Extracting and Classifying for Ear Recognition in Biometrics Knowledge

Myat Thu Zar
University of Computer Studies, Yangon
myatthuzar86@gmail.com

Abstract

Ear detection is an important part of an ear recognition system. This paper proposes ear recognition based on Gabor wavelets and Support Vector Machine (SVM). The framework has three steps. In the first step, the ear is detected from an image of the face. In the second step, Gabor wavelets are used to extract ear feature. The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells, exhibit desirable characteristics of spatial locality and orientation selectivity. In the third step, when the Gabor wavelets features were obtained, classifications were done by SVM. Research of ear recognition and its application is a new subject in the field of authentication. Ear normalization and alignment is a fundamental module in the ear recognition system.

Key words: Ear recognition, Gabor wavelet, support vector machine, multi-classification

1. Introduction

Recent years, biometrics has a rapid development because of its broad applications ranged from identification to security. Fingerprints, irises, faces, gait and speech used as popular biometrics. Because of ideal properties, such as university, uniqueness, permanence and so on, ear become a new class of biometrics[1].Comparing to other population biometrics, ear has not received much attention though it has advantages over other biometrics. For example, ear is rich in features; it is a stable structure which doesn't change with the age; it doesn't change its shape with facial expressions, cosmetics and hair styles. [4], [5].

Biometrics identification methods proved to be very efficient, more natural and easy for users than traditional methods of human identification. In fact, only biometrics methods truly identify humans, not keys and cards they posses or passwords they should remember. The future of biometrics surely leads to systems based on image analysis as the data acquisition is very simple and requires only cameras, scanners or sensors.

In this paper, a framework for ear recognition based on Gabor wavelets and support vector machine was proposed, the framework contains three steps: ear detection, ear feature extraction and ear recognition.

The rest of paper is organized as follows. Section 2 introduces the related work. Section 3 describes our approach to detect ear, Gabor feature extraction and classification of support vector machine. Section 4 gives the experiment results to demonstrate the effectiveness of approached system. Section 5 provides the conclusions and future work.

2. Related Work

Although there are many papers presented about ear recognition in some area, in this section, we only review ear recognition based on Gabor wavelets and Support Vector Machine (SVM) [7]. Chen and Bhanu [9] presented a template matching based detection method for extracting ears from side face range images. The model template is represented by an average histogram of shape index of ears. However this method cannot identify the ear region accurately.

Chang used standard PCA algorithm for ear recognition, and gets the conclusion that ear and face does not have much difference on recognition rate [8]. For ear recognition in 3D, Hui Chen [3] proposed an ICP based 3D ear recognition system. Pin Yan [10] proposed an automatic ear extraction, and ICP based ear recognition. Yuizono et al. (2002) proposed a new framework of recognition by using genetic local search. Localized orientation information is used by Kumar and Wu (2011) as a new feature extraction approach to identify ears.

Wang used a high order moment-invariant method to extract ear features [5]. Moreno used feature points of outer ear contour and information obtained from ear shape and wrinkles for ear recognition [2].

3. Ear Detection

Before ear feature extraction and ear recognition, ear should be detected from a person's side face. In order to extract an image contains only the ear, there are three steps. First skin-tone detection is used to detect a person's side face containing the ear. Second, contour extra extraction is applied to the skin region and removes short and isolate edges. Third, the ear is located by segmentation from other skin region. Then the detected ear is normalized to 45× 80. Figure 1 depicts the ear detection from a face. And the procedure of ear recognition is shown as Figure 2.

3.1 Gabor Feature Extraction

The Gabor wavelets are similar to human vision system, so they have been widely used in recognition application, such as face recognition, fingerprint recognition, character recognition, etc. This feature based method aims to find the important features and represent the corresponding information in an efficient way. In the space domain, the 2D Gabor filter can be considered as a Gaussian kernel modulated by a sinusoidal plane wave. The Gabor wavelets can be defined as follows:

$$\varphi_{u,v}(k,z) = \frac{\|k_{u,v}\|}{\sigma_2} exp\left[-\frac{\|k_{u,v}\|^2\|z\|^2}{2\sigma^2}\right]\left[\exp(ik_{u,v}z) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (1)$$

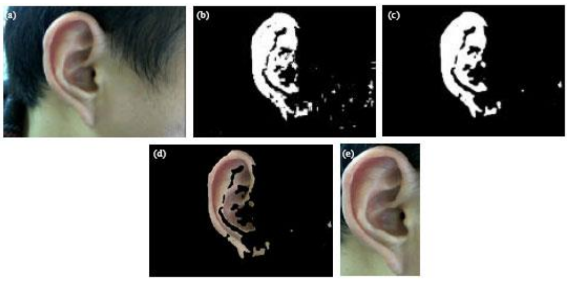

Figure1(a-e): Ear detection, (a) Original image of right face containing right ear, (b) Image with contour extraction and filling, (c) Image after smoothing, (d) Result of segmentation and (e) Normalized ear image

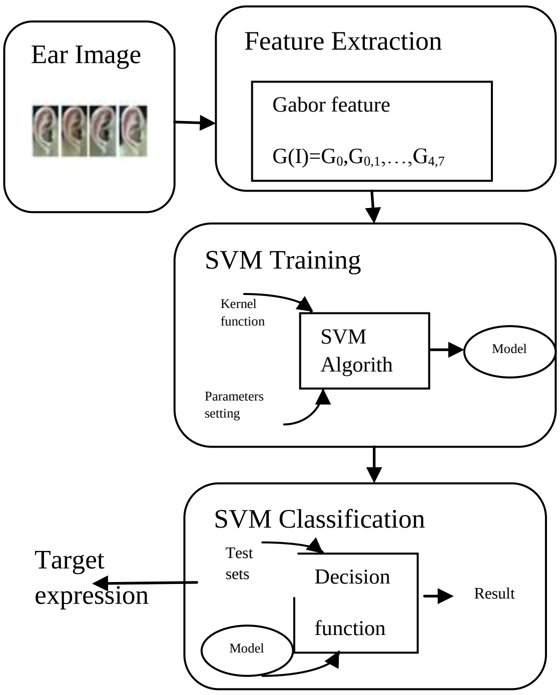

Figure 2 : The stage of ear

The parameter u defined the orientation of the Gabor kernels and the parameter v defines the scale of the Gabor kernels, z= (x, y), ||.|| denotes the norm operator and the wave vector $k_{u,v}$ is defined as follow

$$k_{u,v} = k_v e^{i\varphi} \quad (2)$$

Where, $k_v=k_{max}/f^v$ and $\varphi_u=u\pi/U$, k_{max} is the maximum frequency and f is the spacing factor between kernels in the frequency domain. Each kernel is a product of a Gaussian envelope and a complex plane wave, while the first term in the square brackets in Eq. 1 determines the oscillatory part of the kernel and the second term compensates for the DC value. The parameter σ determines the ratio of the Gaussian window width to wavelength.

v ϵ {0, 1,2,…,V-1} is called label, in most cases the use of Gabor wavelets of five different scales, so V=5. u ϵ {0,1,2,…,U-1} is orientation label and the Gabor wavelets are usually used eight orientations, so U=8. With the following parameters: $k_{max}=\pi/2, f=\sqrt{2}, \sigma=2\pi$ the kernel exhibit desirable characteristics of spatial frequency, spatial locality and orientation selectivity.

Representation of Gabor feature: The Gabor wavelet representation of an image can be obtained by convolving the image with a family of Gabor kernels as defined by Eq.1. Let I(x, y) be the gray level distribution of an image and define the convolution of image I and a Gabor kernel $\varphi_{u,v}$ as follows:

$$G(z) = I(z)^*\varphi(k,z) \quad (3)$$

Where, z = (x, y) and * denotes the convolution operator, G (z) is the convolution result. As described V =5, U=8, 40 Gabor filters are made, the set S= {$G_{u,v}$ (k,z): uϵ {0,…,3}, vϵ {0,…,7} forms the Gabor wavelet representation of the image I(z). Where $G_{u,v}$ (k, z) = |I(z)*G (k,z), all the Gabor features can be described as $G(I)=G=(G_0,G_{0,1}\ldots G_{4,7})$. Taking an image of size 128×128 for example, the Gabor feature vector will be 128×128×5×8=655360 dimension which is incredibly large. Due to the large number of convolution operations, the computation and memory cost of feature extraction is also necessarily high. So each $G_{u,v}$ (k, z) is down sampled by a factor r. For a vector image, the vector dimension is 10240 when the down sampling factor r=64.

3.2 Classification Based on Support Vector Machine

Support Vector Machine (SVM) is based on results from statistical learning theory. The basic idea of SVM is to map the input space to a higher dimensional feature space and to classify the transformed feature by a hyper-plane. SVM has become one of the most useful approaches in machine learning field due to its good performance of resolving classification problems. Consider the problem of separating the set of labeled training vectors belonging to two separate classes:

$$S = \{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}, x_i = R^n, y_i \in \{-1,1\}$$

With a hyper-plane:

$$(w, x)+ b=0 \quad (4)$$

$$y_i[(w,x_i)+b]-1 \geq 0, i=1,2,\ldots,n \quad (5)$$

Learning systems typically try to find a decision function of the form:

$$f(x) = sgn\{\sum_{i=1}^{n} a_i^* \, y_i(x_i, x) + b^* \qquad (6)$$

The coefficients a^*_i and b^* in Eq.6 are the solution of a quadratic programming problem. For non-linearly separable data, a non-linear mapping function Φ that embeds input vectors into feature space, kernels have the form:

$$K(x,z) = (\emptyset(x), \emptyset(z)) \qquad (7)$$

SVM algorithms separate the training data in feature space by a hyper-plane defined by the type of kernel function used. The kernel functions used are:

-Linear kernel : k (x, x.) ~ (x, x.)
- Radial Basis Function (RBF):
K (x, x.) ~ exp $(-\|x-x_i\|^2/2\sigma^2)$
-Polynomial : K $(x, x_i) \sim [(x, x_i) + l\,]^d$
- Sigmoid : K (x, x.) ~ tanh $(a(x, x_i) + c)$

The SVM methodology learns nonlinear functions of the form:

$$f(x) = sgn\{\sum_{i=1}^{n} a_i^* \, y_i K(x_i, x) + b^* \qquad (8)$$

Multi-class classifications: A multi-class classification can be obtained by composition of two- class SVMs. There are two main strategies to deal with multi-class classification. One is the one-against-all strategy to classify between each class and all the remaining, it needs to construct k SVMs where k is the number of classes, in each SVM, and all data must be included in training. Another is the one-against-one strategy to classify between each pair, only two classes data are used in each SVM, as a tradeoff, k (k-1)/2 classifiers have to be constructed. Because of the one-against-all strategy often leads to ambiguous classification, so one-against-one strategy was adopted in our system and the RBF kernel is used.

A bottom up binary tree is constructed for classification. Take an eight-class data set for example, the decision tree is shown in Fig-3. We encode each class with number from 1 to 8 and the numbers are arbitrary without any means of ordering. One class number will be chosen representing the "winner" of the current two classes after comparison between each pair and the selected classes from the lowest level of the binary tree will come to the upper level for another round of tests. Finally, the unique class will appear on the top of the tree.

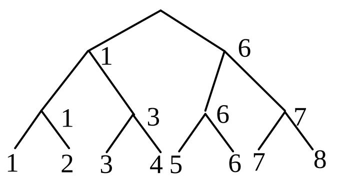

Figure 3: The binary tree with 8 classes

Suppose the number of classes as k, the SVMs will learn k (k-1)/2 discrimination functions in the training stage and each binary tree has k-1 time's comparisons.

4. Implementation for Proposed System

In this paper, the framework chooses 40 volunteers as proposed system's subjects. The system takes 10 images with right face containing ear for each subject. These images are taken under different view conditions and different illumination conditions. Besides, the distance between camera and subject is considered.

The normalized ear images with a resolution of 45 × 80, 40 Gabor filters are used to extract features, the Gabor feature vector of each ear image will be 45×80×5×8=14400 dimensions. Down sampling Gabor feature vector by factor r=64, the vector dimension is reduced to 225. Forty subject take part in the experiment, that means 40 classes, so the number of classes k=40, each class contains 10 images, thus, we have 400 ear images in total. Decomposing 40=32+8, two binary tree are constructed, one with 32 leaves and the other with 8 leaves. Then compare the two outputs to determine the true class in another binary tree with only two leaves. The times of comparisons for one query are 39. Five ear images are selected at random from each subject as training set and the other 5 ear images as test set, both training set and test set have 5×40=200 images. The experiments are repeated for 5 times.

In order to evaluate the performance of the proposed method in ear recognition, some widely-used methods are compared with such as Principal Components Analysis (PCA) (Victor *et al*, 2002; Chang *et al*., 2003, Linear Discriminant Analysis (LDA) (Zhang and Jia, 2007), Fisher Discriminants Analysis (FDA) (Liu *et al*, 2006), and Local Binary Pattern (LBP) (Ahonen *et al*., 2006). Those methods are also tested on same ear dataset. From Table1, the proposed framework performs well in ear recognition.

Table 1: Experimental results of ear recognition via different methods

Methods	Recognition rate (%)
PCA	73.54
LDA	79.63
FDA	83.97
LBP	88.38
Proposed method	93.74

5. Conclusion and Future Work

In this paper, a novel approach of ear recognition are presented, ears are detected from person's side face, then Gabor wavelets are used for feature extraction and multi-class SVM classification with binary tree strategy is performed for ear recognition. The experimental results shows the proposed framework perform better than methods in some literatures. The ear biometrics has the potential to be used in application to identify or recognize humans by their ears. Also, the ear biometrics can be combined with other biometrics as security application. In the further work, proposed method will focus on the recognition with part-covered ear and the dimension reduction of Gabor feature.

REFERENCES

[1]A. Iannarelli, *Ear Identification*, Forensic Identification Series.Paramont Publishing Company, 1989.

[2] Belé Moreno, Ángel Aánchez and José F.Vélez, “Use Outer EarImages for Personal Identification in Security Applications,” in *Proc.of the 33rd Annual 1999 International Carnahan Conference,* Madrid,1999,

[3] Hui. Chen and B. Bhanu, “Human Ear Recognition in 3D,” *IEEETrans. on Pattern Analysis and Machine Intelligence*, vol.29, no.4, Apr.2007, pp. 718-737..

[4] M. Burge and W. Burger, “Ear biometrics in computer vision,” *Proc. Int. Conf. on Pattern Recognition*, vol. 2, pp.822–826, 2000.

[5] K. Chang, K. Bowyer, S. Sarkar, and B. Victor, “Comparison and combination of ear and face images in appearance-based biometrics,” *IEEE Trans. Pattern Analysis and Machine Intelligence*,vol. 25, no. 9, pp. 1160–1165, 2003.

[6] Mu Zhichun, Yuan Li, Xu Zhengguang, “Shape and Structural Feature Based Ear Recognition,” *Advances in Biometric PersonAuthentication,* Guangzhou, China, Dec. 2004, pp. 663-670.

[7] Lei, X. and P. Zhou, 2012. “An intrusion detection model based on GS-SVM Classifier” Inform. Technol. J., 11: 794-798.

[8] Chang, K., Bowyer. K.W., Sarkar, S, Victor, B, “Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, September 2003, Vol. 25, No. 9,

[9] H. Chen and B. Bhanu, “Human ear detection from side face range images,” *Proc. Int. Conf. on Pattern Recognition*, vol. 3, pp. 574–577, 2004.

[10] Ping Yan and Kevin W. Bowyer, “Biometric recognition using three-dimensional ear shape,”