

Ear Recognition Based on Statistical Shape Model

LU Lu, ZHANG Xiaoxun, ZHAO Youdong and JIA Yunde

*Department of computer science and engineering,
Beijing Institute of Technology, Beijing 100081, PR China
{lulu123,jiayunde}@bit.edu.cn*

Abstract

Alfred system suggests that ear shape can be used as a unique and comparable feature for identity recognition. In this paper, we aim to extract ear shape features for recognition. Active Shape Models (ASMs) is applied to model the shape and local appearance of the ear in a statistical manner. In addition, steerable features are extracted from the ear image ahead of ASMs. Steerable features encode rich discriminant information of the local structural texture and provide accurate guidance for shape location. Eigeneareshape is used for final classification. Experiments on an ear database show an encouraging performance. Some experiments on double ears combined for recognition are also carried out, and indicate that the fusion outperforms either ear alone.

1. Introduction

Using ear in identifying people has been interesting in the past few years. There are many methods proposed for ear recognition. Burge and Burger [1] proposed automating ear biometrics with Voronoi diagram of its curve segments. Hurley, Nixon and Carter [3] used force field transformations for ear recognition. Bhangu [4] proposed an approach of human ears recognition in 3D images and presented a new local surface descriptor for surface representation for ears. Victor, Bowyer and Sarkar [6] made a comparison between face and ear recognition using principle component analysis. Chang [5] found that multi-model recognition using both ear and face results in statistically significant important over either individual biometric. Moreno et al. [7] presented multiple identification method, which combines the result from several neural classifier using feature outer ear points.

The two studies from Alfred Iannarelli in 1989 gave supporting evidence for ear using as biometrics. The first study compared over 10,000 ears drawn from a randomly selected sample, and the second study examined fraternal and identical twins. The evidence from these studies supports the hypothesis that the ear

contains unique physiological features [1]. Alfred Iannarelli has created a 12 measurement system “Alfred system” (see Figure 1). The identification consists of 12 measurements and the basic information about people [2].

Alfred system suggests that ear shape can be used as a unique and comparable feature for identity recognition. The objective of this paper is to extract ear

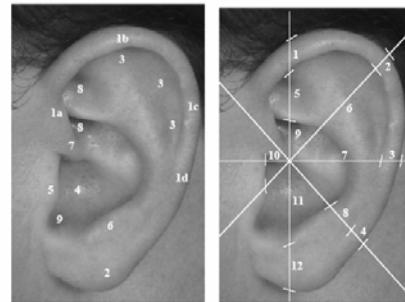


Figure 1 Alfred system

shape features for recognition. Thus, we use ASMs to model the shape and local gray texture of the ear. And steerable features are extracted from the ear image ahead of ASMs to provide accurate guidance for shape location. Eigenshape, which has proved to be successful in gait recognition [12], is used for final classification. Experiments on an ear database show an encouraging performance. Some experiments on double ears combined for recognition are also carried out, and indicate that the fusion outperforms either individual ear.

2. Ear Shape Model

The difficulty of application of the “Alfred system” is the location of the center point in the system. If the first point is not defined accurately, none of the measurements are useful [2]. Alfred system suggests that ear shape can be used as a unique and comparable feature for identity recognition. ASMs, which were proposed by Cootes et al. [8], are used to effectively model both the shape and local appearance of ear in a statistical manner.

2.1. Point Distribution Model

This model is used to describe a shape and its typical appearances. By sampling the coordinates of the n labeled points (landmarks) of each ear image and concatenating into a vector s , which represents an ear shape (see Figure 2):

$$s_i = [x_{i0}, y_{i0}, x_{i1}, y_{i1}, \dots, x_{in-1}, y_{in-1}]^T; 1 \leq i \leq n$$

In order to study the variations of the position of each landmark throughout the set of training images, all shape must be aligned to each other. This is done by changing the pose (scale, rotation, and translation) of consecutive shapes until the whole set is properly aligned. Then we capture the statistic of shape. Principle Component Analysis is applied to the aligned shapes:

$$s = \bar{s} + Pb \quad (1)$$

Where \bar{s} is the mean shape vector, P is the set of orthogonal models of shape variation, and b is a vector of shape parameters. By varying the elements of b we can vary the shape using Equation.

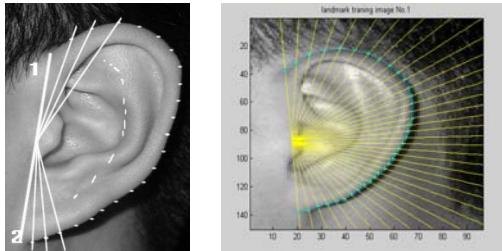


Figure 2 the landmarks of the ear shape, 37 landmarks used to represents a shape

2.2. Local Appearance Model

The local appearance models describe the local gray features around each landmark. It modeled as the first derivative of the sample profiles perpendicular to the landmark contour [8].

We assumed that local gray models are distributed as a multivariate Gaussian. For the j^{th} landmark, we can derive the mean profile \bar{g}_j and the sample covariance matrix S_j from the j^{th} profile example directly. The quality of fitting a feature vector g_s at teat image location s to the j^{th} model is given by calculating the Mahalanobis distance from the feature vector to the j^{th} model mean.

$$f_j(g_s) = (g_s - \bar{g}_j)^T S_j^{-1} (g_s - \bar{g}_j) \quad (2)$$

At the current position s , when searching points, the local appearance models find the best candidate in the neighborhood of the search point, by minimizing $f_j(g_s)$ [8].

3. Ear Shape Model based on Multi-orientation and Multi-resolution

Although ASMs is a powerful statistical tool for shape, it can suffer from changes in illumination and local minima in optimization and the result depends on initialization.

On one hand, if we ensure that the “best candidate point” is included in the profile perpendicular to the landmark contour, the profile must be long enough to include the target point. On the other hand, the more points the profile include, the more consume we cost to find the target. The problem can be solved by a multi-resolution approach. We must generate a pyramid of images with different resolutions and obtain the profile statistics for each landmark and each pyramidal level. The procedure which starts at the highest level is repeated until the lowest (the original image) is reached the convergence.

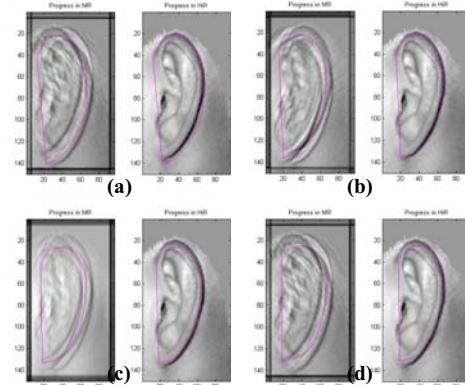


Figure 3 the result of shape search introduced local steerable magnitude features including four orientation: (a) 0 degree,(b) 45 degree,(c) 90 degree ,(d) 135 degree

In ASMs, the local appearance model, which represents the local gray model around each landmark, can not find the accurate shape but get stuck at the local minima. Feng Jiao, Stan Li et al. [9] proposed to apply the Gabor to the ASMs. And they proved the wavelet features can improve the search accurateness. Gabor wavelet feature has achieved great success in search shape with the characteristic of multi-orientation and multi-scale. However, Gabor filters prohibitively time-consuming in image. Steerable filter is a kind of oriented filters and it can capture the local structure corresponding to any orientation on one scale. The magnitude and phase of

Steerable features encode rich discriminative information of the local structural texture of the object as well as Gabor filter and provide accurate guidance for the search. But steerable filter costs much less computation due to its two distinct properties, X-Y separation and symmetry or anti-symmetry [10]. Thus, steerable features are extracted from the ear image ahead of ASMs to improve the performance of ear shape location (see Figure 3).

4. Eigeneearshape

Eigenshape models represent the variability of a set of example shapes in a compact and statistically robust way [11]. Liang Wang et al. have applied Eigenshape to gait recognition [12]. In order to obtain the Eigeneearshape, we re-represent shapes as a set of points in a common complex coordinate [12,13]. That is, each shape can be described as a vector $z = [z_1, z_2, \dots, z_n]^T$, where $z_i = x_i + i * y_i$. For two shapes, z_1 and z_2 , if they are equal through a combination of translation, scaling and rotation, we may consider they are the same shape [12,13].

$$z_1 = \alpha l_k + \beta z_2, \quad \beta = |\beta| e^{j\angle\beta} \quad (3)$$

Where αl_k translates z_2 , and $|\beta| \angle\beta$ scale and rotate z_2 , respectively, we may consider they represent the same shape [12,13].

It is very convenient to center shapes by defining the centered vector $v = [v_1, v_2, \dots, v_n]^T$ $v_i = z_i - \bar{z}$, $\bar{z} = \sum_{i=1}^n z_i / n$. The full Procrustes distance $d_F = (v_1, v_2)$ [12,13] between two configurations can be defined as

$$d_F(v_1, v_2) = 1 - \frac{|v_1^* v_2|^2}{\|v_1\|^2 \|v_2\|^2} \quad (4)$$

Which minimizes

$$\left\| \frac{v_1}{\|v_1\|} - \alpha l_k - \beta \frac{v_2}{\|v_2\|} \right\|^2 \quad (5)$$

Note that the superscript * represents the complex conjugation transpose and $0 \leq d_F \leq 1$ [12,13].

Given a set of n shapes, we can find their mean by finding v that minimizes the objective function

$$\min_{\alpha_i, \beta_i} \sum_{i=1}^n \|v - \alpha_i l_k - \beta_i v_i\|^2 \quad (6)$$

To find v , we compute the following matrix:

$$S_v = \sum_{i=1}^n (v_i v_i^*) / (v_i^* v_i) \quad (7)$$

The mean shape \hat{v} is the dominant eigenvector of S_v , the eigenvector that corresponds to the greatest eigenvalue of S_v [12,13]. This vector is used as Eigeneearshape.

5. Experiments

To demonstrate the efficiency of our method, extensive experiments are carried out on an ear database. Following geometrical normalization, histogram equalization is also performed for intensity normalization in our experiments.

5.1. Ear Database

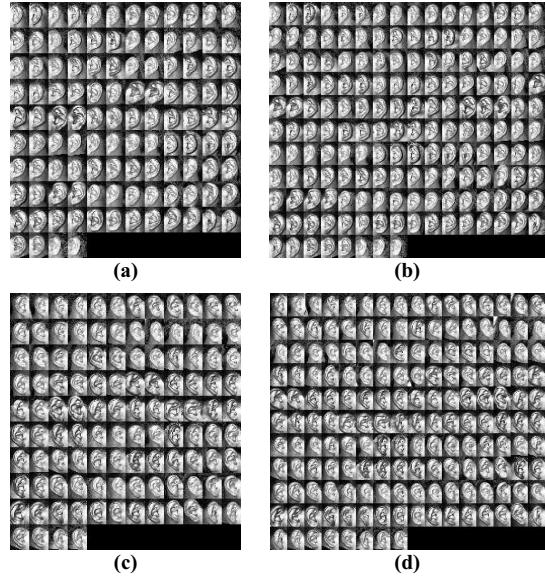


Figure 4 Gallery and probe set of left and right ears. (a) gallery for left ear (b) probe set for left ear (c) gallery for right ear (d) probe set for right ear.

This database mainly concerns ear recognition with slight pose change. Pose change has more influence on performance than other factors such as expression and illumination [14]. A digital camera (Sony DSC-S85) fixed on a tripod is used to capture images. 10 images are captured for individual: 5 images for the left ear and 5 for the right ear, corresponding to 5 different poses. The angle for ear images between two adjacent poses is 5 degree. The ED now contains the images of 56 individuals, namely 560 images totally. The size of original image is 640 by 480, while the size of ear normalized images is 38 by 60 (see Figure 5).

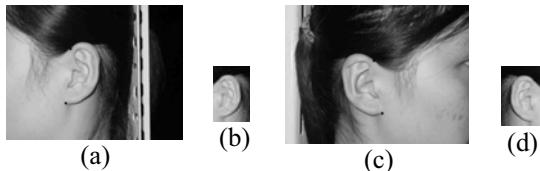


Figure 5 (a)(c) original images (b)(d)normalized images

We divided all ear images into the galleries and the probe sets. Two images with the slight pose change are included in the gallery for training and reference, and other three images are included in the probe set (see Figure 4).

5.2. Ear Recognition

Eareigenshape is applied in ear recognition. Victor et al. have found that the multi-model recognition using both ear and face results in statistically significant important over either individual biometric [5]. We can believe that double ears combined for recognition can obtain a better performance than either one alone. We experiment on the left ear and the right ear respectively. Another experiment was performed to investigate the value of a fusion recognition using left and right ear. Recognition rate for experiment using 168-image training sets is 95.1% for the fusion recognition versus 93.3% for the left ear and 93.3% for the right ear. It indicates that the fusion outperforms either ear alone (see Figure 6).

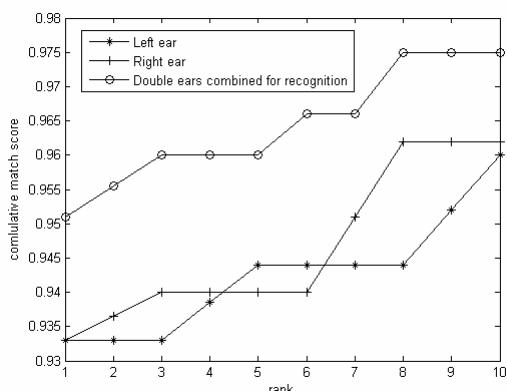


Figure 7 efficiency of the eigeneearshape algorithm

6. Conclusions

In conclusion we have showed that ear biometrics can be used for identification and for the further development. In this paper, we use Active Shape Models to represents the ear shape according to Alfred system. Ahead of ASMs, steerable features are extracted for improving the performance of shape location. Finally, Eigeneearshape is applied in ear

recognition. The experiments on an ear database show an encouraging performance.

7. References

- [1] Mark Burge and Wilhelm Burger. "Ear Biometrics". BIOMETRICS: Personal Identification in a Networked Society. 1999, p.273-286.
- [2] Hanna-Kaisa Lammi. "Ear Biometrics". Lappeenranta University of Technology, 2004
- [3] David J. Hurley, Mark S. Nixon, John N. Carter. "A New Force Field Transform for Ear and Face Recognition", Proceedings of the IEEE International Conference of Image Processing, 2000, p.25-28.
- [4] B.Bhanu and H.Chen, "Human Ear Recognition in 3D", Workshop on Multimodal user Authentication, Dec 2003.
- [5] Kyong Chang, Kevin W. Bowyer, and Sudeep Sarkar, Barnabas Victor. "Comparison and Combination of Ear and Face Images In Appearance-Based Biometrics". IEEE trans. On PAMI Sep. 2003, vol. 25, no. 9.
- [6] B. Victor, K. Bowyer, and S. Sarkar, "An evaluation of face and ear biometrica", IEEE International Conference on Pattern Recognition, pp. 429-432, August 2002.
- [7] Carrira-Perpinan, "Compression neural networks for feature extraction: Application to human recognition from ear images", MSC. thesis, Technical University of Madrid, Spain, 1995.
- [8] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham, Active Shape Models – Their Training and Application., Computer Vision and Image Understanding, 61(1), 38-59, 1995
- [9] Feng Jiao, Stan Li, Heung-Yeung Shum and Dale Schuurmans. Face alignment using statistical models and wavelet features. CVPR 2003.
- [10] Xiaoxun Zhang, Yunde Jia. "Symmetrical Null Space LDA for Object Recognition". International Conference on Intelligence Computing, 2005
- [11] Baumberg, A. and Hogg, D. An adaptive eigenshape model. In Pycock, D. (ed.), 6th British Machine Vision Conf., pp. 87–96. BMVA Press, 1995
- [12] Liang Wang, Huazhong Ning, Weiming Hu, Tieniu Tan. "Gait Recognition based on Procrustes Shape Analysis". IEEE ICIP 2002
- [13] J.Boyd, "Video Phase-Locked Loops in Gait Recognition", in Proc. Of International Conference on Computer Vision, 2001
- [14] P.Jonathon Phillips, Hyeonjoon Moon, Syed A.Rizvi, and Patrick J.Rauss. "The FERET Evaluation Methodology for Face-Recognition Algorithms". IEEE Trans. PAMI, Vol. 22, NO. 10, pp.1090-1104, 2000