The International Journal Of Engineering And Science (IJES)
|| Volume || 3 || Issue || 6 || Pages || 68-74 || 2014 ||
ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805

Multi-Biometric Person Authentication System Using Speech, Signature And Handwriting Features

Girija M K, Sowmya K S
[1,] Department of Electronics &Communication, Don Bosco Institute of Technology Bangalore,
[2,] Associate Professor, Department of Electronics &Communication, Don Bosco Institute of Technology, Bangalore

----------**ABSTRACT**----------
Biometric Technologies are automated methods for verifying or recognizing the identity of a living person based on physiological or behavioral characteristics. Multimodal Biometric Systems are those which utilize more than one physiological or behavioral characteristic for enrollment, verification or identification. Speaker Recognition is the task of recognizing the speakers using their speech signal. Mel Frequency Cepstral Coefficients (MFCCs) of speaker is determined by extracting and analyzing speaker – specific features from the speech signal. Signature Recognition is the task of recognizing signatories by using their signatures. Features like Horizontal Projection Profile (HPP), Vertical Projection Profile (VPP) and Discrete Cosine Transform (DCT) are determined. Handwriting biometric feature can also be used for person authentication, since handwriting has been trained during initial days of learning of language; it is possible to find more regular and reliable results using this feature. In this project, we develop a Multimodal Biometric System using speech, signature and handwriting features, with the objective of improving performance and robustness. For identification and verification, we use MATLAB 7.14 to determine MFCC, HPP, VPP, DCT coefficients.
Keywords- *Biometrics; Speaker recognition; Signature recognition; Handwriting recognition; Multimodal system.*

Date of Submission: 06 June 2014 Date of Publication: 20 June 2014

I. INTRODUCTION

Establishing the identity of a person is becoming critical in our vastly interconnected society. Questions like "Is she really who she claims to be?", "Is this person authorized to use this facility?" or "Is he in the watchlist posted by the government?" are routinely being posed in a variety of scenarios ranging from issuing a driver's license to gaining entry into a country. In the present era of e-commerce more and more services are being offered over the electronic devices and internet. These include banking, credit card facility, e-shopping, etc. To ensure proper use of these facilities only by the authorized or genuine users and avoid any misuse by the unauthorized or imposter users.

Person authentication scheme is embedded into these services. Currently, person authentication is done mostly using one or more of the following means: text passwords, personal identification numbers, barcodes and identity cards. The merit of these schemes is that they do not change their value with respect to time and also unaffected by the environment in which they are used. The main demerit of them is that they can be easily misused or forgotten.

A time more and more services are being offered over the electronic devices and internet. Hence it becomes unmanageable to keep track of the authentication secrets for different services. The alternative that provides relief from all these demerits is the use of biometric features for person authentication. Any physiological and/or behavioral characteristics of human can be used as biometric feature provided it possesses the following properties: universality, distinctiveness, permanence, collectability, circumvention, acceptability and performance [2].

The need for reliable user authentication techniques has increased in the wake of heightened concerns about security and rapid advancements in networking, communication and mobility. Biometrics, described as the science of recognizing an individual based on her physiological or behavioral traits, is beginning to gain acceptance as a legitimate method for determining an individual's identity. Biometric systems have now been deployed in various commercial, civilian and forensic applications as a means of establishing identity. These systems rely on the evidence of fingerprints, hand geometry, iris, retina, face, hand vein, facial thermo gram, signature, voice, etc. to either validate or determine an identity. This type of attack is especially relevant when

behavioral traits such as signature or voice are used. However, physical traits such as fingerprints are also susceptible to spoof attacks.

A biometric sensor works on the inputs provided by any of the human characteristics and applies an algorithm on the scanned biometric data. This is then compared with, and matched to, a template that has already been created earlier and approved by the user. The most specific and reliable biometric data is obtained from the DNA sequencing of any subject. The matching and comparing process creates a „score" based on how closely the sampled biometric matches with the template already obtained, a match score is known as genuine score If it is a result of matching two samples of a biometric trait of the same user. It is known as an imposter score if it is the result of matching two samples of a biometric trait originating from different users. An imposter score that exceeds the predefined threshold results in a false accept, while a genuine score that falls below the predefined threshold results in a false reject.

The False Accept Rate (FAR) of a biometric system is the fraction of imposter scores exceeding the threshold. Similarly, the False Reject Rate (FRR) of a system is defined as the fraction of genuine scores falling below the threshold. Regulating the value of threshold changes the FRR and the FAR values, but for a given biometric system, it is not possible to decrease both these errors simultaneously. In real-world biometric system, biometric measure is referred in terms of FAR and FRR. The FAR measures the percentage of invalid users who are incorrectly accepted of genuine users and the FRR measures the percentage of valid users rejected as imposters. The Equal Error Rate (EER) refers to the point where the FAR equals the FRR. Lower the value of EER, the more accurate the biometric system.

The present work mainly deals with the implementation of multimodal biometric system employing speech, signature and handwriting as the biometric modalities. This includes feature extraction techniques, modeling techniques and fusion strategy used in biometric system. The organization of the paper is as follows: Section II deals with speaker recognition system, signature recognition system and handwriting recognition system using different feature extraction and modeling techniques, and Section III deals with multimodal biometric person authentication system by combining speaker, signature and handwriting recognition systems using fusion strategy. Section IV provides conclusion.

II. DEVELOPMENT OF UNIMODAL SYSTEMS

1. Speaker Recognition System

Speaker recognition is the task of recognizing the speakers using their voices. Speaker recognition can be either identification or verification depending on whether the goal is to identity the speaker among the group of speaker or verify the identity claim of the speaker. Further, speech from the same text or arbitrary text may be used for recognizing the speakers and accordingly we have text dependent speaker identification and verification approaches. The present work approaches text dependent speaker identification and verification of a speaker through identification. In this work, two different feature extraction and modeling techniques are used for text dependent speaker recognition. The feature extraction techniques are: (1) Mel Frequency Cepstral Coefficients (MFCC) are derived from cepstral analysis of the speech signal, (2) a new feature set, is proposed to capture the spectro-temporal source excitation characteristics embedded in the linear predictive residual of speech signal [5]. The modeling technique is used for modeling the person information from the extracted feature is Vector Quantization (VQ).

i. Feature extraction phase

The speaker information is present both in vocal tract and excitation parameters .The MFCCs represent mainly the vocal tract aspect of speaker information and hence take care of only physiological aspect of speech biometric feature. The vocal tract system can be modeled as a time-varying all-pole filter using segmental analysis. The segmental corresponds to processing of speech as short 10 to 30 milliseconds overlapped 5 to 15 milliseconds windows. The vocal tract system is assumed to be stationary within the window and is modeled as an all-pole filter of order P using linear prediction analysis. The feature vectors that are extracted from smooth spectral representations are cepstral coefficients. In the present work we are using MFCC as feature vectors. The cepstral analysis used for separating the vocal tract parameters and excitation parameters of speech signal s(n). This analysis uses the fundamental property of convolution. The cepstral coefficients (C) are derived by using Fast Fourier Transform (FFT) and Inverse FFT (IFFT) which is given by equation (1).

$$C=\text{real}\,(\text{IFFT}(\log|\text{IFFT}(S(n))|)) \quad (1)$$

Human auditory system does not perceive the spectral components in linear scale, but it will perceive on a nonlinear scale. So we can use the nonlinear scale, Mel frequency scale, to extract the spectral information. The critical band filters are used to compute the MFCC feature vectors by mapping the linear spaced frequency spectrum (f HZ) into nonlinearly spaced frequency spectrum (f Mel) using equation (2).

$$f_{Mel}=2595\log_{10}\left(1+\frac{f_{Hz}}{700}\right) \quad (2)$$

When a speech signal is given as an input to the feature extractor, it will truncate entire speech signal into frames of length 10-30 ms to make it quasi-stationary. Hamming window is used for eliminating the Gibbs oscillations, which occur by truncating the speech signal. But, due to windowing, samples present at the verge of window are weighted with lower values. In order to compensate this, we will try to overlap the frame by 50%. A time-frequency vocal source feature extraction by pitch-synchronous wavelet transform, with which the pitch epochs, as well as their temporal variations within a pitch period and over consecutive periods can be effectively characterized. The wavelet transform of time signal x(t) is given by equation (3).

$$w(a,\tau) = \frac{1}{\sqrt{|a|}} \int_{t} x(t)\psi * \left(\frac{t-\tau}{a}\right) \quad (3)$$

Where, and are the mother wavelet function, scaling (or dilation) parameters and translation parameter respectively. Where $\psi\left(\frac{t-\tau}{a}\right)\frac{1}{\sqrt{a}}$) is named the baby wavelets. It is constructed from the mother wavelet by first, scaling $\psi(t)$ which means to compress or dilate $\psi(t)$ by parameter a and then moving the scaled wavelet to the time position of parameter τ. The compression or dilation of $\psi(t)$ will change the window length of wavelet function, thus changing the frequency resolution. Therefore, the ensemble of $\psi\left(\frac{t-\tau}{a}\right)\frac{1}{\sqrt{a}}$ constitutes the time-frequency building blocks of the wavelet transform. The wavelet transform of discrete time signal x(n) is given by equation (4).

$$w(a,b) = \frac{1}{\sqrt{a}} \sum_{n}^{0} x(n)\psi * \left(\frac{n-b}{a}\right) \quad (4)$$

Where= {2k|k=1,2…,K} and = 1,2…,N, and N is the window length. is the conjugate of the fourth-order Daubechies wavelet basis function $\Psi(n)$. K=4 is selected such that the signal is decomposed into four sub-bands at different octave levels. At a specific sub-band, the time-varying characteristics within the analysis window are measured as parameter b changes. To generate the feature parameters for pattern recognition, the wavelet coefficients with specific scaling parameters are grouped is given by equation (5).

$$W_k = \{w(2k,b)|b = 1,2….N\} \quad (5)$$

where N is the window length. Each Wk is called an octave group. Then WOCOR parameters can be derived by using equation (6).

$$WOCOR_M = \{PW_K(m)|P_{k=1,2..4}^{m=1,2...M}\} \quad (6)$$

where ||.|| denotes two-norm operation. Finally, for a given speech utterance, a sequence of WOCORM feature vectors is obtained by pitch-synchronous analysis of the LP residual signal. Each feature vector consists of 4M components, which are expected to capture useful spectro-temporal characteristics of the residual signal. For each voiced speech portion, a sequence of LP residual signals of 30 ms long is obtained by inverse filtering the speech signal. The neighboring frames are concatenated to get the residual signal, and their amplitude normalized within (-1, 1) to reduce intra-speaker variation. Once the pitch periods estimated, pitch pulses in the residual signal are located. For each pitch pulse, pitch-synchronous wavelet analysis is applied with a Hamming window of two pitch periods long. For the windowed residual signal x(n) the wavelet transform is computed using equation (4).

ii. Training Phase

For speaker recognition, pattern generation is the process of generating speaker specific models with the collected data in the training stage. The mostly used modeling techniques for modeling include vector quantization [14] and Gaussian mixture modeling. The VQ modeling involves clustering the feature vectors into several clusters and representing each cluster by its centroid vector for all the feature comparisons. The GMM modeling involves clustering the feature vectors into several clusters and representing all these clusters using a weighted mixture of several Gaussians. The parameters that include mean, variance and weight associated with each Gaussian are stored as models for all future comparisons. After finding the MFCC feature vectors for the entire frame of the speech signal for the individual speaker, we have to find some of the code vectors for the entire training sequence with less number of code words and having the minimum mean square error. To find minimum mean square error with less number of code words by using VQ, we have two most popular methods namely K-means algorithm and Linde-Buzo and Gray (LBG) algorithms. Vector quantization process is nothing but the idea of rounding towards the nearest integer. The second modeling technique we used in our work, the Gaussian Mixture Modeling (GMM), which is most popular generative model in speaker recognition. The template models, VQ codebooks, can also be regarded as a generative model, although it does not model variations. The pattern matching can be formulated as measuring the probability density of an observation given the Gaussian. The likelihood of an input feature vectors given by a specific GMM is the weighted sum over the likelihoods of the M unimodal Gaussian densities which is given by equation (7).

iii. Testing Phase.

In this phase, feature vectors are generated from the input speech sample with same extraction techniques as in training phase. Pattern matching is the task of calculating the matching scores between the input feature vectors and the given models in recognition. The input features are compared with the claimed speaker pattern and a decision is made to accept or reject the claiming. Testing phase in the person authentication system includes matching and decision logic. The testing speech is also processed in a similar way and matched with the speaker models using Euclidean distance in case of VQ modeling. Hence matching gives a score which represents how well the feature vectors are close to the claimed model. Decision will be taken on the basis of matching score, which depends on the threshold value. For testing the performance of speaker recognition system, we have collected the speech database, sampling frequency of 8 kHz. Figure 1 shows speaker 1 sample speech signal of four sentences, which is collected by using microphone.

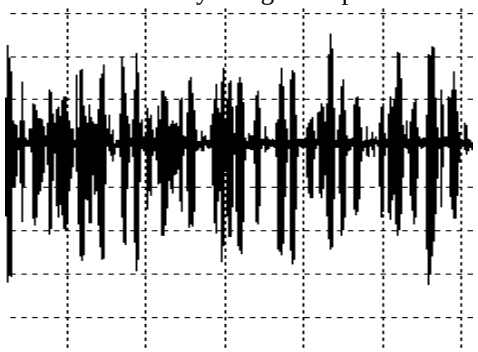

Figure 1 : Sample of speech signals of speaker 1.

2. Signature Recognition system

Signature recognition is the task of recognizing signatories by using their signatures. Signature is a behavioral biometric, the features of signature are variant with respect to time and the forgers can easily fool the system by reproducing the signatures of the correct persons. Irrespective of the above limitations we can still use signature as our best biometric feature, since the signature is a unique identity of an individual and is being used extensively in practical systems. No two signatures can be identical, unless one of them is a forgery or copy of the other . The signature recognition systems find applications in government, legal and commercial areas. Signature verification is the verification of given signature of claimed identity of a person. There are two types of signature verification systems in practice, namely, online and offline [13], [14]. Online signature verification uses information collected dynamically at the time of signature acquisition like timing, acceleration, velocity, pressure intensity and also termed as dynamic signature verification. Offline signature verification uses only the scanned image of signature and also termed as static signature verification.

i. Feature extraction phase

Feature extraction plays a very important role in offline signature verification. Unlike our speaker recognition case, we are not going model the feature vectors up to some codebook level. Here feature vectors itself will give the training sequence. In this work the features of signature are extracted by using Discrete Cosine Transform (DCT) analysis, Vertical Projection Profile (VPP) analysis and Horizontal Projection Profile (HPP) analysis. The VPP and HPP are static features of a signature and DCT is a global feature of a signature image. Since our signature is an image, it will have the gray levels from 0 to 255 and to compute the maximum gray level the histograms of all images are used. VPP and HPP are the kind of histograms. VPP gives the horizontal starting and ending points and HPP gives the vertical starting and ending points of the image. The size of VPP and HPP is equal to the number of columns and the number of rows in the signature image respectively. Since, the size of signature regions are not constant even for a single user, in this work we are taking average value of vertical projection profile as a feature. The signature image intensity A (p, q) at pth row and qth column indices respectively. Where M is number of rows in an image and N is number of columns in image. Equation (11) gives the DCT coefficient corresponding to pth row and qth column of an input signature image. The performance of signature recognition system depends on the way in which the DCT coefficients are considered. The zonal coding of DCT coefficients of signature image are used for better performance, which gives concentration at low spatial frequencies.

ii. Testing Phase

For the identification or verification, same set of features which have been extracted during registration process are extracted from the input samples scanned or recorded using input devices like writing pads, to form the feature vectors. Verification is 1 to 1 matching while identification is 1 to n matching. In verification, the individual claims his/her identity which is verified by comparing these features vectors by the feature vectors of the individual which he/she claimed to be. If the matching score crosses the predefined threshold then the system verifies the individual as authentic user. In identification, the feature vectors of the individual are compared with the feature vectors of every individual stored in the database. If the highest matching score crosses the predefined threshold, then it identifies the individual as the person whose matching score is the highest otherwise the system suggest few top most matches. The matching algorithm is needed to compare the samples and computes the matching score and decide if two samples belong to the same individual or not by comparing the matching score against the acceptance threshold.

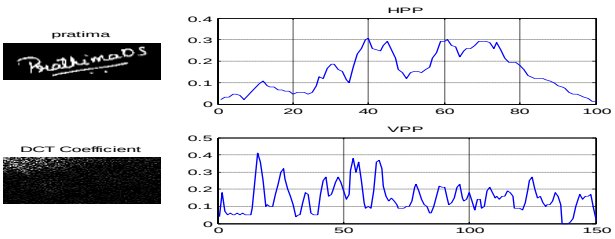

Figure 2 Sample Signature of user 1

During the training session, we considered the signatures of each writer and extract the features from those signatures by using VPP, HPP and DCT analysis. The three feature models are obtained for all 30 users. In testing phase, we have used the remaining 8 signatures for each writer. For the given test signature, we have to extracted the VPP, HPP values and DCT coefficients separately by VPP-HPP-DCT analysis. After getting these values, we found the minimum distance between the VPP-HPP-DCT values and the feature vectors of all the writers corresponding to each of the model. To improve the performance of the signature recognition system, along with the baseline VPP-HPP system the DCT coefficients are used. The two sequences are compared with some distance measures like Euclidean distance at each and every point, so as to obtain the distance matrix. These distances in the matrix are termed as local distances. Let the Matrix be D and the sequences are A, B with lengths M, N respectively. Then D is calculated using equation (12).

D (i, j) = distance (A(i), B(j)) (12)

where i varies from 1 to M and j varies from 1 to N. The distance here considered is Euclidean distance. The modified feature vectors obtained from the signature image A (i, j) of size MxN are given in the equations (13) and (14). Calculate the DTW distance values separately for VPP and HPP vectors from all the users for all the training images to the testing image and obtain distances from each user using average distance method.

3. HANDWRITING RECOGNITION SYSTEM

Handwriting biometric feature can also be used for person authentication. Most of the existing works on handwriting information is for forensic investigation. The scope includes identifying the author of the given handwritten script from the group of available large population. The end result may be a subgroup of most likely population. This subgroup may then be carefully analyzed by the human experts to identity the correct person who might have written the script. Thus using handwriting information in criminal investigation is an age old method. Handwriting biometric feature may also possess several characteristics to qualify it for use in person authentication. Relatively few works have been done in this direction . With the integration of pen-based input devices in PDA and Tablet PCs strongly advocates the use of handwriting information for person authentication due to ease of collection. Handwriting verification can also be done either in online or offline mode as in signature verification. Online handwriting verification exploits similar dynamic features as in signature verification. Thus it is easy to extend the online signature verification approach to handwriting verification. Initially an online handwriting verification system will be developed. However, it should be noted that there is a significant difference between signature and handwriting. Signature is one pattern from hand, but it will not use any language specific information.

i. Feature extraction phase

Mainly dynamic time warping in context of images is used for word matching which uses vectors like normalized upper word and lower profile, back ground ink transitions etc.,. The features from the handwriting image considered in our work are VPP vector and HPP vector. The VPP is an array that contains sum of gray levels of each column in a handwriting image. This feature signifies the variations of Gray level distribution along the length of the image. This VPP vector is a unique feature for a given user and will vary from user to user. Even the same user will have variations. The important and the uniqueness of the information present in the HPP vectors are equally important as that of VPP vectors. So along with the VPP vector extraction, another feature HPP vector is obtained from the handwriting image.

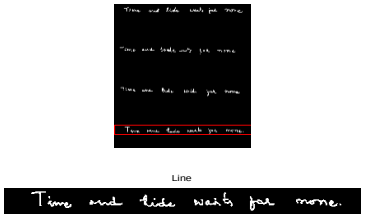

Figure 3 Sample of handwriting of user1

ii. Testing Phase

Writer recognition system is built using the individual words, segmented from the sentence considered for handwriting and combined later for better performance. In order to obtain a correct segmentation, a threshold is calculated that distinguish words and characters. After obtaining the threshold, words are segmented by obtaining the VPP vector and examining its intensity profile. The each word extracted from the sentence now act as the images to be tested. The algorithm proposed for
a full sentence is applied for each word.

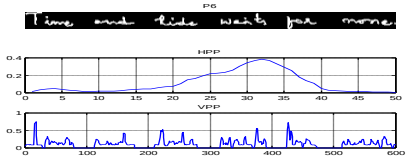

Figure 4 tested handwriting

Consider one word and apply already proposed algorithm for all sentence. Obtain the DTW distances from each user by averaging method. Next, normalize the distances and repeat the same procedure for all the words. The normalized distances are fused using the fusion principle. Obtain minimum distance and its corresponding user there by identifying the user. At the fusion level, distances are fused using sum rule. The similar procedure is used for finding the HPP vectors reason is that, the richest image information obtained from Gray level distribution along the length of the handwriting image. The HPP vectors are equally important as that of VPP vectors which gives the information about the variations of the handwriting along the lateral extent. The combined feature gives the complete behavior of handwriting image of a user, hence VPP-HPP based system is one the unimodal system in our multimodal biometric person authentication.

III. MULTIMODAL BIOMETRIC PERSON AUTHENTICATION SYSTEM

A. Development of Multimodal System

Multimodal or Multi-biometric systems, remove some of the drawbacks of the unimodal systems by grouping the multiple sources of information. These systems utilize more than one physiological or behavioral characteristic for enrollment and identification. Once the unimodal systems are developed, then the next step is to develop multimodal system by integrating them suitably. The unimodal systems using speech, signature and handwriting information are ranked according to their performance. Based on this, the best performing system is used as the baseline system to which other systems are integrated. The integration can be done at any of the following three levels: feature, measurement and score levels [2]. A tight integration is possible if it is done at the feature level them suitably and combine them. Alternatively, the features can be applied to one more level of smoothing using feature modeling techniques to obtain modified features that are similar for both biometric features. These features are then used for modeling. This will result in the development of the multimodal biometric person authentication system using all the possible biometric features.

B. Performance and Robustness Evaluation

There are standard databases for the individual evaluation of the unimodal biometric systems, like YOHO database, IITG database etc... However, such an evaluation is only for finding the performance of the particular unimodal system in an absolute sense. To have comparative study evaluate the strength of multimodal system on common platform, it is proposed to develop a multimodal database for these three biometric features. For this reason we have prepared our own database of 6 users. The database consists of 3 samples of speech information, 6 samples of signature and 6 samples of handwriting for each user. Once the database is developed, then the performance is evaluated first for each of the unimodal systems. The performance is then evaluated for multimodal system using all the three features. Such evaluation provides a systematic comparison between unimodal and multimodal systems. The main features considered in developing a multimodal system are handwriting, signature and speech. The following are the steps involved in the implementation of multimodal biometric person authentication system based on unimodal system performance.

a) Collect the individual matching scores of the unimodal systems for every user.
b) Normalize the matching scores using normalization techniques and integrate the scores by using fusion rules.
c) Assign the multiple biometric to a particular person who produces the minimum score.

IV. CONCLUSION

The trend of multimodal biometrics is spreading for the authentication process to maintain the interests regarding the security as strong as possible. The vital features that encourage the use of multimodal biometrics are the performance and accuracy along with the ability to overweigh the drawbacks of unimodal biometric systems. In this work we demonstrated multimodal biometric person authentication system using three biometric features. We generated our own database of 6 users and effectively using the principle of matching score fusion and normalization technique for developing multimodal system. Further, we combined the multimodal systems using normalization and fusion techniques. This system gives the identification performance is 100% and the verification performance is 0%, in terms of FAR 0% and FRR is 0%. As a result, we implemented multimodal biometric person authentication system using speech, signature and handwriting features which provides 0% error rates.

REFERENCES

[1] L.Gorman, "Comparing passwords, tokens and biometrics for user authentication," IEEE Proc., vol. 91, no.12, Dec. 2003.
[2] A.K. Jain, A Ross and S. Prabhaker, "An introduction to biometric recognition," IEEE Trans. Circuits and Systems for Video Technology, vol. 14, no. 1, pp, 4-20, Jan. 2004.
[3] R. Bruneelli and D.Falavigna, "Person identificationusing multiple cues," IEEE Trans. PAMI, vol. 17, no.10, pp.955-966, oct.1995.
[4] L. Hong and A.K. Jain, "Integrating faces and fingerprints for person identification," IEEE PAMI, vol. 20, no. 12, pp. 1295-1307, Dec. 1998.
[5] V. Ghattis, A.G. Bors and I. Pitas, " Multimodal decision level fusion for person authentication," IEEE Trans. Systems, Man and Cybernatics, vol. 29, no. 6, pp, 674-680, Nov. 1999.
[6] R. W.. Frischholz and U. dieckmann, "Bioid: A multimodal biometric identification system," IEEE Computer Society, pp. 64-68, Feb.2000. Name Stand. Abbrev., in press.
[7] A. Kumar et. al., "Person verification using palmprint and handgeometry biometric," proc. Fourth Int. Conf. AVBPA, pp.668-678, 2003.
[8] S.Ribaric, D. Ribaric and N. Pavesic, "Multimodal biometric user identification system for network based applications," IEEE Proc. Vision, Image and Signal Processing, vol. 150, no.6, pp.409-416, 2003.
[9] A.K. Jain and Ross, "Learning user specific parameters in multibiometric system," Proc. Int. Conf. Image Processing (ICIP), pp. 57-60,2002.
[10] A. K.Jain, L.Hong and Y. Kulkarni, " A multimodal biometric system using fingerprint, face and speech," Proc. Second Int. Conf. AVBPA, pp.182-187, 1999.
[11] S.Ribaric, I. Fratic and K. Kris, " A biometric verification system based on the fusion of palmprint and face features," Proc. Fourth Int. Symposium Image and Signal Processing, pp. 12-17, 2005.
[12] B.Duc et. al., "Fusion of audio and video information for multimodal person authentication," Pattern Recogntion Letters, vol. 18, pp.835-845, 1997.
[13] S. Furu, "Cepstral analysis techniques for automatic speaker verification," IEEE Trans, Acoust., Speech Signal Processing, vol. 29(2), pp.254-272, 1981.
[14] F.K.Soong, A.E.Rosenberg, L.R. Rabiner and B.H. Jvang," A Vector quantization approach to speaker recognition," Proc., IEEE, Int., Conf., Acoust., Speech Signal Processing, vol. 10, pp.387-390, Apr.1985.