World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:2, No:5, 2008

Through Biometric Card in Romania: Person Identification by Face, Fingerprint and Voice Recognition

Hariton N. Costin, Iulian Ciocoiu, Tudor Barbu, and Cristian Rotariu

Abstract**—**In this paper three different approaches for person verification and identification, i.e. by means of fingerprints, face and voice recognition, are studied. Face recognition uses parts-based representation methods and a manifold learning approach. The assessment criterion is recognition accuracy. The techniques under investigation are: a) Local Non-negative Matrix Factorization (LNMF); b) Independent Components Analysis (ICA); c) NMF with sparse constraints (NMFsc); d) Locality Preserving Projections (Laplacianfaces). Fingerprint detection was approached by classical minutiae (small graphical patterns) matching through image segmentation by using a structural approach and a neural network as decision block. As to voice / speaker recognition, melodic cepstral and delta delta mel cepstral analysis were used as main methods, in order to construct a supervised speaker-dependent voice recognition system. The final decision (e.g. "accept-reject" for a verification task) is taken by using a majority voting technique applied to the three biometrics. The preliminary results, obtained for medium databases of fingerprints, faces and voice recordings, indicate the feasibility of our study and an overall recognition precision (about 92%) permitting the utilization of our system for a future complex biometric card.

Keywords**—**Biometry, image processing, pattern recognition, speech analysis.

I. Introduction

BIOMETRIC recognition refers to the use of distinctive *physiological* (e.g., fingerprints, face, retina, voice) and *behavioral* (e.g., gait, signature) characteristics, called *biometric identifiers* (or simply biometrics) for automatically recognizing individuals. All biometric identifiers are a combination of physiological and behavioral characteristics. The objectives of biometric recognition are user convenience (e.g., money withdrawal without ATM card or PIN), better security (e.g., difficult to forge access), and higher efficiency (e.g., lower overhead for computer password maintenance).

This work was supported by a grant from the Romanian Ministry of Education and Research, within a CNCSIS grant, contract No. 61GR/16.05.2006.

H. N. Costin and C. Rotariu are with the Faculty of Medical Bioengineering, Univ. of Medicine and Pharmacy, Iasi, Romania, 700098, str. Universitatii 16, (phone: +40-232-213573; fax: +40-232-211820; e-mail: hcostin@gmail.com), and with the Institute for Informatics, Romanian Academy-Iasi Branch.

I. Ciocoiu is with theTechnical University of Iasi, Faculty of Electronics and Telecommunications, B-dul Carol 11A (e-mail: iciocoiu@etc.tuiasi.ro).

T. Barbu is with the Institute for Informatics, Romanian Academy-Iasi Branch, B-dul Carol 11A, Iasi (e-mail: tudbar@iit.tuiasi.ro).

A *biometric system* is essentially a pattern recognition system that recognizes a person by determining the authenticity of a specific feature of that person. A biometric system may be called either a *verification* system or an *identification* system: (a) *a verification system* authenticates a person's identity by comparing the captured biometric characteristic with her own biometric template(s) pre-stored in the system. Such a system either rejects or accepts the submitted claim of identity (*Am I whom I claim I am ?*); (b) *an identification system* recognizes an individual by searching the entire template database for a match. It conducts one-to-many comparisons to establish the identity of the individual, without the subject having to claim an identity (*Who am I ?*). The term *authentication* is also used in the biometric field, sometimes as a synonym for verification.

Depending on the application, the template may be stored in the central database of the biometric system or be recorded on a *magnetic card* or *smartcard* issued to the individual. The verification task is responsible for verifying individuals at the point of access. The representation of biometric characteristic is fed to the feature matcher, which compares it against the template of a single user (retrieved from the system DB based on the user's PIN). In the identification task, no PIN is provided and the system compares the representation of the input biometric against the templates of all the users in the system database. Each biometric has its strengths and weaknesses and the choice typically depends on the application. No single biometric is expected to effectively meet the requirements of all the applications. Thus, a combination of different biometrics is necessary to develop a reliable system for person identification. We have studied three different biometrics for person verification and identification: face, fingerprints and voice.

II. Face Recognition

Most of the approaches of face identification may be classified into two categories [5]:

a) *template-based techniques*, usually performing a projection of the original (high-dimensional) images onto lower dimensional subspaces spanned by specific basis vectors. Examples include Principal Components Analysis (PCA), Linear Discriminant Analysis (LDA), and their kernel-based variants. Eigenfaces [24] represent a *de facto* standard for this approach and still defines a performance reference against which any new method is compared;

b) *geometric feature-based techniques*, relying on the identification of generic components of a face such as eyes, nose, mouth, and distances among them, followed by computation of specific local features. Elastic Graph Matching [24], active shape models [9], and Local Feature Analysis (LFA) [22] belong to this category of tools.

From another taxonomic perspective, we may identify *holistic* and *parts-based* approaches, which extract specific face "signatures" by processing the entire face image or localized portions of it. In principle, parts-based representation offer advantages in terms of stability to local deformations, lighting variations, and partial occlusion.

The present paper presents a comparative analysis of subspace projection methods yielding localized basis functions, against techniques using locality preserving constraints. In this respect, four distinct local feature extraction techniques and a manifold learning strategy are considered: a) Nonnegative Matrix Factorization (NMF) [18]; b) local NMF (LNMF) [19]; c) Independent Component Analysis (ICA) [3]; d) NMF with sparse constraints (NMFsc) [14]; e) Locality Preserving Projections (Laplacianfaces) [13]. We have taken into account the type of distance metric, the dimension of the feature vectors to be used for actual classification, the sources of face variability. We have conducted extensive computer experiments on "AR face database". Test results indicate that the relative ranking of the methods is highly task dependent, and the performances vary significantly upon the distance metric used.

A. Local Feature Extraction Techniques

Existing parts-based representations aim at decomposing a face image into a linear combination of space-localized components (Fig. 1). The individual images form a basis, and the set of coefficients may be interpreted as the face "signature" related to the specific basis. The available N training images are organized as a matrix **X**, where a column consists of the raster-scanned *p* pixel values of a face. We denote by **B** the set of *m* basis vectors, and by **H** the matrix of projected coordinates of data matrix **X** onto basis **B**. If the number of basis vectors is smaller than the length of the image vectors forming **X**, we get dimensionality reduction. In the opposite case we obtain overcomplete representations. As a result, we may write:

$$\mathbf{X} \approx \mathbf{BH} \tag{1}$$

where $\mathbf{X} \in \Re^{pxN}$, $\mathbf{B} \in \Re^{pxm}$, and $\mathbf{H} \in \Re^{mxN}$. Different linear projections techniques impose specific constraints on **B** and/or **H**, and some yield spatially localized basis images.

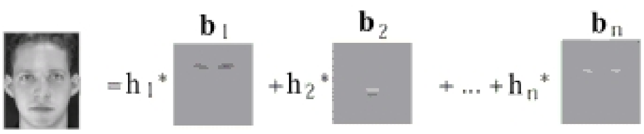

Fig. 1 Principle of parts-based representation of faces

B. Local Non-Negative Matrix Factorization (LNMF)

NMF was recently introduced as a linear projection technique that imposes non-negativity constraints on both **B** and **H** matrices during learning [11]. The method resembles matrix decompositions techniques such as positive matrix factorization [12], and has found many practical applications including chemometric or remote-sensing data analysis. The basic idea is that only *additive* combinations of the basis vectors are allowed. Referring to (1), NMF imposes the following restrictions:

$$\mathbf{B}, \mathbf{H} \geq \mathbf{0} \tag{2}$$

Unlike simulation results reported in [11], the basis images provided by NMF algorithm still maintain a holistic aspect, particularly in case of poorly aligned images, as was previously noted by several authors. In order to improve localization, we used a local version of the algorithm [11] that imposes the following constraints: a) maximum sparsity of coefficients matrix **H**; b) maximum expressiveness of basis vectors **B**; c) maximum orthogonality of **B**. The following equations describe the updating procedure for **B** and **H**:

$$\begin{aligned} H_{aj} &\leftarrow \sqrt{H_{aj} \sum_i \left[\mathbf{B}^T\right]_{ai} \frac{X_{ij}}{\left[\mathbf{BH}\right]_{ij}}} \\ B_{ia} &\leftarrow B_{ia} \sum_j \frac{X_{ij}}{\left[\mathbf{BH}\right]_{ij}} \left[\mathbf{H}^T\right]_{ja} \\ B_{ia} &\leftarrow \frac{B_{ia}}{\sum_j B_{ja}} \end{aligned} \tag{3}$$

C. Independent Components Analysis (ICA)

The redundancy of natural images provides knowledge [3], and that the role of the sensory system is to develop factorial representations in which the dependencies between pixels are separated into statistically independent components. While in PCA and LDA the basis vectors depend only on pairwise relationships among pixels, higher-order statistics are necessary for face recognition. ICA is an example of a method sensible to such statistics: given a set of linear mixtures of several statistically independent components, ICA aims at estimating the mixing matrix based on the assumption of statistical independence of the components.

There are two distinct possibilities to apply ICA to face recognition [3]. The first one organizes the database into a large matrix whereas *every image is a different column*. In this case images are random variables and pixels are outcomes (independent trials). We are interested in independence of images or functions of images. Two *i* and *j* images are independent if when moving across pixels, it is not possible to predict the value taken by the pixel on image *i* based on the value taken by the same pixel on image *j*. This approach yields a set of spatially independent basis images, roughly associated with the components of typical faces such as eyes, nose, and mouth. A systematic analysis has been reported in

[3] in terms of algorithms and architectures used to implement ICA, the number of subspace dimensions, distance metric, and recognition task (facial identity vs. expression). Results indicate that specific ICA design strategies are significantly superior to standard PCA. We have tested ICA using the InfoMax approach (MATLAB code available at http://inc.ucsd.edu/~marni/code.html).

D. NMF with Sparseness Constraints (NMFsc)

A random variable is called *sparse* if its probability density is highly peaked at zero and has heavy tails. Standard NMF does yield a sparse representation of the data, but there is no effective way to control the degree of sparseness. Augmenting standard NMF with the sparsity concept proved useful for dealing with overcomplete representations. Sparsity is taken into account in LNMF and non-negative sparse coding [14]. In fact, the later enables the control over the (relative) sparsity level in **B** and **H** by defining an objective function that combines the goals of minimizing the reconstruction error and maximizing the sparseness level. Yet, the optimal values of the parameters describing the algorithm are to be set by extensive trial-and-error experiments. This shortcoming is eliminated by using the method termed *NMF with sparseness constraints (NMFsc)* [14]. The sparseness degree may be set explicitly through the use of a special projection operator that sets the L_1 and L_2 norms of the basis components. The address http://www.cs.helsinki.fi/patrik.hoyer/ yields the MATLAB code.

E. Locality Preserving Projections (LPP)

Linear projection techniques such as PCA or LDA are unable to accurately approximate data lying on nonlinear submanifolds hidden in the face space. Although several nonlinear solutions to unveil the structure of such manifolds have been proposed (LLE, Laplacean Eigenmaps), these are defined only on the training set data points, and the extension to cover new data remains largely unsolved [4]. An alternative solution is to preserve the *local structure* of the manifold after subspace projection. One such method is Locality Preserving Projections (LPP) [12], that represents a linear approximation of the nonlinear Laplacean Eigenmaps introduced in [4]. It aims at preserving the intrinsic geometry of the data by forcing that neighboring points in the original data space are mapped into closely projected data. In this respect, a special objective function is constructed based on a weighted adjacency graph, including terms that penalize points that are mapped far apart. Basically, the approach finds a minimum eigenvalue solution to a generalized eigenvalue problem (MATLAB code available at http://people.cs.uchicago.edu/~xiaofei). When applied to face image analysis the method yields so-called "Laplacianfaces".

F. Experimental Results

1) Image Database Preprocessing

AR database contains images of 116 individuals (63 males and 53 females). Original images are 768x576 pixels in size with 24-bit color resolution. 13 conditions with varying facial expressions, illumination and occlusion were used. In Fig. 2 we present examples from this database.

As in [11], we used as training images two neutral poses of each person captured in different days (labeled $AR01_1$ and $AR01_2$), while the testing set consists of pairs of images for the remaining 12 conditions, AR02…AR13, respectively. More specifically, images AR02, 03, and 04 are used for testing the performances of the analyzed techniques to deal with expression variation (smile, anger, and scream), images AR05, 06, and 07 are used for illumination variability, and the rest of the images are related to occlusion (eyeglasses and scarf), with variable illumination conditions.

Fig. 2 Example of one individual from the AR face database: (1) neutral, (2) smile, (3) anger, (4) scream, (5) left light on, (6) right light on, (7) both lights on, (8) sunglasses, (9, 10) sunglasses left/right light, (11) scarf, (12, 13) scarf left/right light

2) Comparative Performance Analysis

The performances are given in terms of recognition accuracy, and are compared to results obtained by performing standard PCA. The considered design items are: a) the distance metric used: Euclidean (L2), Manhattan (L1), cos (cosine of the angle between the compared vectors, $\cos(\mathbf{x},\mathbf{y}) = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$) ; b) projection subspace dimension: the dimension of the feature space, equal to the number of basis vectors used, is set to 50, 100, 150, and 200.

In order to make the evaluation, we conducted a rank based analysis as follows: for each image/dimension combination, we ordered the performance rank of each algorithm/distance measure combination (the highest recognition rate got rank 1, and so on) regardless the subspace dimension. This yielded a total of 11 rank numbers for each case: expression variation, illumination variation, glasses, and scarf. Then, we computed a sum of ranks for each of the algorithms over all the cases, and ordered the results (lowest sum indicates best overall performance).

TABLE I
RANK BASED ANALYSIS RESULTS

Algorithm / Distance	Expression rank	Illumination rank	Glasses rank	Scarf rank	Sum of ranks
ICA-COS	13	5	3	6	21
ICA-L2	10	8	9	3	25
ICA-L1	12	7	6	9	27
LNMF-L1	26	18	18	17	61
PCA	14	25	24	39	63
NMFsc-L1	32	10	30	16	65
Laplacean	13	31	21	31	72
NMFsc-COS	19	26	27	31	72
NMFsc-L2	21	29	27	27	77
LNMF-L2	16	40	34	32	93
LNMF-COS	24	34	38	32	128

Some of the conclusions revealed by the rank based analysis results (Table I) are as follows:

- ICA implemented by the InfoMax algorithm seems best suited for the overall face recognition task, outperforming clearly all other solutions;
- methods yielding localized basis images perform better than solutions based on manifold learning;
- PCA compares favorably to most local-based representations and is even better than more sophisticated algorithms such as NMFsc and Laplaceanfaces;
- cosine and L_1 metrics are almost always superior to L_2, and this agrees with previously reported results;
- the dependence of the recognition rates (not shown here) on the projection subspace dimension is not always clear, but larger dimensions tend to be generally favored.

Some important aspects are to be tackled if parts-based approaches become important tools in face recognition. Reliable selection of significant basis vectors is still an open problem, if the number of training images per class is small. Basis vectors exhibiting invariance to common transformations such as translations and rotations would be desirable. Also, identification of the conditions under which correct decompositions of faces into significant / generic parts emerge [8] is a key problem to be further addressed.

III. FINGERPRINT RECOGNITION

In recent years, many efforts have been made in the research of minutiae-based person identification, both for access control and in law enforcement. Minutiae (small graphical characteristic patterns) matching is certainly widely used for fingerprint identification.

The access control uses small and medium databases and, due to the very small computing time needed for decision, the identification is not based on minutiae recognition, but on correlation, directional images, Fourier transform or different distance measures [21]. A much more difficult application refers to large and even huge databases, as those found in policy activity, where the verification must be made only by the perfect resemblance of, e. g., 12 minutiae belonging to two different fingerprint images. The great difficulty mainly consists on the poor quality, both of impresses (from the database) and especially of "on-line" fingerprint patterns (imprints). Our system is based on the *neural network* (*NN*) approach for minutiae recognition [6]-[23], and proves satisfactory results concerning the rate of recognition.

The block diagram of our system appears in Fig. 3. Our approach is to implement a *combined* (*hybrid*) *identification system* in order to increase the recognition rate: *a structural detection subsystem going in parallel with a neural network minutiae classifier*. The first one implies an image gray-levels skeletonization [17], followed by a fingerprint post-processing (purification) block [25], which solves the problem of false minutiae occurring after preprocessing. Then, minutiae extraction by scanning the image via windows is performed. For this task we used a simple structural method [15].

The NN path begins with a local binarization (inside of a 16x16 moving window), necessary for a good and fast feature extraction. As local features, the first 47 *Zernike moments* of the window are chosen [16]. Thus, a good rotation and translation invariance is obtained. The minutiae classification is accomplished by a five-layer NN and back-propagation (BP) algorithm.

The structural classifier path also yields a minutiae chart, where all minutia (true and false) locations are registered in a database. So, the NN classification is performed only onto the windows centered on these locations.

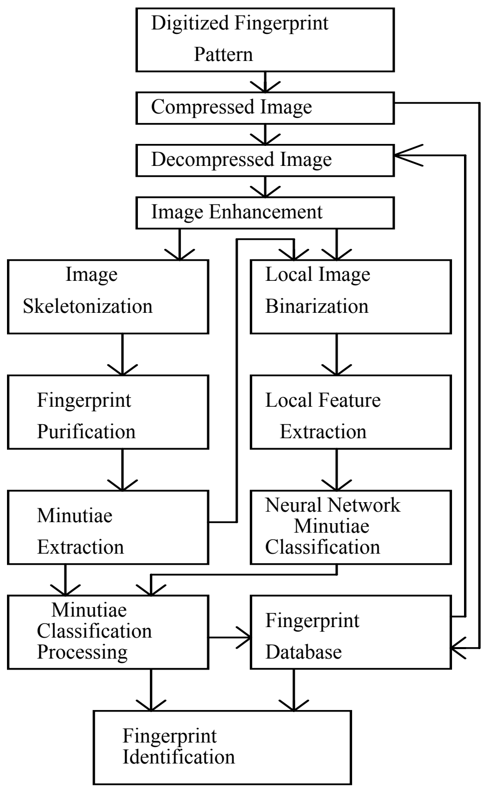

Fig. 3 Our fingerprint identification system

The overall minutiae classification processing consists in comparing the two lists produced by the mentioned classifiers and retaining the first 20 identical patterns, if available (especially for imprints). Moreover, minutiae are registered according to their identifying power that is in the reverse order of a priori average frequency of appearance. The significant minutiae are called: dots, islands, spurs, crossovers, bridges, short ridges, interruption and overlap.

Though the corresponding system is still under development, the experimental results proved its effectiveness. The overall minutiae classification accuracy is about 95% and was obtained on a 500 fingerprints database. Due to the combined minutiae classification approach, the rate of fingerprint identification is about 91% on average, and the average computing time per print is 11 sec.

IV. Supervised Speech-Dependent Speaker Recognition

Speaker recognition (*voice recognition*) methods can be divided into *text-dependent* and *text-independent* techniques [2]-[7]-[10]. Text-dependent recognition are usually based on template-matching techniques. Many of them utilize *dynamic time warping* (DTW) algorithms or *hidden Markov models* (HMM). Our supervised recognition technique is based on *mel-cepstral* (MFCC) sound analysis and contains methods for both identification and verification of the speaker. As usually, it consists of two main parts: *feature extraction* and *feature vector classification.*

A. Speech Feature Extraction

The *Mel Frequency Cepstral Coefficients* (MFCC) are the dominant features used for speech and speaker recognition [20]. Considering a vocal signal *S* to be featured, MFCC speech feature vector extraction uses a short-time analysis of the involved audio signal. Thus, the signal is divided in overlapping frames having the length 256 and overlaps of 128 samples. Then, each resulted segment is windowed, by multiplying it with a Hamming window of length 256. The spectrum of each windowed sequence is then computed, by applying DFT (Discrete Fourier Transform) to it. *Melodic spectrum* is obtained by converting that frequency spectrum on the melodic scale that is described as:

$$mel(f) = 2595 \cdot \log_{10}(1 + f/700), \tag{4}$$

where *f* represents the physical frequencies and *mel*(*f*) the *mel frequencies*. Next, the *mel cepstral* acoustic vector is computed by applying first the logarithm, then the unidimensional DCT (Discrete Cosinus Transform) to the previously obtained mel spectrum. These acoustic vectors can work as feature vectors but we want to obtain more powerful speech features. Therefore, a derivation process is then performed on the MFCC acoustic vectors.

Delta mel cepstral coefficients (DMFCC) are computed as the first order derivatives of mel cepstral coefficients. Then, the *delta delta mel frequency cepstral coefficients* (DDMFCC) are obtained as the second order derivatives of MFCCs. Thus, a set of DDMFCC acoustic vectors result for the initial voice signal. Each of them is composed of 256 samples, but the speech information is codified mainly by the first 12 coefficients. The 12-row DDMFCC acoustic matrix thus created constitutes a powerful speech discriminator, so it can be used as the feature vector of the vocal sound *S*. Let note it as *V*(*S*). The second size of this feature vector, its number of columns, depends always on the length of signal *S*. Because of their different dimensions, these speech feature vectors cannot be compared using linear metrics such as the most known Euclidean distance. For this reason a special nonlinear metric is introduced by us, which is able to compute the distance between different sized matrices having a single common dimension, like the acoustic matrices representing our speech feature vectors [1]. It derives from the Hausdorff metric for sets, described as

$$h(A,B) = \max_{a \in A}\{\min_{b \in B}\{dist(a,b)\}\}, \tag{5}$$

where *dist* is any proper metric between the points of sets *A* and *B* (for example, the Euclidean distance). By further processing (5), we obtain the Hausdorf-based distance

$$d(A,B) = \max\left\{\sup_{1\le k\le p}\inf_{1\le j\le m}\sup_{1\le i\le n}\left|b_{ik} - a_{ij}\right|, \sup_{1\le j\le m}\inf_{1\le k\le p}\sup_{1\le i\le n}\left|b_{ik} - a_{ij}\right|\right\}, \tag{6}$$

where $A = (a_{ij})_{n\times m}$, $B = (b_{ij})_{n\times p}$, and *n* represents their rows number. This metric given by (6) constitutes a satisfactory discriminator between speech feature vectors, therefore it could be successfully used in the next classification process.

B. Speaker Classification and Verification

We provide a *minimum mean distance classification* approach that represents an extended variant of *minimum distance classifier*. A set of registered speakers is established first. Then, a *training set* is obtained as a collection of spoken utterances, provided by these speakers and filtered for noise removal. Each speech signal from the training set constitutes a *vocal prototype*. As our recognition system is a speech-dependent one, each of these signals represents the same speech. The feature vectors obtained from these prototypes make the *feature training set.*

For *N* advised speakers, then the resulted training set is $M = \{M_1,...,M_N\}$, where each $M_i = \{s_1^i,...,s_{n(i)}^i\}$ represents the set of signal prototypes corresponding to the *i*th speaker. For each s_j^i, where $i = \overline{1,N}$, $j = \overline{1,n(i)}$, the previously described vocal feature extraction is then performed, and the obtained sequence $\{\{V(s_1^1),...,V(s_{n(1)}^1)\},...,\{V(s_1^N),...,V(s_{n(N)}^N)\}\}$ represents the *feature training set* of our classifier.

Then, we consider a sequence of input vocal utterances to be classified by speaker criteria, using these *prototype vectors*. Let it be $\{S_1,...,S_n\}$, each signal S_i, for $i = \overline{1,n}$, corresponding to the same spoken text (speech). The feature extraction process is performed on them, and then the feature set $\{V(S_1),...,V(S_n)\}$ is obtained. There must be *N* classes, each of them corresponding to a different registered speaker. Our procedure inserts each input vocal sequence in the class of the closest registered speaker, i.e. the speaker corresponding to the smallest mean distance between the feature vector of the input signal and the prototype vectors of the speaker. Thus, the mean distance between the input S_i and the training subset M_j, related to the *j*-th speaker, is computed as $\frac{\sum_{k=1}^{n(j)} d(V(S_i),V(s_k^j))}{n(j)}$. Therefore, the closest speaker is identified as the p_i-th registered speaker, where

$$p_i = \arg\min_j \frac{\sum_{k=1}^{n(j)} d(V(S_i),V(s_k^j))}{n(j)}, \forall i \in [1,n], \quad (7)$$

where d is the metric given by (7). Obviously, signal S_i has to be inserted in the p_i -th class, where $p_i \in [1, N]$. Of course, it is possible to get $p_x = p_y$ for $x \neq y$.

This classification result, the N classes of vocal utterances, represents also the result of the speaker identification process. For each input speech, the closest advised speaker is thus identified. The next stage of the recognition process, the speaker verification, must decide if that identified speaker is the one who really produced it.

Therefore, a *verification operation* should be performed for each previously obtained speaker class. Let these classes be $C_1,...,C_N$. Among various techniques that might be used, the most used are the *thresholding* methods and we used such an approach too. Thus, we set a threshold value T and then compare the resulted minimum mean distance values with it. Therefore, the following condition has to be tested:

$$\forall i \in [1,N], \forall S \in C_i \mid \frac{\sum_{k=1}^{n(i)} d(V(S),V(s_k^i))}{n(i)} \leq T. \quad (8)$$

If condition (8) becomes true for a voice sequence S and a class C_i, then the utterance S is accepted by the recognition system as an advised vocal input generated by the i-th registered speaker. Otherwise, S is rejected by our system, as being provided by an unadvised speaker. The speaker verification procedure ends when all spoken utterances produced by unregistered system users are rejected. As to the threshold value, we propose an automatic method, considering the overall maximum distance between any two prototype vectors belonging to the same training feature subset, as a threshold value. Consequently, a satisfactory threshold is obtained from the following relation:

$$T = \max_{i \leq N} \max_{k \neq t} \sum_{k=1}^{n(i)} d(V(s_k^i),V(s_t^i)). \quad (9)$$

As results, the main contributions of this approach are as follows: the DDMFCC matriceal representation of the speech feature vectors, the proposed minimum mean distance classifier, the Hausdorff-derived metric and the threshold-based verification technique. Our system produces high recognition rates, around 90%. Therefore, it is able to provide a proper recognition and identification of any human user.

Acknowledgment

This work was supported by a grant from the Romanian Ministry of Education and Research, within a CNCSIS grant, contract No. 61GR/16.05.2006.

References

[1] Barbu, T. Discrete speech recognition using a Hausdorff-based metric. In *Proc. of the 1st Int. Conf. of E-Business and Telecomm. Networks*, *ICETE 2004*, Setubal, Portugal, Vol. 3, (Aug. 2004), pp.363-368.

[2] Barbu, T. *Modelarea Sistemelor Informatice Multimedia* (*Modeling of Multimedia Informatic Systems*). Editura Academiei Române, 2006.

[3] Bartlett, M.S., Movellan, J.R., Sejnowski, T.R.. Face recognition by independent component analysis. *IEEE Trans. on Neural Networks*, 13, (2002), pp. 1450-1464.

[4] Belkin, M., Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15, (2003), 1373–1396.

[5] Brunelli, R., Poggio, T. Face recognition: features versus templates. *IEEE Trans. Pattern Anal. Machine Intell.* 15, (1993), 1042-1052.

[6] Costin, H., *et al*. Fingerprint identification system using neural network approach, în *Advances in Modeling and Analysis*,B,36,No.2,(1995),21-27.

[7] Costin, M., Barbu, T., Zbancioc, M., Grichnik, A. Automatic speaker recognition: inference of new decision tools. *Journal of Fuzzy Systems and Artificial Intelligence*, Romanian Academy, 2005.

[8] Donoho, D., Sodden, V. When does non-negative matrix factorization give a correct decomposition into parts? In: *Proc. Neural Inform. Process. Syst.* (2003).

[9] Edwards, G.J., Taylor, C.J., Cootes, T. Face recognition using the active appearance model. In: *Proc. European Conf. on Computer Vision*, (1998), pp. 581-595.

[10] Furui, S. Speaker-independent isolated word recognition using dynamic features of the speech spectrum, *IEEE Transactions on Acoustic Speech and Signal Processing*, (1986), Vol ASSP-34, No.1, pp. 52-59.

[11] Guillamet, D., Vitrià, J. Classifying faces with non-negative matrix factorization. In: *Proc. 5th Catalan Conf. for Artif. Intell.*, (2002), pp.24-31.

[12] He X., Niyogi, P., 2003. Locality preserving projections. In: *Proc. Neural Inform. Process. Syst.* 16, Canada.

[13] He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.J. Face recognition using Laplacianfaces. *IEEE Trans. P A M I,* 27, (2005), 328–340.

[14] Hoyer, P.O. Non-negative matrix factorization with sparseness constraints. *Journal of Machine Learning Research* 5, (2004), 1457-1469.

[15] Hrechak, A., McHugh, J. Automated fingerprint recognition using structural matching, *Pattern Recognition,* 23, no. 8, (1990), pp. 893-904.

[16] Khotanzad, A., Hong, Y.H. Rotation invariant image recognition using features selected via a systematic method, *Pattern Recognition*, vol. 23, no. 10, (1990), pp. 1089-1101.

[17] Lam, L. *et al.* Thinning methodologies - a comprehensive survey, *IEEE Trans. on PAMI*-14, 9, (1992), pp. 869-885.

[18] Lee, D.D., Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* 401, (1999), pp. 788-791.

[19] Li, S.Z., Hou, X.W., Zhang, H.J. Learning spatially localized, parts-based representation. In: *Proc. IEEE Internat. Conf. on Computer Vision and Pattern Recognition*, (2001).

[20] Logan B. Mel frequency cepstral coefficients for music modelling, In *Proc Int. Symp. on Music Information Retrieval (ISIMIR).* Plymouth, MA, 2000.

[21] Maltoni, D., *et al. Handbook of Fingerprint Recognition*, Springer, (2003).

[22] Penev, P., Atick, J. Local feature analysis: a general statistical theory for object representation. *Network: Comput. in Neural Syst.*,7,(1996),477-500.

[23] Srinivasan, V.S., Murthy, Neural network detection of singular points in fingerprint images, *Pattern Recognition*, 25, no. 2, (1992), pp. 139-153.

[24] Wiskott, L., Fellous, J.M., Kruger, N., v.d. Malsburg, C. Face recognition by elastic graph matching. *IEEE Trans. Pattern Anal. Machine Intell*. 19, (1997), pp. 775-779.

[25] Xiao, Q., Raafat, H. Fingerprint image postprocessing: a combined statistical and structural approach, *Pattern Recognition*, vol. 24, no. 10, (1991), pp. 985-992.