University of London

Imperial College of Science, Technology and Medicine

Department of Computing

Title of my Ph.D. Thesis

Joe Bloggs

Abstract

Text of the Abstract.

Acknowledgements

I would like to express (whatever feelings I have) to:

- My supervisor

- My second supervisor

- Other researchers

- My family and friends

Contents

List of Tables

List of Figures

x

Chapter 1

Introduction

1.1   Motivation and Objectives

Motivation and Objectives here.

1.2   Literature Reviews

1.2.1   The Early Work towards Ear Forensics by Iannarelli

In 1949, Alfred Iannarelli was well-known in American as an ear identification export who providing ear evidence as a strong personal identification in the context of forensic science[3]. He also developed a measurement system in order to representing the ear by numbers, which was used by American law enforcement agencies at that time[4]. As shown in Figure 1.1, he aligned the ear into 4 reference lines which divided the image into equally 45° intervals. The intersection of those lines is located on the tragus and the crus of helix (see section 4.1 for a description of the anatomical parts of the ear), therefore the ear image can be represented based on the edge markers of each reference line.
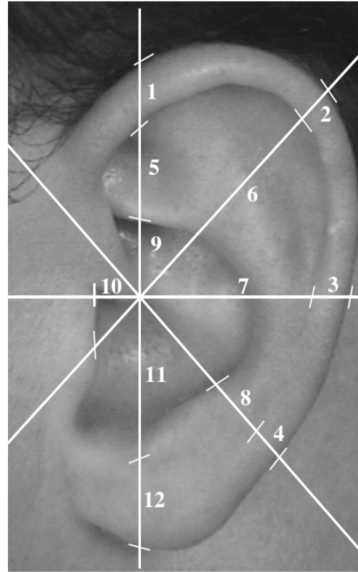
Need change ences

Figure 1.1: Iannarellis manual ear measurement system.

According to Iannarelli's book, through 38 years of research and application in earology on more than 10,000 ear images, no two ears were found to be identical, not even the two ears from the same person[4].

Although this system was based on a man-power visual measurement which may be not very accurate, the usage time of the successful application has proved itself to be very useful. It also inspired a lot of researches towards ear recognition using more advanced technology later.

1.2.2 Burge and Burgers' method

In the early time of the 21st Century, Burge and Burger started to use machine vision for ear recognition system. They were the first who managed to use machine model each individual ear with an adjacency graph.

The main step of their method can be shown as figure 1.2. They used the canny edge detection firstly to extract the "ear print", then reconstructed it into a Voronoi diagram which looks like a segmentation of the ear parts. Finally, join the center of each segment to form the "N-graph" which is used to authenticating a person.
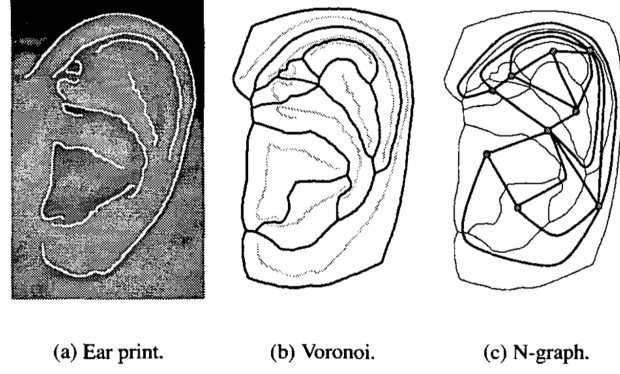
(a) Ear print.   (b) Voronoi.   (c) N-graph.

Figure 1.2: Burge and Burgers' ear model method.

1.2.3   Reduced Hough Transform for Ear Detection

Hough Transform (HT) is a very classic algorithm in image processing widely used for feature extraction and patten recognition. It is useful to find the imperfect instances of objects in certain shapes which is very suitable for ear detection as the ear is just like a ellipse and remains that way. Hence, in 2007 Arbab-Zavar et al.[1] used it to design an algorithm for automatic ear detection. Despite the advantages, HT has certain drawbacks such as high computational requirement and memory usages, therefore they used a reduced Hough Transform to overcome those problems as it was specified for only detect ellipse using the known geometrical properties of it to decompose the parameter space from 5D to 2D.

(a)   (b)   (c)   (d)
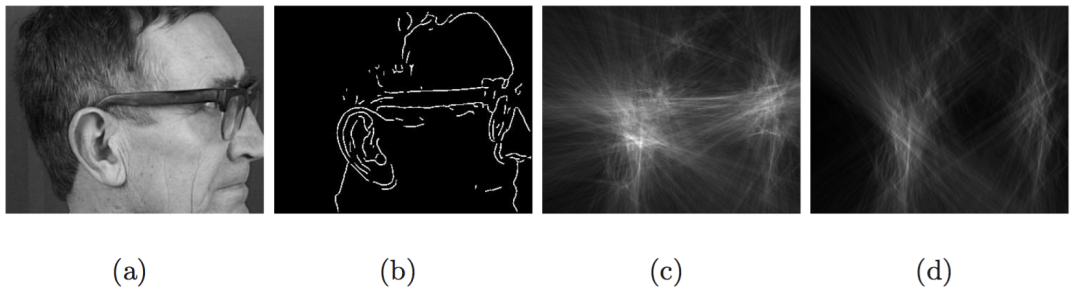
Figure 1.3: The process of Arbab-Zavar's method [1]. (b)Canny edge (c)accumulator of ellipse (d)reduce the horizontal vote

Firstly, a Canny operator is applied to get a smoothed edge detected image, then the reduced HT transform reconstruct it into an accumulator space as it shown in figure 1.3

(c). The locations of the peaks will provide the coordination of the best matching ellipse. However, there are some mismatches as well, such as the presence of the spectacles. The way to eliminate most of them is to get rid of the horizontal vote by HT, due to the ear shape is mostly a vertical ellipse in the database.

The result shows that this algorithm achieves error-free on the XM2VTS derived, 252-image database. But not that good in UND database, given that it include more backgrounds informations. When applying occlusions, despite the HT was known as tolerate to noise and imperfect, the detection accuracy still drops below 80% after 40% occlusion percentages and reaches 30% at 70% partial occlusion [1]. This algorithm can be seen as very successful under the condition of head profile only database, without backgrounds and large hair covers. It is a very good example of how to maximize the usage of classic methods, however the speed of detection needs to be evaluated if the algorithm needs to be improved.

1.2.4   Haar-like Features for Ear Detection

The Haar-like feature was a very successful algorithm when applied for human face detection by Viola at 2001[5]. Therefore, in 2009 Yuan et al.[6] designed an algorithm which used Haar-like features for ear detection. There are some extended asymmetric Haar-like features as in figure 1.4 were added as the ear structure is different from face.
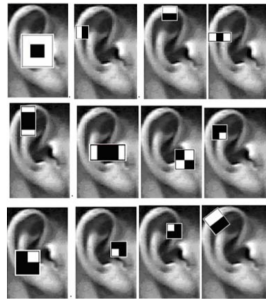

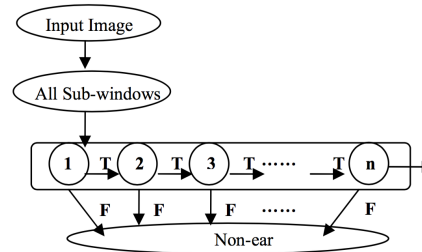**Figure 1.4:** Haar-like features reflects the local features

Figure 1.5: The procedure of classify ears

They trained several strong classifiers with AdaBoost algorithm and then cascaded them together into a multi-layer classifier [6]. As all the sub-windows of the image pass through

each one of the single classifier, false response will immediately reject the corresponding sub-window. Only the one went through all classifiers will be marked as an ear....

1.2.5   SIFT Point Matching for Ear Detection

Scale-invariant feature transform also known as SIFT is a famous computer vision algorithm for detecting and describing local features in images. It extract the interest point of the object in an image to provide "feature description" of the object, then a chosen distance can be used to decide if two description are the same between two images. Advantages of the SIFT for ear detection is that it is scale-invariant and not sensitive with noise and illumination.

Based on the method written by Brown et al.[7], they try to created a homography transform between a probe image and a known gallery object image using SIFT matches. If an homography can be created, means that the probe contains the gallery object. In addition, 4 matched SIFT points were used to align if they lies in one plane , due to the unreliability of more points. Although it can provided an accurate result, a RANSAC algorithm was used to select the best match.

The detection result was under the XM2VTS derived, 252-image database shown that it achieved 96% rank-1 detection accuracy [8]. However it requires the predefined galley of object image which needs a manual mask to locate the object.

1.2.6   Shaped Wavelets for Ear Detection

Due to the ear image mainly contains a lot of curvilinear structures, Ibrahim et al.[2] convolve the image with some curved wavelet filters called "banana wavelet" shown as the figure 1.6 to perform a generalized template matching to detect the location of ear.

The Banana wavelets are a generalization of Gabor wavelets and it can be parameterized by four variables: frequency, curvature, orientation and size. They use 8 filters which
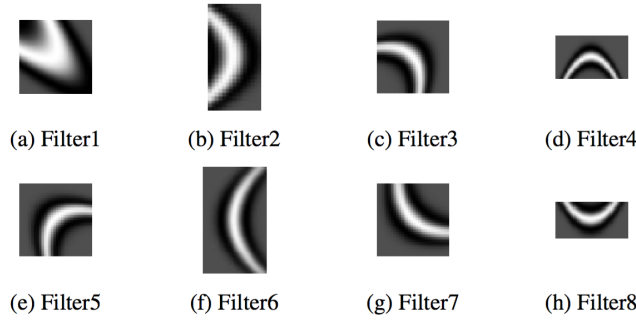
(a) Filter1   (b) Filter2   (c) Filter3   (d) Filter4

(e) Filter5   (f) Filter6   (g) Filter7   (h) Filter8

Figure 1.6: Banana wavelets used in this method[2] .

prove to be sufficient for ear detection as it shown above in figure 1.6. Initially, they do the convolution between image and the banana wavelet which resulting the magnitude of the filter response. Then the local maxima of the magnitude should be the position where ear has similar curvature, size and orientation to the specific corresponding banana wavelets, as it shown below in figure 1.7. Finally, some threshold and anti-overlapping algorithms can be applied to make the decision where is the ear.
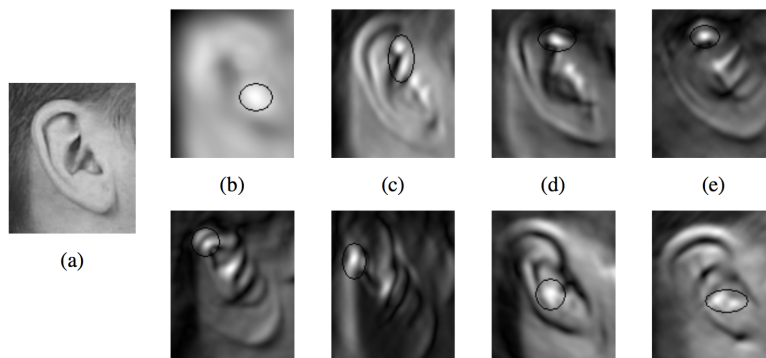

Figure 1.7: (a) Input image, and (b)-(i) after convolution with 8 banana filters[2]

The results demonstrate that this is a very promising methods which achieve 100% detection rate on the XM2VTS database and above 98% when the Gaussian noise ($\sigma = 100$) is presented. However, when testing on the SOTON database with some occlusion, the detection accuracy drop dramatically to 44.7% with partial head and small occlusion[2] .

Although the algorithm treat the ear as the combination of some curved lines and focus on finding those lines, it cannot achieve better detection rate when occlusion occurs.

The amount of calculation must be very big due to the several fully convolution of whole image with 8 wavelet filters, but the author did not mention anything about speed.

1.2.7 Active Contour for Ear Detection

In 2011, Kumar et al.[9] tried to develop an online application which can use ear biometric as an authentication method. They build a wooden stand for acquiring the image as it shown in figure 1.8, so there are no extra illumination management require, the ear was captured inside a isolated box with utilized camera flash light. Which should make the detection more easy, however can not handle the real scenario image.
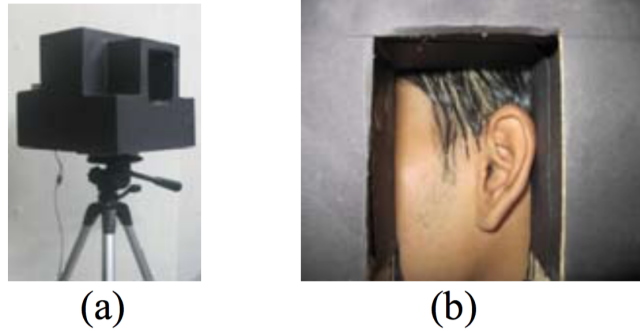

(a) (b)

Figure 1.8: (a)Imaging setup (b)Sample captured image

First of all, they use Gaussian classifier to detect the skin region, then apply LoG(Laplacian of Gaussian) for edge detection and remove the false edge until only ear edge left. By using the top and bottom pixel of the ear edge, they can manage to rotate the ear with reference to the vertical axis. Finally, a localized region based active contour model is applied to extract the ear part only from the ear-ROI(Region of Interest). The whole process is shown in figure 1.9.

The database they created is 100 users with 7 pictures each. However, the detection rate from the ear-ROI is about 94.2%, which means the correct ear-contours extract is 660 images, the blurry images and the presence of hair are the ones failed [9]. As they only uses the specific controllable database, it is no doubt that this method will failed more when applying in real life images.
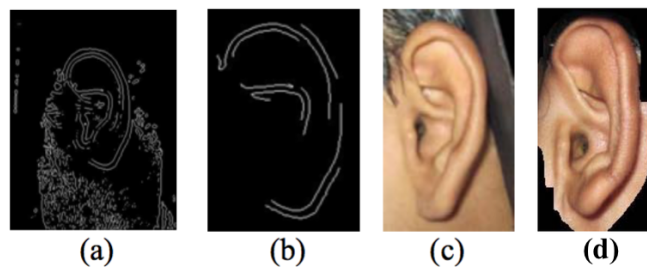
Figure 1.9: (a)Ear Edge by LoG (b)False edge removal (c)ear-ROI (d)Ear-contours

Bibliography

[1] B. Arbab-Zavar and M. S. Nixon, "On shape-mediated enrolment in ear biometrics," *Advances in visual computing*, 2007.

[2] M. I. S. Ibrahim, M. S. Nixon, and S. Mahmoodi, "Shaped Wavelets for Curvilinear Structures for Ear Biometrics." *ISVC*, pp. 499–508, 2010.

[3] S. v. David Wayne Kunze, Court of Appeals of Washington, Division 2. 97 Wash. App. 832, 988 P.2d 977, 1999.

[4] B. Arbab-Zavar, "On guided model-based analysis for ear biometrics," Ph.D. dissertation, School of Electronics and Computer Science, University of Southampton, 2009.

[5] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, 2001.

[6] L. Yuan and F. Zhang, "Ear detection based on improved AdaBoost algorithm," in *2009 International Conference on Machine Learning and Cybernetics (ICMLC)*. IEEE, 2009, pp. 2414–2417.

[7] M. Brown and D. Lowe, "Invariant Features from Interest Point Groups," in *British Machine Vision Conference 2002*. British Machine Vision Association, 2002, pp. 23.1–23.10.

[8] J. D. Bustard and M. S. Nixon, "Robust 2D Ear Registration and Recognition Based on SIFT Point Matching," in *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems.* IEEE, 2008, pp. 1–6.

[9] A. Kumar, M. Hanmandlu, M. Kuldeep, and H. M. Gupta, *Automatic Ear Detection for Online Biometric Applications.* IEEE, 2011.