2010 International Conference on E-Health Networking, Digital Ecosystems and Technologies

A New Fingerprint Image Recognition Approach Using Artificial Neural Network

Zou Lihua
Computer Engineering, Dongguan Polytechnic, Dongguan ,China

***Abstract*—In this article, the application algorithm in fingerprint image recognition based on artificial neural network and wavelet transform is put forward. In this algorithm, we have constructed a flatten-structure element utilizing the low-frequency parameters, which can affect the high-frequency parameters with the flatten-structure element. In this way we can enhance the fingerprint image with keeping the global texture feature of fingerprint. The experiment results indicate that this increase effect of image-pretreatment in fingerprint image recognition is good and the algorithm is effective.**

Keywords- Fingerprint Image Recognition ; Artificial Neural Network; Wavelet Transform

I. INTRODUCTION

The fingerprint recognition problem can be grouped into two sub-domains; fingerprint verification and fingerprint identification. In addition, different from the manual approach for fingerprint recognition by experts, the fingerprint recognition here is referred as AFRS (Automatic Fingerprint Recognition System), which is program-based. Fingerprint verification is to verify the authenticity of one person by his fingerprint. The user provides his fingerprint together with his identity information like his ID number. The fingerprint verification system retrieves the fingerprint template according to the ID number and matches the template with the real-time acquired fingerprint from the user. Usually it is the underlying design principle of AFAS (Automatic Fingerprint Authentication System).

In our algorithm feature extraction from MFCC composite with wavelet transform of the image will assist in achieving a higher recognition rate. The Neural Network (NN) classification technique is used in the proposed algorithm. The comparison between the use of the wavelet transform with MFCCs and the MFCCs only is presented in the paper. The rest of the paper is organized as follows. Section 2 gives an overview on the structure of identification system. Section 3 discusses the process of feature extraction. Feature matching is discussed in section 4. In Section 5, the proposed speaker identification method is introduced. Section 6 gives the experimental results. Finally, Section 7 summarizes the concluding remarks.

II. RECOGNITION SYSTEM

The process of performing this method of identification consists of two phases: a training or enrollment phase followed by a testing or evaluation phase. During the training phase, each signal in the set is modeled using a set of training data. Features are extracted from the training data essentially stripping away all unnecessary information in the training signals samples leaving only the characteristic information with which signals models can be constructed. In the testing phase, a sample of signals, which converted from an image with unknown person, is subjected to feature extraction and the resulting information is compared to the models in the fingerprint image database allowing the unknown one to be identified. It is clear that the feature extraction process (obtained from discriminatory information) and classification process (using the features to determine the correct signal. A new fingerprint Recognition system is composed feature extraction and feature matching for the purpose of classification as shown in Fig. (1). This operates in two modes; training and recognition modes, as indicated above. Both of them include feature extraction, sometimes called the front end of the system. The feature extractor converts

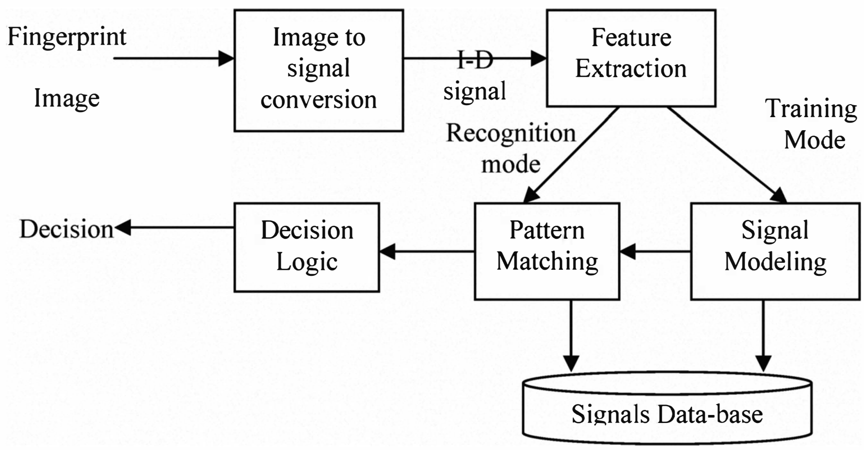

Figure 1. components of automatic identification system

the digital 1-D signal into a sequence of numerical descriptor called feature vectors. Several features extraction techniques used in signal recognition system such linear prediction coefficients (LPC), linear predictive cepsral coefficients (LPCC), perceptual linear predictive analysis (PLP), Mel -Frequency cepstrum coefficients (MFCC), MFCC is the most popular and will be introduced in this paper. Classification is a process which has two phases: image modeling and pattern matching. For successful classification, each image is modeled using a set of data samples in the training mode, from which a set of feature vectors is generated and saved in a database. Features are extracted from the training data essentially striping away all unnecessary information in the training samples leaving only the characteristic information with which image models can be constructed. When a sample of data from some unknown fingerprint arrives, pattern matching techniques are used to map the features from the input sample to a model corresponding to a known fingerprint.

III. Extraction Of Polynomial Coeffeicients With Mel_ Frequency Cepestral Coeffients

The Mel is a unit of measure of perceived pitch or frequency of a tone. The Mel- scale is therefore a mapping between the real frequency scale (Hz) and the perceived frequency scale (Mels). The Mapping is virtually linear below 1 KHz and logarithmic above as given in the following equation:

$$f_{mel} = 2595\log_{10}(1+\frac{f_{linear}}{700}) \qquad (1)$$

MFCC is based on the short term analysis, and thus for each frame a MFCC vector is computed.

A. Windowing and framing

For feature extraction to take place the signal must first be broken up into small sections each of N samples. These sections are called frames. Where the number of samples per frame N will depend on the sampling rate of the data. To avoid a loss of information frame overlap is used: each frame begins at some offset of L samples with respect to the previous frame where L $\leq$ N.

For each frame a windowing function is usually applied to increase the continuity between adjacent frames. Windowing in time domain is a point wise multiplication of the frame and the window function. According to the convolution theorem, this corresponding to convolution of the short term spectrum with the window function response. A good window function has a narrow main lobe and low side lobe levels in their transfer functions, so the most commonly used window function is the hamming window, which defined as:

$$w(n) = 0.54 - 0.46\cos(\frac{2n\pi}{N-1}) \qquad (2)$$

where n=0, 1...N-1 and N is size of the window or frame.

When a data is framed and windowed, the data at the ends of the frame is much likely to be reduced to zero. This will represent a loss of information. An approach to tackle this problem is to allow overlapping in the sections between frames. Overlapping will allow adjacent frames to include portions of data in the current frame. This will mean the edges of the current frame will be included as the center data of adjacent frames. Typically, around 60% of overlapping is sufficient to embrace the lost information .

B. Discrete Fourier Transform (DFT)

The Discrete Fourier Transform (DFT) of the frame is computed to obtain the magnitude spectrum. The DFT is mathematically defined as:

$$S[k] = \sum_{n=0}^{N-1} s(n) e^{-j2\pi\frac{k}{N}n} \qquad (3)$$

The inverse transformation is defined as:

$$s(n) = \frac{1}{N}\sum_{k=0}^{N-1} X[k] e^{j2\pi\frac{k}{N}n} \; Where\, n = 0,1,...N-1 \qquad (4)$$

Fourier Transform theory is based on the idea that any signal can be represented as the sum of properly chosen sinusoidal waves. This is of profound importance in digital signal processing since the transform effectively decomposes a signal into its component frequencies and their amplitudes.

C. Discrete Cosine Transform

The last stage involves performing a Discrete Cosine Transform on the log of the Mel- spectrum. This replaces the inverse DFT stage in practice for increasing computational efficiency. The output of the kth Mel- filter is the sk, the Mel Frequency Cepstral Coefficients (MFCC) are given as:

$$c(n) = \sqrt{\frac{2}{N}}\sum_{k=1}^{N} \log(s_k)\cos(\frac{n}{N}(k-0.5)) \qquad (5)$$

where n=0,1,...M-1 , N is the number of filters , M is the number coefficients and c(n) are the Mel- frequency cepstral coefficients . The number of resulting Mel-frequency cepstral coefficients are chosen between 12 and 20, since most of the signal information is represented by the first few coefficients. The 0th coefficient is usually dropped out because it represents the average log energy of the frame and carries only little speaker specific information.

To use the polynomial coefficients, Time functions of the cepstral coefficients are expanded by orthogonal polynomial representation over 90ms intervals every 10ms. The 90ms interval length seemed adequate for preserving transitional information between phonemes. The first three orthogonal polynomials are used:

P0j =1 , P1j = j-5 , P2j = j2 - 10 j +55/3

Thus, if the control function samples for an utterance within the segment being measured are X_j(j=1,2,.....,9), then the first three coefficients of the orthogonal polynomial representation are:

$$a = \sum X_j / 9 \qquad (6)$$

$$b = \left(\sum_{j=1}^{9} X_j P_{1j}\right) / \sum_{j=1}^{9} P_{1j}^2 \qquad (7)$$

$$c = \left(\sum_{j=1}^{9} X_j P_{2j}\right) / \sum_{j=1}^{9} P_{2j}^2 \qquad (8)$$

These coefficients represent means value, slope, and curvature of the time function of each cepstrum coefficient in each segment, respectively.

IV. Pattern Matching Using Artificial Neural Network (ANN):

Neural Networks are widely used for feature matching. The classification step in automatic identification systems is in fact a feature matching process between the features of a new fingerprint images and the features saved in the database. Multi-layer perceptrons (MLPs) are a popular type of neural network and consist of an input layer, one or more hidden layers and one output layer as shown in Fig.2. The structure shown will be used for feature matching with our algorithm. Processing elements or neurons in the input layer only act as buffers for distributing the input signal x_i to neurons in the hidden layer each neuron j in the hidden layer sums up its input signals xi after weighting them with the strengths of the respective connections w_{ji} from the input layer and computes its output y_j as a function f of the sum

$$Y_j = f(\sum w_{ji} x_i) \tag{9}$$

Training a network consists of adjusting its weights using a training algorithm. The training algorithms adopted the weights by attempting to minimize the sum of squared difference between the desired and actual values of the output neurons:

$$E = \frac{1}{2} \sum_f (Ydj - Yj)^2 \tag{10}$$

Where Y_{dj} is the desired value of the output neuron j and Y_j is the actual output of that neuron. Each weight w_{ji} is adjusted by adding an increment. W_{ji} is selected to reduce E as rapidly as possible.

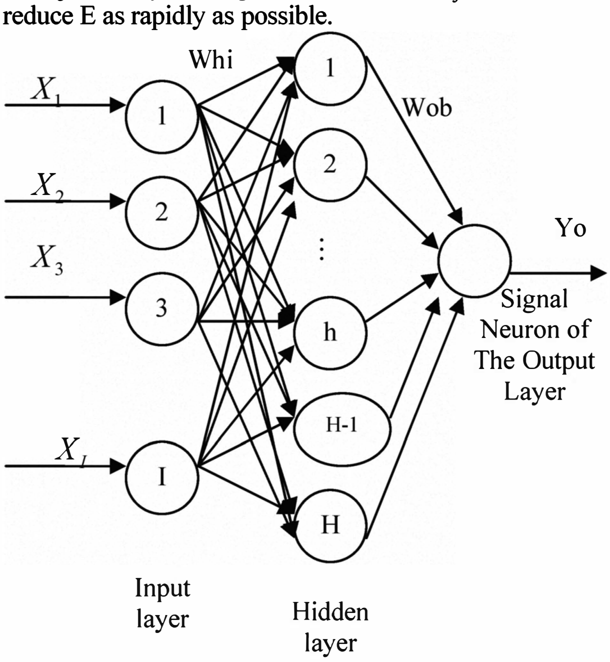

Figure 2. A multilayer perceptron neutral network.

V. The Proposed Fingerprit Image Recognition Mrthod

The proposed algorithm is illustrated in Fig.3. The image, which converted to 1-D signal, is used to extract the MFCC features. In addition, the wavelet transformed signal is used to extract additional MFCC features which can be used to assist the MFCC features extracted from the original degraded signal. Wavelet denoising can be used in the cases of low SNR values.

A. Discrete wavelet Transform (DWT):

The Discrete Wavelet Transform (DWT) is a very popular tool for the analysis of non-stationary signals. The idea of it is to represent a signal as a series of approximations (low pass version) to the signal and details(high pass version) at different resolutions.

B. Wavelet –Based De-noising:

The noise robustness can be improved by the wavelet based de-noising. The following steps are applied to get the wavelet de-noising.

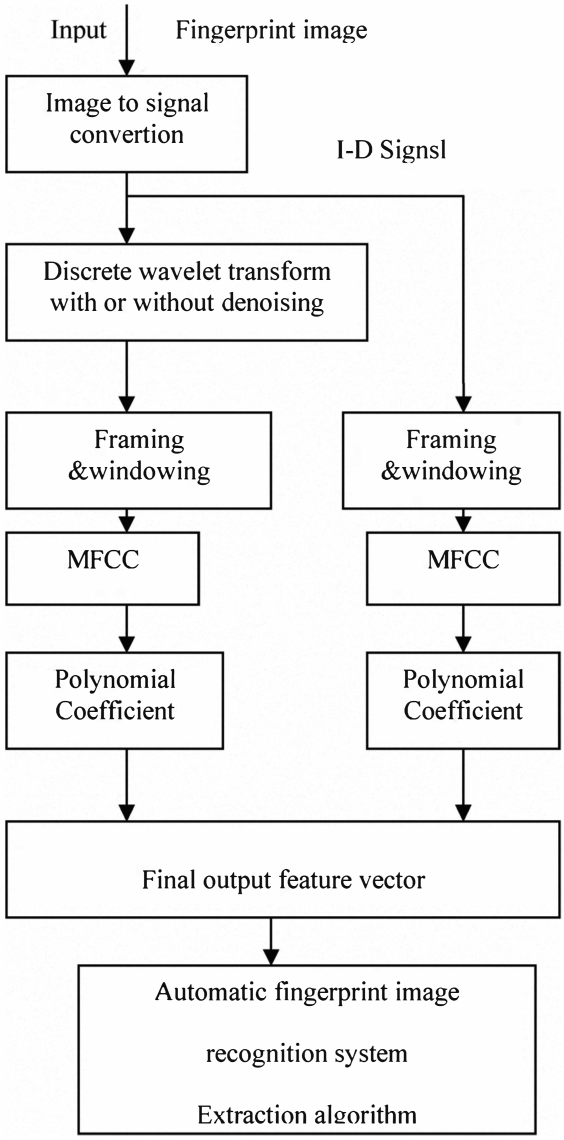

Figure 3. wavelet de-noising

Mother wavelet should be selected carefully to better approximate and capture the transient spikes of the original signal,. The mother wavelet will not only determined how the original signal be estimated in terms of the shape of the PD spikes, but also, it will affect the frequency spectrum of the de-noised signal. The choice of mother wavelet can based on correlation γ (4.3) between the signal of interest and the wavelet de-noised signal

$$\gamma = \frac{\sum (X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum (X-\bar{X})^2 (Y-\bar{Y})^2}} \tag{11}$$

Where $\bar{X}$ and $\bar{Y}$ are the mean value of set X and Y, respectively.Or based on the cumulative energy over some interval where PD spikes occur:

$$E = \sum X^2 \tag{12}$$

Where E is the energy and X is the signal vector. Two rules are generally used for thresholding the wavelet coefficients (soft/ hard thresholding). Hard thresholding sets zeros for all wavelet coefficients whose absolute value less than the specified threshold limit. The hard thresholding provides an improved signal to noise ratio. The equation of the hard thresholding is:

$$f_{hard} = \begin{cases} x & |x| \geq T \\ 0 & |x| < T \end{cases} \tag{13}$$

On the other hand, that of soft thresholding is given by:

$$f_{soft}(x) = \begin{cases} x & |x| \geq T \\ 2x - T & T/2 \leq x \leq T \\ T + 2x & -T < x \leq -T/2 \\ 0 & |x| < T/2 \end{cases} \tag{14}$$

where T denotes the threshold value and x represents the detail coefficients of the DWT.

VI. EXPREMENTAL RESULTS AND ANALYSIS

In the training phase of the automatic fingerprint identification system, a database is first composed. All fingerprint images in the database are transformed into 1-D signals. Thus, these signals are used to generate MFCCs and polynomial coefficients to form the feature vectors of the database. In the testing phase, each one of these fingerprints is degraded by noise and features are extracted from the degraded images.

The features used in all experiments are 13 MFCCs and 26 polynomial coefficients forming feature vectors of 39 coefficients for each frame of the fingerprint signals. Five methods for extracting these features are adopted in the paper. In the first method, the MFCCs and the polynomial coefficients are extracted from the fingerprint signals only. In the second one, the features are extracted from the DWT of the fingerprint signals. In the third method, the features are extracted from both the original fingerprint signals and the DWT of these signals and concatenated in a single feature vector. In the fourth method, denoising is applied to the noisy signals in the testing phase only to reduce noise prior to feature extraction from the fingerprint signals. In the last method, denoising is applied and features are extracted from both the denoised signals and the DWT of these denoised signals. The results of these experiments are given in Figs.(4) to (5) .

VII. CONCLUSION

This paper presented a robust fingerprint identification method based on MFCCs and the wavelet transform. The experiments illustrate that in the presence of noise, fingerprint identification can be accomplished by the use of MFCCs from the 1-D fingerprint signals and the wavelet transform of these signals. This approach for fingerprint recognition as 1-D signals has avoided the tedious tasks of geometrical feature extraction from fingerprint images.

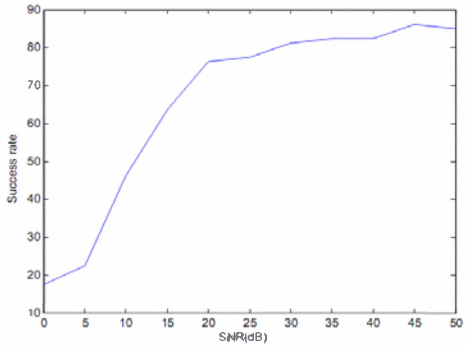

Figure 4. Recognition rates for features extracted from clean fingerprin ts in train and noisy fingerprints in test.

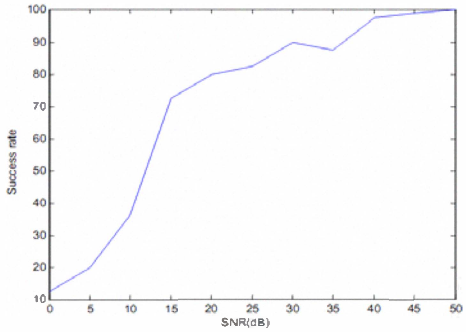

Figure 5. Recognition rates for features extracted from the wavelet transformed clean fingerprints in train and noisy fingerprints in test.

REFRENCES

[1] A.K. Jain, L. Hong and R. Bolle, "On-line Fingerprint Verification," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 19, No. 4, pp. 302-314, 1997.

[2] B. Moayer, K. Fu, "A tree system approach for fingerprint pattern recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, 8(3)(1986) 376-388.

[3] L. Hong, A. Jain, S. Pankanti, and R. Bolle," Fingerprint enhancement," Proceedings 3rd IEEE Workshop on Applications of Computer Vision, pp. 202-207, 1996.

[4] Sch. of Eng., Blekinge Inst. of Technol., Ronneby, Sweden;, " The successive mean quantization transform, " Proceedings on Acoustics, Speech, and Signal Processing, (ICASSP '05). IEEE International Conference ,Vol. 4, pp. iv/429- iv/432 ,2008

[5] B. C. Jong, "Wavelet Transform Approach For Adaptive Filtering With Application To Fuzzy Neural Network Based Speech Recognition", PhD Dissertation, Wayne State University, 2009.