DFA) so as to obtain the final and more reliable decision [11,12]. The degree of correlation among the decisions is a critical parameter for consideration when addressing the DFA [13–17]. The decisions obtained from the words were found to be correlated. This is due to the fact that they resulted from words written by the same person, containing similar line attributes and in some cases, the same letters [18]. The Neyman–Pearson formulation is applied in the DFA since it is regarded as the optimal scheme [13,17], compared to the Bayesian approach [21]. An existing procedure [17] is employed to evaluate the efficiency of the N–P test, when the second-order correlation coefficients are indexed by a simple parameter. This single parameter, called the correlation index, was found 0.15 for the decisions obtained from our database. Experimental results display a discrimination error smaller than 1% for a five-word sentence. This error can be considered satisfactory since both the probability of false alarm and the probability of detection for the word level decisions were poor (0.1 and 0.9, respectively). Improvement in the DFA performance can be achieved using more discriminative features to enhance the quality of the word-level decisions, as well as, an increased number of words in the sentence employed. Simulation results describe the dependence of this error on the number of words in the sentence as well as the correlation among the decisions.

This work is organized in the following way. Section 2 presents an overview of the proposed verification system structure and the employed database. Section 3 gives a brief description of the feature used and addresses the word-level decision procedure. In Section 4 analysis of the Neyman–Pearson decision fusion rule is provided. In Section 5 a summary of the proposed method is presented using two flow chart diagrams. A model which describes the way the correlation between the decisions affects the performance of the fusion procedure is described in Section 6. Experimental and simulation results are given in Section 7, while the conclusions are drawn in Section 8.

2. The database

The proposed system architecture for increased reliability in writer verification is shown in Fig. 2. Each person uses a specific PIN number as an index in order to enter the database in which his/her personal handwriting information has been recorded. Then, the writer is requested to write down a short five-word secret sentence along with his signature. Each word is preprocessed and features are extracted in order to produce the local binary decision u_i. The set of local decisions u_i is then combined using the fusion rule and the final decision is made. In this work, the signature is not taken into consideration and consequently the decision u_0 is skipped. This is due to the fact that the signature is a different type of curve, containing information irrelevant of the handwriting attributes of the words. On the other hand, attention is paid to the information content of the words and the existing degree of correlation.

The developed database consists of two different small sentences, one written in Greek and the other one in English. Twenty persons were employed for this purpose within a time period of three months. Each sentence was written by each writer 120 times. Consequently, 4800 sentences were recorded in our database containing a total of 24,000 words. The database is accessible freely via the Internet [19]. The sentence given in Fig. 3(a), is the Greek equivalent of “New method in graphological analysis”. The Greek language, being our native language, was used in order to maintain constant handwriting characteristics. The sentence is made up of two words with relatively small length (three letters), two medium length words (seven letters) and a lengthy word (11 letters). The English sentence shown in Fig. 3(b), was selected in order to further test the proposed writer verification procedure. Each word was recorded inside a specific field so that preprocessing and feature extraction procedures could easily be carried out. The number of sentences written by each person is considered adequate

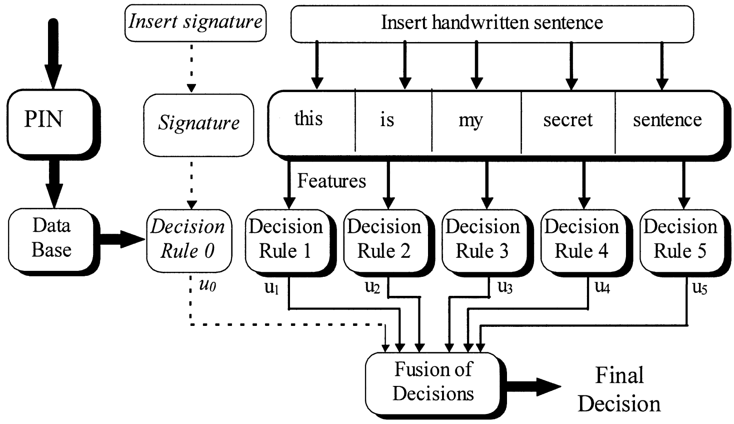

Fig. 2. The proposed system architecture. The signature (dashed) path is optional.

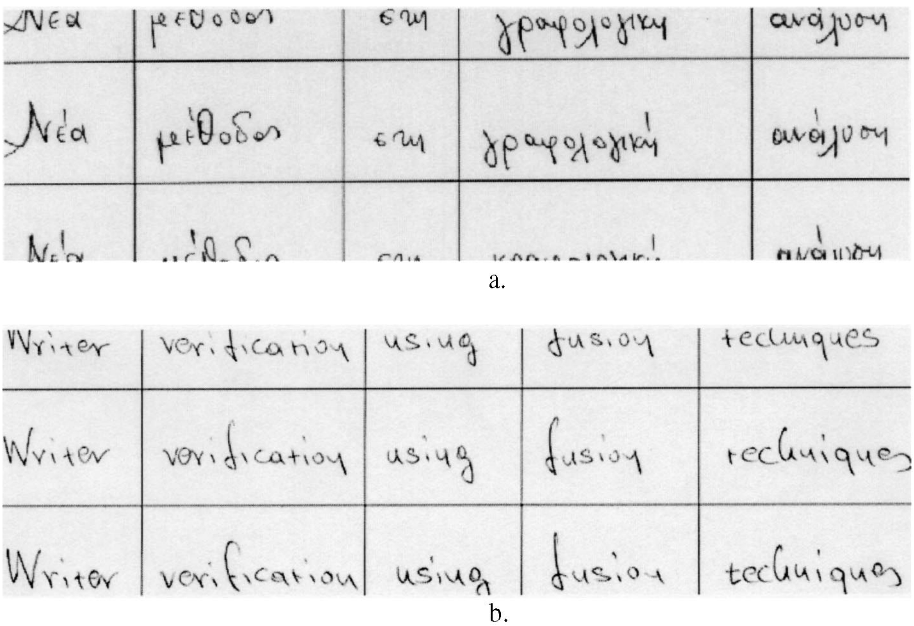

Fig. 3. Samples of the database constructed to apply decision fusion for improving reliability in writer verification. (a) The Greek sentence and (b) the English sentence. 2400 sentences of each type was recorded containing a total of 24,000 words.

for studying the correlation among local decisions. Simulated data consisted of a large number of correlated decisions, were also created to further test the Data Fusion Algorithm (DFA).

To the knowledge of the authors no other database, freely available in the Internet, was found suitable for testing the proposed method. The publicly available databases do not contain any kind of text written by specific persons repeatedly, as the database described previously.

The preprocessing stage incorporates thresholding to binarize the images and thinning of each word in order to obtain a one-pixel-width trace. This is because we need to remove the redundancies that originate form sources like ink absorption, background scanning noise, and type of pen used. The training samples were enrolled in the database as follows: For each writer we used 60 samples representing the genuine class, denoted hereafter as H_1, and 1140 samples from the rest nineteen writers representing the forged class, denoted as H_0. The remaining samples were used to test the efficiency of the system by means of measuring miss and false alarm errors.

3. Word-level decision

3.1. Feature extraction

Among the various shape descriptors that have been used for handwritten pattern representation and signature analysis are granulometrics [6]. A granulometric feature vector is employed in this work for word representation [9]. It contains spatial information about the orientation of the line segments in a handwritten pattern. Accordingly, the binary image of each word is partitioned into sub-blocks. The partition $W(n,m)$ of the word is defined as the division of the original image into a grid of $n \times m$ equal rectangular blocks. Each sub-block is designated by the indexes i, j as follows:

$W_{i,j}(n,m)$ or $W_{i,j}$: *The* (i,j) *sub* $-$ *block of the partition* $W(n,m)$.

Next the projection functions $f_{i,j}(n,m)$ (or simpler $f_{i,j}$) of the partition $W(n,m)$ are defined. The vertical projection function $f^V_{i,j}$ (VPF) is defined as the sum of the black pixels with the same abscissa k inside the $W_{i,j}$ sub-block:

$$f^V_{i,j}(k) = \sum \textit{black pixels with abscissa } k,$$
$$i = 1,\ldots,n, \quad j = 1,\ldots,m.$$

Similarly, the horizontal projection function $f^H_{i,j}$ (HPF) is defined as the sum of the black pixels inside the $W_{i,j}$ sub-block with the same ordinate l:

$$f^H_{i,j}(l) = \sum \textit{black pixels with ordinate } l,$$
$$i = 1,\ldots,n, \quad j = 1,\ldots,m.$$

Obviously for the $W(n,m)$ partition a total of $2 \times n \times m$ projection functions are evaluated. Fig. 4 demonstrates the family of the projection waveforms obtained when the partition $W(2,2)$ is applied to the Greek word 'graphologic'. The final feature vector is obtained when two successive morphological openings $f_{i,j} \circ g_k$ are performed on the projection functions $f_{i,j}$ with a line structuring element (SE) g_k having two different lengths. Fig. 5 illustrates the effect of two morphological openings on the vertical projection function of a real image with SE length three and seven. As a result, the corresponding parameters e (fine details) and c (coarse details) are derived, which measure the gradual reduction in the area of each waveform according to

$$e^V_{i,j} = \left[\frac{mes(f^V_{i,j}) - mes(f^V_{i,j} \circ g_1)}{mes(f^V_{1,1})}\right], \tag{1a}$$

$$e^H_{i,j} = \left[\frac{mes(f^H_{i,j}) - mes(f^H_{i,j} \circ g_1)}{mes(f^H_{1,1})}\right], \tag{1b}$$

$$c^V_{i,j} = \left[\frac{mes(f^V_{i,j} \circ g_1) - mes(f^V_{i,j} \circ g_2)}{mes(f^V_{1,1})}\right], \tag{1c}$$

$$c^H_{i,j} = \left[\frac{mes(f^H_{i,j} \circ g_1) - mes(f^H_{i,j} \circ g_2)}{mes(f^H_{1,1})}\right], \tag{1d}$$

where *mes*(.) is the area under the function in the argument and $f^V_{1,1} = f^H_{1,1} = f_{1,1}$ are the projection functions

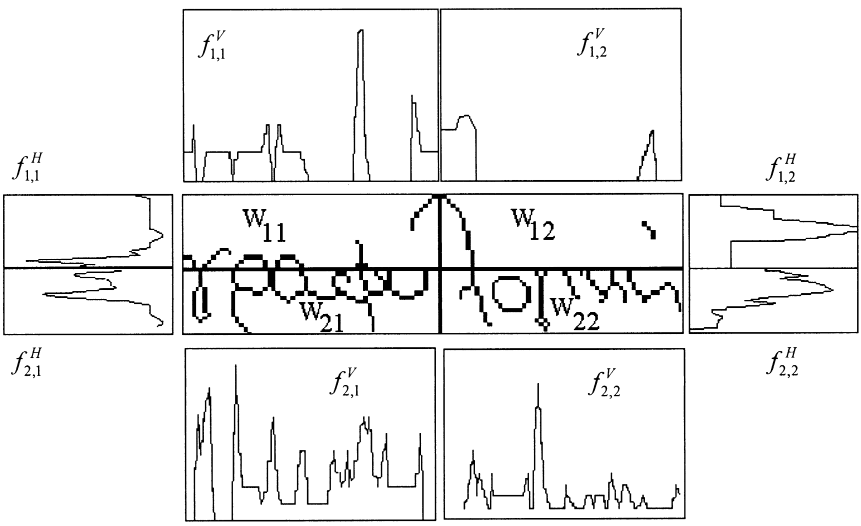

Fig. 4. Family of waveforms derived from the binarized image of the Greek word 'graphologic' when a partition level (2, 2) is applied.

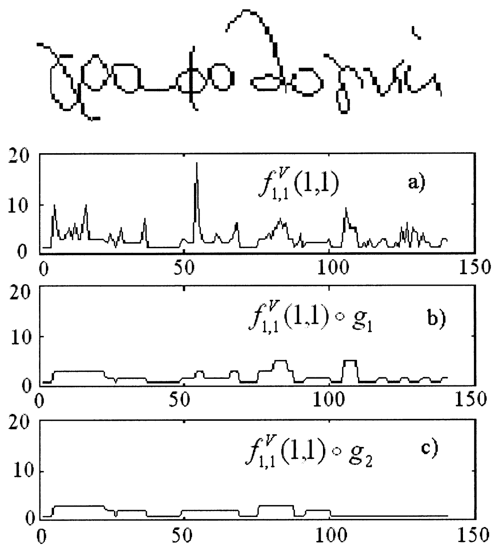

Fig. 5. Morphological openings on the projection function of a real image. (a) Initial projection function $f^V_{1,1}(1,1)$. (b) opening with line SE g_1 of length 3 provides the $e^V_{1,1}$ parameter, (c) opening with line SE g_2 of length 7 provides the $c^V_{1,1}$ parameter.

corresponding to the primary (1,1) partition. The normalization factor $f_{1,1}$, which corresponds to the number of pixels in the initial image, is used in order to achieve finite dynamic range for the obtained feature components as well as to make intraclass dispersion insensitive to natural variations of the genuine class. The set of all parameters $\{e^V_{1,1}, e^H_{1,1}, c^V_{1,1}, c^H_{1,1}, e^V_{1,2}, e^H_{1,2}, c^V_{1,2}, c^H_{1,2}, \ldots, e^V_{n,m}, e^H_{n,m}, c^V_{n,m}, c^H_{n,m}\}$ constitutes the feature vector corresponding to each word. It is obvious that the feature space dimensionality is determined by the partition level. In the general case of a $n\times m$ partition, the procedure results in a $(4\times n\times m)$-dimensional feature vector.

3.2. Local decisions

The verification efficiency using only one word depends on both the partition level and the size of the SE used. If a relatively small partition is employed (i.e. 1,1) then the classification efficiency is poor. This is due to the low dimensionality of the feature vector and its low space discrimination capabilities. On the other hand, a large partition (i.e. 5,5), especially when applied to a short word, yields a feature vector with high dimensionality and quite sensitive components. The length of the SEs used must be chosen so that the correlation between the components of the feature vector be kept as low as possible. In the experimental procedure the partition levels between $(1,1)$ up to $(3,3)$ were tested, while the length l of the SEs g_1 and $g_2(l_2 > l_1)$ got values from the set $\{3, 5, 7, 9, 11\}$. Experimentally, the best verification performance was achieved when the partition level (2, 2) was applied to the short words, while the partition (3, 2) to the long words. This resulted in a feature dimensionality of 16 and 24 respectively. However, using eigenvalue analysis, the intrinsic dimensionality of the obtained feature space is found to be much smaller. The optimal length for the SEs was found to be the same for all the configurations and was set to three and nine.

A specific word sample is classified either as genuine (H_1: the specific writer is present) or forger (H_0: the specific writer is not present). For each word, the training samples form the required feature space. The cluster which corresponds to hypothesis H_1 is a multidimensional pdf $p(x|H_1)$ which is usually well known, since it comprises information about the genuine writer. The rest of the clusters (forgers) form the pdf $p(x|H_0)$, which is generally unknown. The clusters corresponding to both classes have almost the same mean vectors since the samples are generated from the same words. However, they have different covariance matrices making the shape of the cluster a hyperellipsoid. Thus a single hypothesis scheme can be employed [20]. Typically, we measure the distance y of a sample $\mathbf{x}$ from the mean (M_1) of H_1 class (normalised by class covariance matrix C_1) according to the equation

$$y = \{(\boldsymbol{x} - \bar{M}_1)C_1(\boldsymbol{x} - \bar{M}_1)^{\mathrm{T}}\}^{1/2}. \quad (2)$$

The single hypothesis scheme transforms the hyperellipsoid in the feature space into a donut in the distance space [20]. Classification of the unknown feature vector $\mathbf{x}$ involves the selection of two suitable thresholds y_1 and y_2 $(y_1 < y_2)$ to decide upon the validity of either H_1 or H_0:

$$H_1:\quad y_1 < y < y_2,$$

$$H_0:\quad \text{elsewhere.} \quad (3)$$

The total classification error equals to the weighted summation of the P_{fa} and P_m:

$$\textit{Total classification error} = \tfrac{1}{2}(P_{fa} + P_m) = \tfrac{1}{2}(1 + P_{fa} - P_d). \quad (4)$$

P_{fa} is the probability to accept hypothesis H_1 while hypothesis H_0 is true,

$$P_{fa} = \sum (\text{samples classified as } H_1 \text{ but belong to } H_0)/1140 \quad (5)$$

and P_m is the probability to accept hypothesis H_0 while hypothesis H_1 is true,

$$P_m = \sum (\text{samples classified as } H_0 \text{ but belong to } H_1)/60 = 1 - P_d, \quad (6)$$

where P_{d} is the corresponding probability of detection.

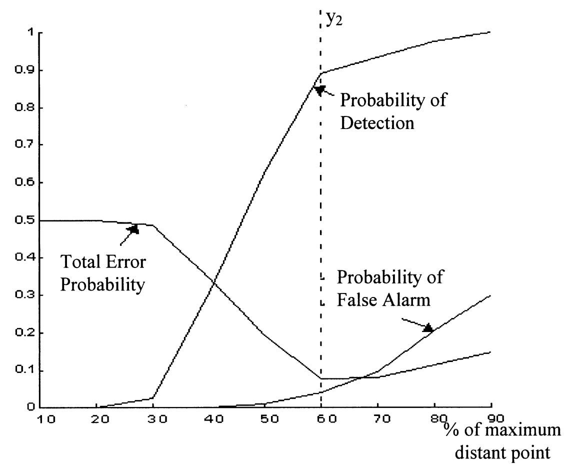

Fig. 6. Operating characteristics for a specific writer and word. The threshold y_2 is selected at the position of the minimum total error, which in this case is at 0.6 of the maximum distant point of cluster $p(y|H_1)$.

For an unknown feature vector **x** the described classifier will decide whether hypothesis H_1 is valid ($u_i = 1$) or not ($u_i = 0$). These hard decisions u_i, $i = 1, \ldots, 5$, are extracted for all five words of each sentence and for the entire set of writers. Each u_i corresponds to a different word and it is associated with the individual P_{fa_i} and P_{m_i} (local operating points). The above quantities were evaluated for every single writer so that the minimum classification error is achieved when the thresholds y_1 and y_2 are located at some portion of the maximum distant point of the genuine cluster. This portion was experimentally found between 60 and 70% for the y_2 threshold, while it was kept at 10% for y_1, for almost all clusters formed in the experimental procedure. In Fig. 6, the P_d, P_{fa} and the minimum classification error are shown for a specific writer and word.

4. Fusion of decisions

The decisions u_i cannot be considered independent since the information content of each word is not totally different from that of the others. This is because the same type of curves, letters or even syllables are common in the words of the sentence. The classification algorithm which maps the feature space to the decision space, conveys the correlation from the vectors to the individual decisions u_i. The vector

$$U = [u_1, u_2, u_3, u_4, u_5]$$

of the correlated decisions u_i, is used in the decision fusion algorithm (DFA) by means of the proper fusion rule in order to obtain a more reliable decision about the presence of a writer.

4.1. The fusion procedure

The optimum combining scheme for the DFA, when the local decisions are provided along with their correlation, is the Neyman–Pearson (N–P) approach [17]. Since the input space, as represented by the local decisions is discrete, a randomized N–P decision rule is to be used. This is defined as follows:

$$\theta(U) = \begin{cases} 1 & \text{if } T(U) > t, \\ \eta & \text{if } T(U) = t, \\ 0 & \text{if } T(U) < t, \end{cases} \tag{7}$$

where $\theta(U)$ corresponds to the final decision and expresses the probability of accepting the presence of a writer (H_1), given that the DFA observes U. The quantity $T(U)$ is the likelihood ratio

$$T(U) = \frac{P_1(U)}{P_0(U)}, \tag{8}$$

where $P_j(U)$ is the probability of U under hypothesis H_j, $j = 0, 1$. The randomization constant η and the threshold t must be chosen so that the overall system probability of detection:

$$P_D(\theta) = E_1[\theta(U)] = P_1(T(U) > t) + \eta P_1(T(U) = t) \tag{9}$$

and the overall system probability of false alarm:

$$P_F(\theta) = E_0[\theta(U)] = P_0(T(U) > t) + \eta P_0(T(U) = t) \tag{10}$$

satisfy the N–P criterion:

$$P_F \leqslant \beta \quad \text{and} \quad P_D \geqslant \max_i (P_{d_i}) \tag{11}$$

where β is a pre-specified upper bound for the false alarm probability and P_D the achieved probability of detection at the DFA. P_{d_i} is the probability of detection corresponding to the individual verification procedures determined in the previous section.

Without loss of generality we consider known the conditional probabilities $P_j(U)$. If we enumerate all the possible states U_i, $1 \leqslant i \leqslant 2^N$ (N is the number of words) that the feature vector U can obtain, we can evaluate the likelihood $T(U_i)$ for every case. The quantities $T(U_i)$ represent the scalar abscissae which will be used to determine the statistic $P_j(T(U_i))$. Accordingly, all $T(U_i)$ have to be ordered so that

$$T(U_1) \leqslant T(U_2) \leqslant \cdots T(U_{2^N}). \tag{12}$$

Making use of the property

$$P_j(T(U_i)) = P_j(U_i) \tag{13}$$

it follows that

$$P_0(T(U) > t) = \begin{cases} 1 & \text{if } t < T(U_1), \\ \lambda_i & \text{if } T(U_i) \leqslant t < T(U_{i+1}), \\ 0 & \text{if } t \geqslant T(U_{2^N}), \end{cases} \quad (14)$$

where

$$\lambda_i = \begin{cases} 1 - \sum_{j=1}^{i} P_0(U_j) & \text{if } 1 \leqslant i \leqslant 2^N, \\ 1 & \text{if } i = 0. \end{cases} \quad (15)$$

If the predetermined upper bound β for the required probability of false alarm lies in the interval $\lambda_i \leqslant \beta \leqslant \lambda_{i-1}$, then the selection of the threshold t as well as the randomization constant η are determined by the following relations [17]

$$t = \begin{cases} T(U_i) & \text{if } \lambda_i \leqslant \beta \leqslant \lambda_{i-1}, \\ 0 & \text{if } \beta = 1, \end{cases} \quad (16)$$

$$\eta = \begin{cases} \dfrac{\beta - \lambda_i}{P_0(T(U) = t)} & \text{if } \lambda_i \leqslant \beta \leqslant \lambda_{i-1}, \\ \text{arbitrary} & \text{if } \beta = 1. \end{cases} \quad (17)$$

The overall probability of detection is determined by means of (9), (16) and (17) as follows:

$$P_D = \begin{cases} 1 - \sum_{j=1}^{i} P_1(U_j) + (\beta - \lambda_i) \dfrac{P_1(T(U) = t)}{P_0(T(U) = t)} & \text{if } \lambda_i \leqslant \beta \leqslant \lambda_{i-1}, \\ 1 & \text{if } \beta = 1. \end{cases} \quad (18)$$

4.2. The conditional statistics

From all the above mentioned, it is obvious that a critical issue towards realizing the randomized N–P test is the determination of the conditional statistics $P_j(U_i)$. In order to express the $P_j(U_i)$ in a convenient form the Bahadur–Lazarsfeld expansion series approach is employed [15,21]. Introducing the normalized variables z_i, corresponding to the individual u_i, as

$$z_i = \frac{u_i - p_i}{\sqrt{p_i q_i}} \quad \text{with } p_i = P(u_i = 1), \quad q_i = 1 - p_i \quad (19)$$

and given that

$p_i^0 \to P(u_i = 1|H_0) = P_{fa_i}$, $p_i^1 \to P(u_i = 1|H_1) = P_{d_i}$ we obtain for $z_i \to z_i^{H_j}$:

$$z_i^0 = \frac{u_i - P_{fa_i}}{\sqrt{P_{fa_i}(1 - P_{fa_i})}}, \quad z_i^1 = \frac{u_i - P_{d_i}}{\sqrt{P_{d_i}(1 - P_{d_i})}}. \quad (20)$$

The variable z_i^0 is the way u_i is transformed, assuming that hypothesis H_0 holds, whereas z_i^1 corresponds to normalised u_i when hypothesis H_1 is valid. Using the variables z_i the Bahadur–Lazarsfeld polynomials have as follows:

$$\varphi_i(U) = \begin{cases} 1 & i = 0, \\ z_1 & i = 1, \\ z_2 & i = 2, \\ \vdots & \\ z_n & i = n, \\ z_1 z_2 & i = n + 1, \\ \vdots & \\ z_1 z_2 z_3 & i = n + 1 + \dfrac{n(n-1)}{2}, \\ \vdots & \\ z_1 z_2 \ldots z_n & i = 2^n - 1. \end{cases} \quad (21)$$

These polynomials are orthogonal, i.e. $\sum_U \varphi_i(U)\varphi_j(U)Q(U) = \delta_{ij}$ with respect to the kernel function $Q(U)$ given by

$$Q(U) = \prod_{i=1}^{n} p_i^{u_i} q_i^{1-u_i}. \quad (22)$$

The kernel function is the probability density function of U under the independence assumption. It is known that each function $P(U)$ of the binary vector $U = [u_1, u_2, \ldots, u_n]$, such as the fusion rule, can be expressed by employing the above polynomials as follows:

$$P(U) = \sum_{i=0}^{2^n - 1} a_i \phi_i(U) \quad (23)$$

or in the form

$$P(U) = Q(U) \sum_{i=0}^{2^n - 1} \gamma_i \phi_i(U). \quad (24)$$

The coefficients γ_i are interpreted as correlation coefficients given by

$$\gamma_i = \sum_U \varphi_i(U) P(U) = E[\varphi_i(U)]. \quad (25)$$

Depending on the order of $\varphi_i(U)$ the set of the correlation coefficients $\{\gamma\}$ can be seen as zero order, first order, second order, etc., as follows:

$$\begin{aligned} &\gamma_{ij} = \textstyle\sum_U z_i z_j P(U) && \text{second} - \text{order correlation coefficient,} \\ &\gamma_{ijk} = \textstyle\sum_U z_i z_j z_k P(U) && \text{third} - \text{order correlation coefficient,} \\ &\vdots && \\ &\gamma_{12\ldots n} = \textstyle\sum_U z_1 z_2 \ldots z_n P(U) && n\text{th} - \text{order correlation coefficient.} \end{aligned} \quad (26)$$

4.3. Fusion rule formation

The individual conditional probabilities $P_j(U)$ are formed by means of Eq. (24) in the following way:

$$P_1(U) = P(U|H_1) = \prod_{i=1}^{5} p_{d_i}^{u_i}(1-p_{d_i}^{1-u_i})\left[1+\sum_{i<j}\gamma_{ij}^1 z_i^1 z_j^1 + \sum_{i<j<k}\gamma_{ijk}^1 z_i^1 z_j^1 z_k^1 + \cdots + \gamma_{12\ldots5}^1 z_1^1 z_2^1 \ldots z_5^1\right], \quad (27a)$$

$$P_0(U) = P(U|H_0) = \prod_{i=1}^{5} p_{fa_i}^{u_i}(1-p_{fa_i}^{1-u_i})\left[1+\sum_{i<j}\gamma_{ij}^0 z_i^0 z_j^0 + \sum_{i<j<k}\gamma_{ijk}^0 z_i^0 z_j^0 z_k^0 + \cdots + \gamma_{12\ldots5}^0 z_1^0 z_2^0 \ldots z_5^0\right], \quad (27b)$$

and finally, the likelihood ration $T(U)$ which is employed in the decision rule described by Eq. (7), is as follows:

$$T(U) = \frac{\prod_{i=1}^{5} p_{d_i}^{u_i}(1-p_{d_i}^{1-u_i})\left[1+\sum_{i<j}\gamma_{ij}^1 z_i^1 z_j^1 + \sum_{i<j<k}\gamma_{ijk}^1 z_i^1 z_j^1 z_k^1 + \cdots + \gamma_{12\ldots n}^1 z_1^1 z_2^1 \ldots z_5^1\right]}{\prod_{i=1}^{5} p_{fa_i}^{u_i}(1-p_{fa_i}^{1-u_i})\left[1+\sum_{i<j}\gamma_{ij}^0 z_i^0 z_j^0 + \sum_{i<j<k}\gamma_{ijk}^0 z_i^0 z_j^0 z_k^0 + \cdots + \gamma_{12\ldots n}^0 z_1^0 z_2^0 \ldots z_5^0\right]} \quad (27c)$$

In case that the coefficients $\{\gamma\}$ are close to zero, Eqs. (27) result in the relations valid for independent decisions.

5. Outline of the proposed method

The writer verification procedure described in the previous sections is revisited here by means of two flow chart diagrams. In this way, the steps required to train the system and thus, verify the presence of a specific writer are clearly stated. Firstly, the system is to be updated with the data corresponding to a new writer. Accordingly, the training stage described in Flowchart 1 (Fig. 7), starts with the data recording. In the first step the specific writer records his/her own secret sentence a number of times (60 in our experiments), in a formatted paper, like the one shown in Fig. 3. The second step involves feature extraction. In this step the images of the recorded words are preprocessed (thresholding and thinning) and each word is partitioned as described in Fig. 4. The obtained projections are morphologically processed in the way described in Fig. 5 and the feature vector is obtained according to Eqs. (1a)–(1d). The derived vectors are stored in the database. For each specific word of the sentence a cluster is formed in a separate feature space. In this way step 3 of the procedure is completed.

Step 4 deals with the derivation of the local decisions u_i. For this purpose, the distances y_1 and y_2 are selected using the appropriate feature space and the decision rule described by Eq. (3). In this step the quality of the decisions u_i is also evaluated. This is strictly related with the quantities P_{fa_i} and P_{m_i} given by Eqs. (5) and (6) and characterise each decision. It is worth mentioning that each time a new writer is added in the database the quantities P_{fa_i} and P_{m_i} for all existing writers have to be re-evaluated. In the fifth step the correlation between the decisions u_i is evaluated and expressed by means of the coefficients γ_{ij}. After that the performance of the fusion rule is specified evaluating the randomisation constant η and the threshold t, using Eqs. (9)–(18). These parameters along with the correlation coefficients γ_{ij} represent each individual writer and they are used to update the database of the system.

According to Flowchart 2 (Fig. 8) the verification stage is much simpler. The writer to be verified records his/her own simple sentence. The feature extraction procedure is repeated and the local decisions u_i are derived. The final decision is obtained using the vector $U = [u_i]_{i=1,\ldots,N}$ and the fusion rule given by Eqs. (7) and (27c).

6. Correlation impact on system performance

Eq. (27a)–(27c) implies that the parameters required to design the randomized N–P test, are the local operating points P_{fa_i} and P_{m_i} along with the set of correlation coefficients $\{\gamma\}$. According to the discussion in Section 3, the probabilities P_{fa_i} and P_{m_i} are determined from the discriminative behaviour of the corresponding feature vector. Estimation of the correlation coefficients $\{\gamma\}$ was experimentally carried out using Eq. (26) and the available local decisions u_i. Experimentation, using the described features on both types of sentences in our database, resulted in considerable values only for the second order coefficients. The higher order correlation coefficients can be neglected since they were found very small and, consequently, they can hardly affect the final value of $T(U)$ in Eq. (27a)–(27c).

A correlation index [17] is employed in this section to represent the second order correlation parameters γ_{ik}^j. Thus, with the use of a single correlation parameter the efficiency of the Fusion procedure can be evaluated. This single parameter is derived in the following way. Let us consider the second order correlation coefficient ρ_{ik}^j between two local decisions u_i and u_k under hypothesis H_j:

$$\rho_{ik}^j = \frac{E_j[u_i u_k] - E_j[u_i]E_j[u_k]}{\sqrt{\{E_j[u_i^2] - E_j[u_i]^2\}\{E_j[u_k^2] - E_j[u_k]^2\}}} = \frac{E_j[u_i u_k] - E_j[u_i]E_j[u_k]}{\sqrt{E_j[u_i]E_j[u_k]\{1-E_j[u_i]\}\{1-E_j[u_k]\}}}, \quad (28)$$

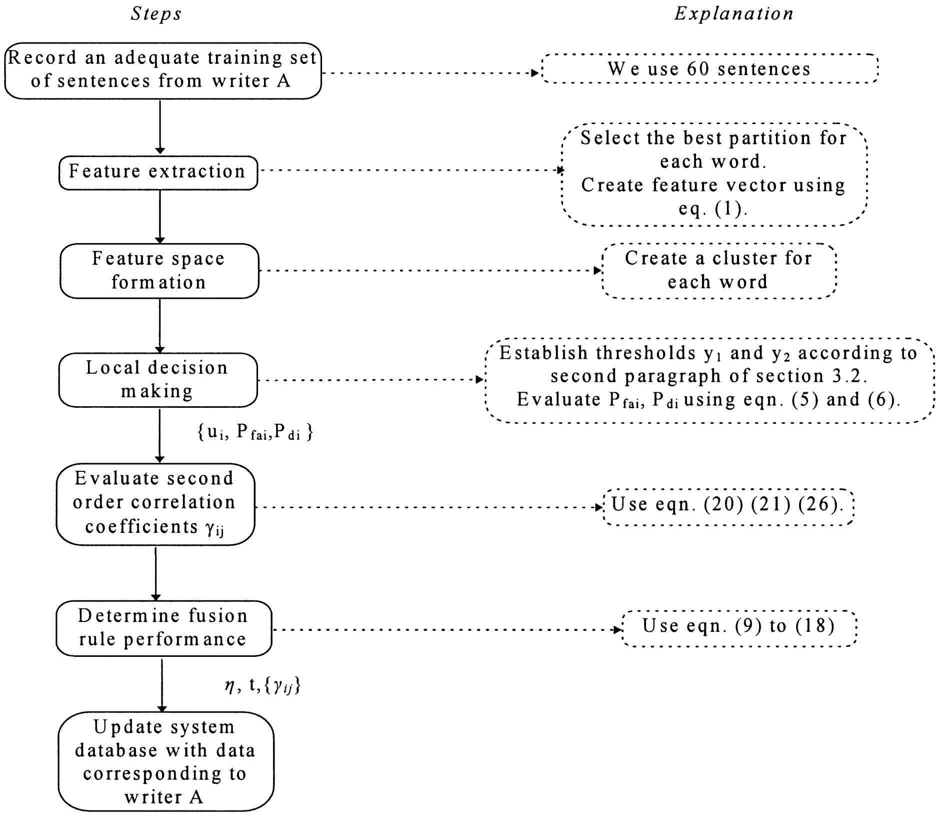

Fig. 7. Flowchart 1. Training stage: train system to incorporate a specific writer into systems database.

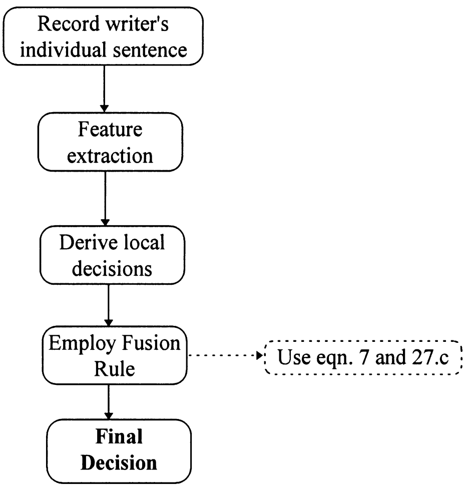

Fig. 8. Flowchart 2. Verification stage: system verifies the presence of a specific writer.

where $j = 0, 1$, $1 \leqslant i, k \leqslant 2^N$, $i \neq k$. The second equality is obtained since $u_i \in \{0, 1\}$ and thus $u_i^2 = u_i$. Since $E_j[u_i u_k] \leqslant \min\{E_j[u_i], E_j[u_k]\} = E_j[u_i]$ and assuming, without loss of generality, that $E_j[u_i] \leqslant E_j[u_k]$ we have

$$\rho_{ik}^j \leqslant \frac{E_j[u_i] - E_j[u_i]E_j[u_k]}{\sqrt{E_j[u_i]E_j[u_k]\{1 - E_j[u_i]\}\,\{1 - E_j[u_k]\}}}$$

$$= \sqrt{\frac{E_j[u_i]\{1 - E_j[u_k]\}}{E_j[u_k]\{1 - E_j[u_i]\}}} \equiv \sigma_{ik}^j. \tag{29}$$

We note that $\sigma_{ik}^j \leqslant 1$ with the equality being valid only if $E_j[u_i] = E_j[u_k]$. The σ_{ik}^j parameter depends only on the operating points P_{fa_i} and P_{m_i} as Eq. (29) implies. The second-order coefficients γ_{ik}^j are equal to the coefficients ρ_{ik}^j since

$$\gamma_{ik}^j = E_j[z_i z_k]$$

$$= E_j\left\{\frac{u_i - E_j[u_i]}{\sqrt{E_j[u_i][1 - E_j[u_i]]}}\,\frac{u_k - E_j[u_k]}{\sqrt{E_j[u_k][1 - E_j[u_k]]}}\right\}$$

$$= E_j\left\{\frac{u_i u_k - u_k E_j[u_i] - u_i E_j[u_k] + E_j[u_i]E_j[u_k]}{\sqrt{E_j[u_i]E_j[u_k]\{1 - E_j[u_i]\}\{1 - E_j[u_k]\}}}\right\}$$

$$= \frac{E_j[u_i u_k - u_k E_j[u_i] - u_i E_j[u_k] + E_j[u_i]E_j[u_k]]}{E_j[\sqrt{E_j[u_i]E_j[u_k]\{1 - E_j[u_i]\}\{1 - E_j[u_k]\}}]}$$

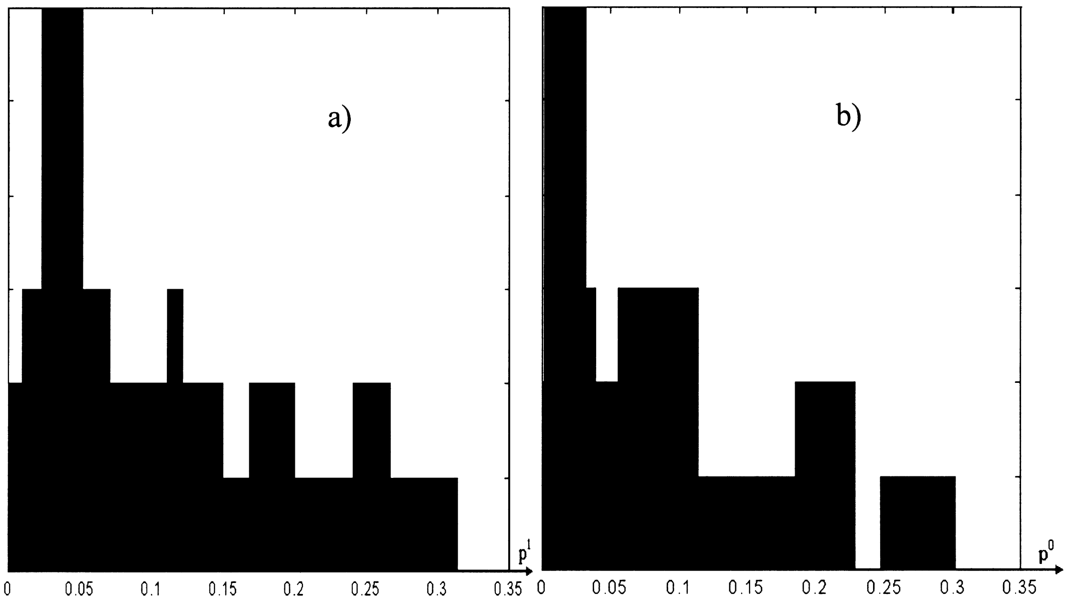

Fig. 9. Histograms of correlation index ρ^j for all writers and the Greek sentence; (a) Hypothesis H_1; (b) Hypothesis H_0.

$$= \frac{E_j[u_i u_k] - E_j[u_k]E_j[u_i] - E_j[u_i]E_j[u_k] + E_j[u_i]E_j[u_k]}{\sqrt{E_j[u_i]E_j[u_k]\{1 - E_j[u_i]\}\{1 - E_j[u_k]\}}}$$

$$= \frac{E_j[u_i u_k] - E_j[u_i]E_j[u_k]}{\sqrt{E_j[u_i]E_j[u_k]\{1 - E_j[u_i]\}\{1 - E_j[u_k]\}}} = \rho^j_{ik} \quad (30)$$

If we assume that $\rho^j_{ik} \geqslant 0$ then from Eq. (29) a non-negative correlation index ρ^j independent of the i, k can be defined as follows:

$$\rho^j = \frac{\rho^j_{ik}}{\sigma^j_{ik}} = \frac{\gamma^j_{ik}}{\sigma^j_{ik}} \quad (31)$$

Fig. 9 shows the histogram of the correlation index ρ^j, evaluated for the entire group of writers and the Greek sentence. Similar results were also obtained for the English sentence. The experimental and simulation results carried out for the fusion procedure are discussed and analyzed taking into consideration the fact that the correlation index ρ^j is relatively low (< 0.3). The cases where ρ^j is large are not of practical use and therefore will not be addressed to. The mean value of the correlation index ρ^j was found to be 0.15, and thus can be used to represent the correlation properties of the employed databases.

Figs. 10 and 11 demonstrate the way the value of the correlation index affects the performance of the fusion procedure. Specifically, the abscissas $T(U_i)$ given by Eq. (27a)–(27c) tend to cluster when the correlation index and, consequently, the γ^j_{ik} coefficients increase. High correlation makes the system operate poorly since $P_1(U)$ and $P_0(U)$ are concentrated on the same region of $T(U)$. In the extreme case, that $\gamma^j_{ik} = 1$ and the operating points for each local decision are equal, it is expected that the statistics for both hypothesis will degenerate to that of a single local decision. On the other hand, a low correlation index unfolds each $P_1(U)$ and $P_0(U)$ and consequently moves them apart. This fact results in improved fusion performance.

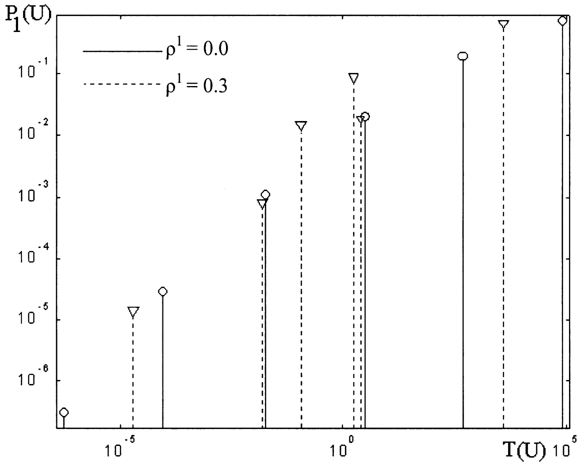

Fig. 10. Probability distribution conditioned on hypothesis H_1 for independent and correlated approaches. Five decisions topology with similar operating points.

The efficiency of the DFA is measured by examining the receiver operating characteristics (ROC, i.e. the total

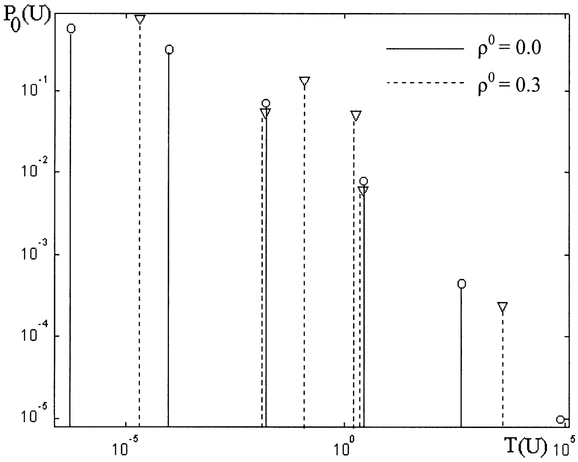

Fig. 11. Probability distribution conditioned on hypothesis H_0 for independent and correlated approaches. Five decisions topology with similar operating points.

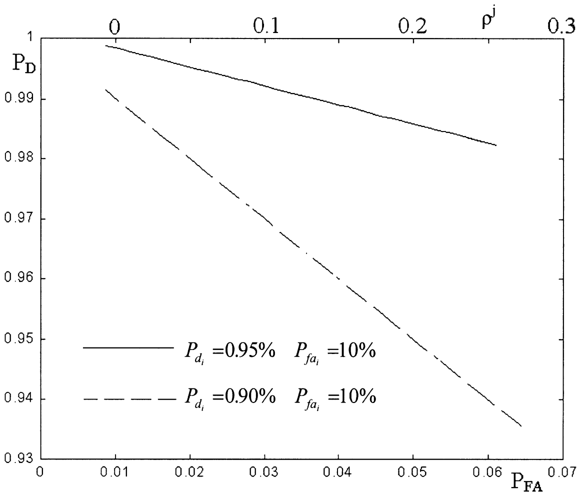

Fig. 12. ROC curves for systems designed with similar operating points.

probability of detection vs. the total probability of false alarm) with the correlation index ρ^j as a parameter. Fig. 12 shows the ROC curves for a fusion system designed so that all local procedures possess the same operating points (P_{fa_i} and P_{m_i}). Fig. 13 corresponds to a fusion system with local procedures working under different operating points. Each case was treated on the assumption that ρ^j is the same for both hypotheses. It is clearly seen that the increase in the index ρ^j results in a linear decrease of the overall system efficiency.

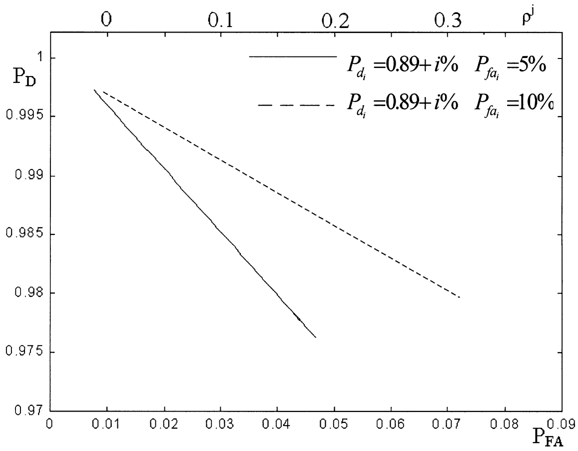

Fig. 13. ROC curves for systems designed with different operating points.

7. Experimental results

The performance of the decision fusion algorithm (DFA) was tested experimentally using two different ways. Firstly, our database was employed using separately the Greek and the English sentences. The obtained local decisions accompanied by the corresponding P_{fa} and P_d are used as inputs to the DFA. According to the second way, the correlated decisions u_i are derived using a simulation approach. This second approach, gives the possibility to accurately test the DFA since the number of the created vectors U is very large.

Actually, in the experimental procedure half of the 120 words from each case were used for cluster formation (training), while the rest 60 were used for local decision making. According to Fig. 5, for obtaining each u_i the threshold y_2 was set at the position of the minimum total error for all local verification problems. The threshold y_1 was kept constant at 10% of the maximum distant point of the corresponding cluster. The operating parameters P_{fa_i} and P_{m_i} were evaluated as well, and accompany the corresponding decisions u_i. Table 1 shows the individual P_{fa_i} and P_{d_i} for 3 of the writers and every word for the Greek sentence. Next, the entire set of the correlation parameters were evaluated in order to form the conditional probability densities and subsequently the likelihood ratio $T(U)$. In the fusion procedure, the required overall probability of false alarm P_F was selected to be at most half of the smaller local P_{fa_i} ($P_F \leqslant \frac{1}{2}\min\{P_{fa_i}\}, i = 1, \ldots, 5$). When this condition is fulfilled the corresponding maximum overall P_D is obtained according to Eqs. (16)–(18). In Table 2 are provided the corresponding overall verification results from the DFA. Similar results are given in Tables 3 and 4 for the English sentence.

Table 1
P_{fa_i} and P_{d_i} for each word and three of the writers for the Greek sentence

Writer no.	Op. points	word 1	word 2	word 3	word 4	word 5
1	P_{fa_i}	0.046	0.059	0.165	0.056	0.071
	P_{d_i}	0.966	0.850	0.900	0.983	0.967
2	P_{fa_i}	0.203	0.273	0.119	0.147	0.197
	P_{d_i}	0.833	0.833	0.900	0.817	0.933
3	P_{fa_i}	0.153	0.141	0.270	0.178	0.013
	P_{d_i}	0.900	0.833	0.983	0.967	0.967

Table 2
Identification results from the DFA using data of Table 1

Writer no.	Minimum P_{fa_i}	Maximum P_{d_i}	Overall P_{FA}	Overall P_D
1	0.046	0.983	0.000	1.000
2	0.119	0.933	0.030	0.985
3	0.013	0.983	0.003	0.990

The fourth and fifth columns in Tables 2 and 4 show off the reliability of the final decision in terms of the P_F and P_D. The DFA gives an improved verification performance in all cases. This is obvious when the local parameters P_{fa_i} and P_{d_i} are not satisfactory (0.119 and 0.933, respectively). The DFA results in significant improvement regarding the overall probabilities of miss and false alarm. For the Greek sentences the total verification error was found to be 0.97%, whereas for the English ones this error amounts to 0.53% due to the longer words contained. In order to examine more precisely the performance of the DFA and the way it is affected from the correlation of decisions, simulated data were devised.

The simulated data comprise decision vectors U with various local operating points (P_{fa_i} and P_{d_i}) and correlation index. Each time 10,000 vectors U were derived for testing the DFA. The procedure used for deriving the correlated components u_i of the vector U, simulates exactly the experimental approach. Table 5 demonstrates the effect of the correlation index ρ on the final performance of the DFA for the two different local operating points. The threshold t at the DFA was selected so that the obtained global P_F to be smaller than half of the local P_{fa_i}. The improvement in the P_D is better for smaller ρ. The results verify those presented in Figs. 9 and 10.

The verification performance of the DFA for different number of words is shown in Table 6, for the same local operating points ($P_{fa_i} = 0.1$ and $P_{d_i} = 0.9$) and correlation index $\rho = 0.3$. From the table is obvious that for the same global P_F the corresponding probability of miss $(1 - P_D)$ is improved when the number of words increases.

Finally, Table 7 presents all possible pairs of (P_F, P_D) obtained from the DFA when the required global P_F can vary according to the discrete nature of $T(U)$. Obviously, the DFA procedure offers a large variety of choices for the final operating point (P_F, P_D) depending on the number N of the words used. When the local operating points are different, 2^N distinct thresholds are available at the DFA. For the experimental results given in Table 7 the value of P_{d_i} is different for each of the five words (0.88, 0.89, 0.90, 0.91, 0.92, respectively) resulting in 32 different thresholds at the DFA.

Table 3
P_{fa_i} and P_{d_i} for each word and three of the writers for the English sentence

Writer no.	Op. points	word 1	word 2	word 3	word 4	word 5
1	P_{fa_i}	0.073	0.044	0.079	0.114	0.039
	P_{d_i}	0.967	0.983	0.900	0.850	0.983
2	P_{fa_i}	0.122	0.098	0.165	0.153	0.083
	P_{d_i}	0.933	0.900	0.850	0.833	0.967
3	P_{fa_i}	0.035	0.158	0.136	0.150	0.199
	P_{d_i}	0.800	0.917	0.950	0.867	0.933

Table 4
Identification results from the DFA using data of Table 3

Writer no.	Minimum P_{fa_i}	Maximum P_{d_i}	Overall P_{FA}	Overall P_D
1	0.038	0.983	0.000	1.000
2	0.083	0.967	0.011	0.995
3	0.035	0.950	0.018	0.991

Table 5
The effect of correlation index on the performance of the DFA

	$P_{fa_i} = 0.1$ $P_{d_i} = 0.9$		$P_{fa_i} = 0.05$ $P_{d_i} = 0.95$	
ρ^j	P_F	P_D	P_F	P_D
0.0000	0.0086	0.9914	0.0012	0.9988
0.0500	0.0183	0.9817	0.0036	0.9964
0.1000	0.0280	0.9720	0.0060	0.9940
0.1500	0.0377	0.9623	0.0085	0.9915
0.2000	0.0454	0.9526	0.0109	0.9891
0.2500	0.0489	0.9428	0.0133	0.9867
0.3000	0.0510	0.9370	0.0158	0.9842

Table 6
The performance of the DFA for different number of words, and the same local operating points ($P_{fa_i} = 0.1$ and $P_{d_i} = 0.9$) and correlation index $\rho = 0.1$ and 0.3

	No. words	P_F	P_D
	3	0.04	0.9526
$p^j = 0.1$	5	0.04	0.9839
	7	0.04	0.9945
	9	0.04	1
	3	0.04	0.9310
$p^j = 0.3$	5	0.04	0.9602
	7	0.04	0.9877
	9	0.04	0.9910

8. Conclusions

Using the proposed decision fusion method, security systems based on handwritten signatures can gain further reliability in writer verification. This is achieved by means of a short handwritten sentence. The words of the sentence are used separately to derive decisions about the authenticity of the writer, and then fused for achieving higher verification performance.

Three different factors can affect the verification performance of the proposed fusion procedure. The first important design parameter is the selection of a discriminative feature vector for modelling the shape of the words. After that, the correlation of the decisions must be studied and appropriately incorporated into the fusion algorithm. Finally, the verification performance of the fusion algorithm is improved when the number of the words in the sentence increases.

Experimental results were obtained using 4800 handwritten sentences and a total of 24,000 words. The correlation among the decisions was experimentally found to be 0.15. Furthermore, simulated data were employed to test the data fusion algorithm. According to these data the achieved verification error can be very small depending on the local operating points and the selected threshold at the DFA. The proposed decision fusion method improves the efficiency of writer verification systems by means of the following two aspects. Firstly, the sentence employed is secret and can be changed by the writer. Secondly the fusion algorithm provides an adequate number of operating points to work with.

9. Summary

Fusion of the information comprised in handwritten patterns for improving verification performance is proposed in this work. A substantial number of person verification systems incorporates handwriting which is a behavioral biometric. A great deal of work for writer verification by means of signature analysis has been reported in the literature so far. Signature analysis and verification is usually carried out by modeling the genuine and false samples using an adequate shape descriptor. Despite the fact that signature is considered an attribute which uniquely characterizes a person, it may lead to a number of drawbacks when it is employed in verification procedures. Handwriting variability as well as the capability of forgers to produce good quality specimens diminish the reliability of the verification system itself. Accordingly, other handwritten patterns could be used on a complementary basis to enhance the overall system efficiency.

In this paper a method for improving the reliability of an automated handwritten signature verification system (AHSVS) using a short sentence is proposed. The employment of handwritten words is justified by the fact that a handwritten text contains stable and significant information for the handwriting of a writer. Additionally, the use of a sentence drawn up by the writer himself will further increase the reliability of the verification system. In our experiments a five-word sentence is used for writer verification. For this purpose, a large database was created asking 20 persons to record two different types of sentences, containing a total of 24,000 words. In the verification procedure, a granulometric feature was selected in order to describe the general characteristics of the line patterns involved in each word. This feature is

Table 7
The performance of the DFA for all possible thresholds used in the N–P approach. The local decisions differ in the value of P_{d_i}

	$p^j = 0.1$			$P^j = 0.3$		
	Abscissa	P_F	P_D	Abscissa	P_F	P_D
1	1.4951e-004	1	1	2.7191e-004	1	1
2	9.4998e-003	0.3505	0.9995	3.5068e-002	0.2849	0.9990
3	1.0556e-002	0.3045	0.9990	3.8991e-002	0.2608	0.9980
4	1.1653e-002	0.2586	0.9984	4.3016e-002	0.2367	0.9970
5	1.2723e-002	0.2127	0.9979	4.6840e-002	0.2127	0.9959
6	1.3608e-002	0.1668	0.9972	4.9796e-002	0.1886	0.9947
7	1.9493e-001	0.1208	0.9954	2.7492e-001	0.1646	0.9914
8	2.1985e-001	0.1113	0.9933	3.1088e-001	0.1527	0.9877
9	2.4234e-001	0.1019	0.9910	3.4144e-001	0.1408	0.9836
10	2.4497e-001	0.0924	0.9887	3.4704e-001	0.1288	0.9795
11	2.5697e-001	0.0829	0.9863	3.5699e-001	0.1169	0.9752
12	2.6998e-001	0.0734	0.9837	3.8108e-001	0.1050	0.9707
13	2.8633e-001	0.0640	0.9810	3.9859e-001	0.0931	0.9660
14	2.9484e-001	0.0545	0.9782	4.1471e-001	0.0812	0.9610
15	3.1212e-001	0.0450	0.9752	4.3265e-001	0.0693	0.9559
16	3.2850e-001	0.0355	0.9721	4.4884e-001	0.0574	0.9505
17	2.9144e + 000	0.0261	0.9653	2.1714e + 000	0.0455	0.9417
18	3.3950e + 000	0.0237	0.9573	2.5870e + 000	0.0414	0.9312
19	3.6401e + 000	0.0214	0.9487	2.6720e + 000	0.0374	0.9204
20	3.8657e + 000	0.0190	0.9397	2.9867e + 000	0.0333	0.9082
21	4.1459e + 000	0.0167	0.9299	3.0905e + 000	0.0293	0.8957
22	4.3253e + 000	0.0143	0.9198	3.1502e + 000	0.0252	0.8829
23	4.3943e + 000	0.0120	0.9094	3.3689e + 000	0.0211	0.8692
24	4.6372e + 000	0.0096	0.8985	3.4868e + 000	0.0171	0.8551
25	4.9142e + 000	0.0073	0.8870	3.5563e + 000	0.0130	0.8406
26	5.1533e + 000	0.0049	0.8749	3.5733e + 000	0.0090	0.8261
27	6.0401e + 001	0.0026	0.8450	6.4574e + 000	0.0049	0.8200
28	7.5625e + 001	0.0021	0.8076	1.5004e + 001	0.0040	0.8059
29	8.9731e + 001	0.0016	0.7632	2.2131e + 001	0.0030	0.7849
30	1.0262e + 002	0.0011	0.7124	2.7717e + 001	0.0021	0.7588
31	1.1417e + 002	0.0006	0.6558	3.1632e + 001	0.0011	0.7289
32	7.2071e + 003	0	0	4.0269e + 003	0	0

based on morphologically processing the projective profiles of the words.

The extracted feature is used to tackle an individual verification problem for each word. Thus, five decisions are obtained from each sentence concerning the identity of a specific person (word-level decision). Each decision is taken using single hypothesis testing with weighted distance measures. These five individual decisions are combined by means of a decision fusion algorithm (DFA) so as to obtain the final and more reliable decision. The degree of correlation among the decisions is a critical parameter for consideration when addressing the DFA. The decisions derived from the words were found to be correlated. This is due to the fact that they resulted from words written by the same person, containing similar line attributes and, in some cases, the same letters. The Neyman–Pearson formulation is applied in the DFA since it is regarded as the optimal scheme compared to the Bayesian approach.

Experimental results display a discrimination error smaller than 1% for a five-word sentence. This error can be considered quite small provided that the mean probability of false alarm for the local decisions equals 0.1, whereas the corresponding mean probability of detection is 0.9. Improvement in the DFA performance can be achieved using more discriminative features to enhance the quality of the word-level decisions, as well as, an increased number of words in the sentence employed. Simulation results describe the dependence of this error on the number of words in the sentence as well as on the correlation among the decisions.

References

[1] B. Carter, Biometric technologies, what they are and how they work, Proceedings of CTST' 97, Orlando, FL, 1997.

[2] J. Woodward, Biometrics: privacy's foe or privacy's friend? IEEE Proc. 85 (9) (1997) 1480–1492.

[3] M.C. Fairhurst, Signature verification revisited: promoting practical exploitation of biometrics technology, Electron. Commun. J. (1997) 273–280.

[4] F. Leclerc, R. Plamondon, Automatic signature verification-the state of the art 1989-1993, Int. J. Pattern Rec. Artif. Intell. 8 (1993) 643–659.

[5] R. Plamondon, G. Lorette, Automatic signature verification and writer identification–the state of the art, Pattern Recognition 22 (2) (1989) 107–131.

[6] R. Sabourin, G. Genest, F. Preteux, Off-line signature verification by local granulometric size distributions, IEEE Trans. Pattern Anal. Mach. Intell. 19 (9) (1997) 976–988.

[7] M. Ammar, Y. Yoshida, T. Fukumura, Off-line preprocessing and verification of signatures, Int. J. Pattern. Rec. Artif. Intell. 2 (4) (1993) 589–602.

[8] J.P. Drouhard, R. Sabourin, M. Godbout, A neural approach to off-line signature verification using directional PDF, Pattern Recognition 29 (3) (1996) 415–424.

[9] E.N. Zois, V. Anastassopoulos, Contour features using morphological transforms, Electron. Lett. 34 (1) (1998) 56–57.

[10] Y. Qi, B.R. Hunt, Signature verification using global and grid features, Pattern Recognition 27 (12) (1994) 1621–1629.

[11] R. Viswanathan, P.K. Varsney, Distributed detection with multiple sensors: Part I-Funtamentals, IEEE Proc. 85 (1) (1997) 54–63.

[12] P.K. Varsney, Multisensor data fusion, Electron. Commun. J. (1997) 245–253.

[13] S.C.A. Thomopoulos, R. Viswanathan, D.C. Bougoulias, Optimal decision fusion in multiple sensor systems, IEEE Trans. Aerospace Electron. Systems 23 (5) (1987) 644–653.

[14] Z. Chair, P.K. Varshney, Optimal data fusion in multiple sensor detection system, IEEE Trans. Aerospace Electron. Systems 22 (1) (1986) 98–101.

[15] M. Kam, Q. Zhu, W.S. Gray, Optimal data fusion of correlated decisions in multiple sensor detection systems, IEEE Trans. Aerospace Electron. Systems 28 (3) (1992) 916–920.

[16] V. Aalo, R. Viswanathan, On distributed detection with correlated sensors: two examples, IEEE Trans. Aerospace Electron. Systems 25 (3) (1989) 414–421.

[17] E. Drakopoulos, C.C. Lee, Optimum multisensor fusion of correlated local decisions, IEEE Trans. Aerospace Electron. Systems 27 (4) (1991) 593–605.

[18] E.N. Zois, V. Anastassopoulos, Fusion of correlated decisions for writer identification, Proceedings of the IEE Third European Workshop on Handwriting Analysis and Recognition, Brussels, 1998.

[19] E.N. Zois, V. Anastassopoulos, Database I (HIFCD1), ftp://anemos.physics.upatras.gr/pub/handwriting/HIFCD1.

[20] K. Fukunaga, Statistical Pattern Recognition, Academic Press, San Diego, 1990.

[21] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1970.