INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

Ear localization in 2D and 3D Profile Face Images-A Survey

Devang Vyas[1], Kinjal Joshi[2], Astha Baxi[3]

[1]PG student, Department of Information Technology, Parul Institute of Engineering & Technology, Vadodara, Gujarat
[1]vyasdev2005@gmail.com

[2]Assistant Professor, Department of Computer Engineering, G H Patel College of Engineering and Technology, Vallabh Vidyanagar, Gujarat
[2] kinjalvjoshi@gmail.com

[3]Assistant Professor, Department of Computer Engineering, Parul Institute of Engineering & Technology, Vadodara, Gujarat
[3] adbaxi@gmail.com

***Abstract**: Ear is a new class of relatively stable biometrics which is not affected by facial expressions, cosmetics and eye glasses and aging between 8 to 70 years. To use ear biometrics for human identification, ear detection is the first part of an ear recognition system. In this paper, the survey of different techniques for ear detection from 2D images and 3D range images has been included. Here overview of few 2D methods and 3D methods are given. At the end, comparison of different methods is given with the details of different used databases, Number of tested images and percentage wise accuracy. Occlusion by hair or earrings, profile face orientation and image quality are open research problem for ear detection. Here some standard methods are briefly explained. For ear detection from 2D image five techniques: 1) Active contour method 2) Skin-color and template based technique and 3) Cascaded Adaboost method 4) Method based on reduced hough transform and 5) Method based on skin color and graph matching and for ear detection from 3D image two approaches 1) Template matching based ear detection and 2) Ear shape model based detection have been described.*

Keywords: *Active Contour, Ear Detection, Hough transform, Skin-Segmentation*

1. INTRODUCTION

Ear is a viable new class of biometrics since ears have desirable properties such as universality, uniqueness and permanence. The ear has certain advantages over other biometrics. For example, ear is rich in features. It is a stable structure which does not change with the age. It does not change its shape with facial expressions. It has reduced spatial resolution and uniform distribution of color. Furthermore, the ear is larger in size compared to fingerprints and can be easily captured although sometimes it can be hidden with hair and earrings. It has fixed background. For face recognition, when image is side face, only the ear is unique feature from which person can be identified. Although it has certain advantages over other biometrics, it has received little attention compared to other popular biometrics such as face, fingerprint etc. Human ear detection is the first task of a human ear recognition system and its performance significantly affects the overall quality of the system. The number of recent research shows that face recognition is possible and effective for side faces by detecting and recognizing components such as ears.

Burge and Burger [1] located the ear using deformable contours on a Gaussian pyramid representation of the image gradient. Then edges are computed using the Canny operator, and edge relaxation is used to form larger curve segments, after which the remaining small curve segments are removed. Ansari and Gupta [2] used outer helix curves of ears moving parallel to each other as feature for localizing ear in an image. Using the Canny edge detector, edges are extracted from the whole image. These edges are segmented into convex and concave edges. From these segmented edges,

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

expected outer helix edges are determined. They assembled a database of 700 side faces, and reported an accuracy of ~ 93%.Since this technique solely relies on the parallelism between the outer helix curves, it fails if the helix edges are not proper. Abdel Mottaleb and Zhou [3] segmented the ear from a face profile based on template matching, where they modeled the ear by its external curve. Yuizono et al. [4] also used a template matching technique for detection. They used both a hierarchical pyramid and sequential similarity computation to speed up the detection of the ear from 2D images. HajSaid et al. [5] addressed the problem of a fully automated ear segmentation scheme by employing morphological operators. They used low computational-cost appearance-based features for segmentation and a learning-based Bayesian classifier for determining whether the output of the segmentation is incorrect or not. They achieved 90% accuracy on 3750 facial images corresponding to 376 subjects in the WVU database. Watabe et al. [6] introduced the notion of "jet space similarity" for ear detection, which denotes the similarity between Gabor jets and reconstructed jets obtained via Principal Component Analysis (PCA). They used the XM2VTS database for evaluation; however they did not report their algorithm's accuracy. Cummings et al. [7] used the image ray transform, based upon an analogy to light rays, to detect ears in an image. This transformation is capable of highlighting tubular structures such as the helix of the ear and spectacle frames. By exploiting the elliptical shape of the helix, this method was used to segment the ear region. This technique achieved a detection rate of 99.6% using the XM2VTS database. Yuan and Mu [8] have proposed an ear detection method based on skin-color and contour information. They assume the ear shape elliptical and fit ellipse to the edges to get the accurate position of the ear. The assumption of considering shape of the ear elliptical is not true in general and does not help is detecting the ear in all cases. In [9], Sana et al. have proposed an ear detection scheme based on wavelet based templates. In real scenario, ear occurs in various sizes and the pre estimated templates are not sufficient to handle all the situations, and automatic resizing need to be done. Liu [10] performed ear segmentation using histogram based K-means clustering and Hough transformation for ear detection.

In the context of 3D ear detection: Zhou et al. [11] introduced a novel shape-based feature set, termed the Histograms of Categorized Shapes (HCS), for robust 3D ear detection. They used a sliding window approach and a linear Support Vector Machine (SVM) classifier. They reported a perfect detection rate, i.e., a 100% detection rate with a 0% false positive rate, on a validation set consisting of 142 range profile images from the UND, collection F, database.

Except above methods some standard methods for any object detection like template matching, active contour, adaboost are also used for ear detection. Here these methods and other two methods, one method based on reduced Hough transform and another method based on skin color and graph matching for 2D image are described in section 2. For ear detection from 3D image two approaches Template Matching Based Ear Detection and Ear Shape Model Based Detection are described in section 3.Comparison of different methods is shown in tabular form in section 4. At last conclusion and future work is shown.

2. EAR DETECTION METHODS FOR 2D SIDE FACE IMAGES

Detecting ears from arbitrary side face 2D images is a challenging problem due to the fact that ear images can vary in appearance under different viewing and illumination conditions. The advantage of ear detection methods for 2D image over methods for 3D side face range image is that it does not require additional requirements to capture 3D data.

2.1 Active Contour Method

It is proposed byBalaji, Vasan, Srinivasan.[12]. First skin segmentation is applied on input side face image. Then nose tip is detected. In order to obtain the nose tip, the first non-black pixel of output of skin filter is noted as a nose point only if it is surrounded by non-black pixels as well. So the first non black pixel in each column is noted which is considered as nose tip. Once the nose tip is identified, a rectangular region is extracted from the side face with dimensions 10cm × 20cm. The probability of finding the ear in this region is high. Once this region is extracted, the classical snake function [13] is used to detect the ear. Active Contour Method (or Snakes) aims at minimizing an energy function subject to certain constraints. Initial guess is taken to locate ear pit from which the contour starts. Figure 1 shows the output of the Active contour algorithm. The outermost curve (in green) shows the initial contour and the innermost curve (in blue) is the final contour.

ISSN 2320-6802

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

The intermediary path (in black) taken by the snake is also shown.

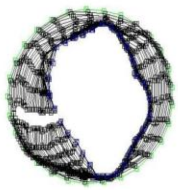


Figure 1: Output of the Active contour method [12]

Analysis: The limitation of this method is the initial guess taken in locating the ear pit. Since only the 2D information is used to locate it, the initial guess might go wrong and this might pull the curve to any edge other than the ear.

2.2 A Skin-Color and Template Based Technique

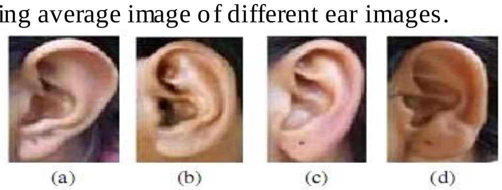
This method is proposed by Surya Prakash, Umarani Jayaraman and Phalguni Gupta[14]

This technique includes 3 steps: Skin Segmentation, Ear Localization and Ear Verification

A color based skin segmentation model is used to segment skin. Ear Localization includes three steps I) Ear Template Creation II) Resizing Ear Template and III) Localization.

- Ear Template Creation:

Human ear shape can broadly be categorized into four classes: triangular, round, oval, and rectangular as shown in figure 2. For creation of ear template all types of ear are considered to obtain a good representative template. Ear template is created by taking average image of different ear images.



(a) Triangular, (b) Round, (c) Oval and (d) Rectangular

Figure 2: Different types of ears [14]

- Resizing Ear Template:

To handle the detection of ears of various sizes, ear template need to be resized according to the width of input profile face image.

- Localization:

To search an ear in the image I, ear template T is moved over the image and normalized cross correlation coefficient (NCC) is computed at every pixel.NCC at point (x; y) is defined by Eqn-1

Eqn-1

are the average of brightness values of the portion of the target image underthe template and template image respectively. Values ofNCC lie between -1.0 and 1.0. When it is typically above apreestimated threshold, It isaccepted that an ear exists in the region.Otherwise it is rejected. Value of NCCcloser to 1 indicates a better match.

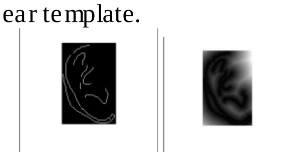
After ear localization verification is done. To measure the similarity, Euclidian distance between the two sets of Zernike moments is used. If the value of distance is less than a preestimated threshold, detection is accepted, otherwise it is rejected.

Analysis: Accuracy of the localization is defined by:(genuine localization/total sample) ×100. It is found to be 94% for this technique. Localization method has failed in some cases, especially for the images which are of poor quality or heavily occluded due to hair.

2.3 Method based on Distance Transform and Template Matching

This method is also proposed by Surya Prakash, Umarani Jayaraman and Phalguni Gupta[15]. There is little modification in previously introduced template matching method.

This technique first segments skin and non-skin regions in the face and then uses template based approach to find the ear location within the skin regions. Ear detection proceeds as follows. First, edge map of the skin regions is computed and further processed to eliminate the spurious edges based on length and curvature based criterion. After getting the clean edge map, its distance transform is obtained on which ear localization process is carried out. Distance transform image of the edge map of an off-line created ear template is employed for ear localization. Figure 3 shows Edge map and distance transform of edge map of ear template.



(a) (b)

Figure 3: (a) Edge map, (b) distance transform of (a)

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

Analysis: Here accuracy is 95.2%. This method is not efficient for noisy images and in the cases where ear is immensely occluded by the hair.

2.4 Cascaded AdaBoost Method

It is proposed by S. M. S. Islam, M. Bennamoun and R. Davies[16].Here ear detection framework is proposed as follows:

- Weak Classifiers:

Weak classifier is defined as a "rough and moderately inaccurate" predictor. In proposed ear detection framework, weak classifiers are built up based on some rectangular features. These features are derived from the idea of Haar wavelets, a natural set of basic functions that compute the difference of intensity in neighboring regions. Similar features with three and four rectangles are used for different types of lines and curves. Finally, the centre-surround feature (F in Figure 4) is used to detect the ear pit.

- Strong Classifiers :

A 'strong' classifier is constructed by combining a set of selected 'weak' classifiers using the AdaBoost (Adaptive Boost) algorithm. The algorithm uses supervised learning with the wrapper method of feature selection. It selects the best weak classifier with respect to a given weighted error of the input samples at each iteration. These classifiers are tuned in favor of those samples which are misclassified by weak classifiers.

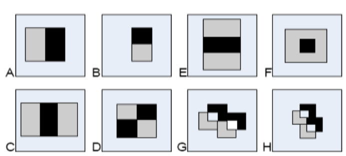


Figure 4: The features used in training the AdaBoost [15]

- Ear Detection with the Cascaded Classifiers :

To speed up detection by using only a small number of features, a cascaded detector is used.
The trained classifiers of all the stages are used in a cascaded manner to build ear detector. To detect various sizes of ears, the detector is scaled up along with the corresponding features. The two ears of a subject are essentially bilaterally symmetric. Therefore, instead of training the system for both ears, it is sufficient to use either right or left ears for training. The features constituting the detector can be flipped to detect remaining ears.

- Multi-Detection Integration:

Since the detector scans over a region in the test image with different scales and shift sizes, there is the possibility of multiple detections of the ear or ear-like regions. Such multiple detections are integrated using clustering algorithm based on percentage of overlap. In this algorithm, pairs of rectangles representing the detected sub windows are clustered together.
Analysis: The proposed ear detector also works well in the presence of partial occlusions involving hair and earrings. The detector is also robust to degradation of images such as the motion blur. The detector fails only in the cases where the majority of the ear is occluded. Method is also fast enough to be used for real-time applications.

2.5 Method based on reduced Hough transform

This method is proposed by Banafshe Arbab-Zavar and Mark S. Nixon [17].
This method proposes the use of the Hough Transform (HT), which can extract shapes with properties equivalent to template matching and is tolerant of noise and occlusion. It finds the elliptical shapes in 2D face profile images to locate the ear regions by using a Hough transform to gather votes for putative ellipse centers in an accumulator, which will go through a refinement process that will eliminate some of the erroneous votes; the location of the peak in this accumulator gives the coordinates of the best matching ellipse.
Analysis: The assumption of elliptical ear shape for all subjects may not be true and may not help in detecting the ear, in general. This technique may correctly approximate the ear boundaries for round and oval shapes but may fail in case of triangular and rectangular shapes. Moreover, assumption of elliptical ear shape restricts the use of these techniques to controlled environment, as the presence of background objects with oval shape may produce false positives. It is found that without occlusion method has 91% accuracy. With 40 % occlusion it has 83%accuracy.

3. EAR DETECTION METHODS FOR 3D SIDE FACE RANGE IMAGES

The most important advantage of methods for ear detection from 3D range images compared to 2D images is that it is not affected under different

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS…..

shading & lighting condition and drawback is that it requires additional requirements to capture 3D data.
For 3D side face range images 2 approaches: Template matching based ear detection and Ear shape model based detection are proposed by H. Chen and B. Bhanu [18].

3.1 Template Matching Based Ear Detection

Here off-line model template is built using shape index. Shape index can be calculated using concepts of principal curvatures. The model template is represented by an averaged histogram of shape index. On side face range image different operations like edge detection, threshold setting, dialation, connected component labeling is applied. So ear region can be found. On that region shape index can be calculated and averaged histogram can be made. Then template matching is applied.
Since histogram can be thought of as an approximation of probability density function, and the x^2 divergence function is used for template matching. Finally over all of the candidate regions, the one with the minimum detected dissimilarity is selected.

3.2 Ear Shape Model Based Detection from 3D Side Face Range Images

In this method first of all Shape model is built. The ear shape model's' is defined as 3D coordinates {x, y, z} of n vertices that lie on the ear helix and anti-helix parts. The shape model 's' is represented by a 3n×1 vector {x1 , y1 , z1 , x2 , y2 , z2 ,…, xn, yn, zn} Figure 5shows the 3D side face range image with textured appearance, in which the ear shape model s marked by vertices is overlaid.

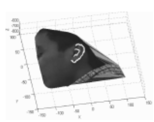


Figure 5: The textured 3D face and overlaid ear shape model (The units of X, Y and Z are mm) [18]

Then Step edges can be detected by gradients. After edge detection, dilation is applied for smoothing. After that thinning is done. Then the edge segments are labeled by running connected component labeling algorithm and some small edge segments are removed. Since the ear region contains several edge segments, the edge segments which are close to each other are grouped into different clusters. For each cluster, the ear shape model is registered with the edges. The region with the minimum mean registration error is declared as the detected ear region. The ear helix and anti-helix parts are identified in this process.
Analysis: The Template matching approach is simple, effective and easy to implement. The Ear shape model approach performs slightly better and more accurate and it is a little bit slower.

4. COMPARISON OF DIFFERENT METHODS OF EAR LOCALIZATION

Various approaches of ear detection are described in section 1, 2 and 3.Obtained accuracy for different approaches is shown in table-1 with information like used database and Number of testing images.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a survey of the ear detection methods. We have discussed 14 methods for 2D images and 3 methods for 3D images with their characteristics and reported their performance. Variation in illumination, occlusion due to hair or earrings, side face orientation is still open research problems for ear localization. For any researcher who wants to start working on ear biometrics, we have presented a systematic discussion that includes available databases and detection techniques with accuracy. In future, work on solving open research problems of ear detection methods can be done. Then ear recognition can also be done for person identification when only side face is available.

Table-1: Accuracy of various ear detection methods

Sr. No	Technique	Database Used	No of Testing Images	Accur-acy
Ear detection methods for 2D side face images				
1.	Deformable Contour based	N/A	N/A	N/A
2.	Outer Helix Curve Based	IITK	700	~93%
3.	Template Matching	N/A	103	N/A
4.	Genetic local search	N/A	110	N/A
5.	Morphological Operations Based Method	WVU	376	90%

6.	Jet space similarity Operations Based Method	XM2VTS	181	N/A
7.	Shape of low-level features	XM2VTS	252	99.6%
8.	Skin color and template based	IITK	150	94%
9.	Distance transform and template matching	IITK	150	95.2%
10.	Active Contour Based	N/A	N/A	N/A
11.	Adaboost	UND	203	100%
12.	Shape based	UND	142	100%
13.	Method based on reduced Hough transform	XM2VTS	252	91%
14.	Histogram Based K-means Clustering and Hough Transformation	CVL	180	90%
Ear detection methods for 3D side face images				
1.	Template matching	UCR	312	92.4%
2.	Ear shape model	UCR	312	92.6%
3.	Histogram based method	UND	142	100%

REFERENCES

[1] BURGE, M. AND BURGER, W. 2000. Ear biometrics in computer vision.In Proc. of the 15th International Conference on Pattern Recognition ICPR. Barcelona, Spain, 826–830.

[2] S. Ansari and P. Gupta. Localization of ear using outer helix curve of the ear. In Proc. of ICCTA, 2007, pages 688–692,2007.

[3] ABDELMOTTALEB, M. AND ZHOU, J. 2006.Human ear recognition from face profile images.In Proc. of the 2nd International Conference on Biometrics ICB. Hong Kong, China, 786–792.

[4] YUIZONO, T., WANG, Y., SATOH, K., AND NAKAYAMA, S. 2002. Study on individual recognition for ear images by using genetic local search. In Proc. of the Congress on Evolutionary Computation CEC. Washington-DC, USA, 237–242.

[5] HAJSAID, E., ABAZA, A., AND AMMAR, H. 2008.Ear segmentation in color facial images using mathematical morphology.In Proc. of the 6th Biometric Consortium Conference BCC. Tampa-FL, USA.

[6] WATABE, D., SAI, H., SAKAI, K., AND NAKAMURA, O. 2008.Ear biometrics using jet space similarity.In Proc. of the IEEE Canadian Conference on Electrical and Computer Engineering CCECE. Niagara Falls, Canada.

[7] CUMMINGS, A., NIXON, M., AND CARTER, J. 2010. A novel ray analogy for enrollment of ear biometrics. In Proc. of the Biometrics: Theory, Applications, and Systems BTAS. Washington-DC, USA.

[8] L.Yuan and Z.-C.Mu. Ear detection based on skin color and contour information. In Proc. of ICMLC,2007, vol. 4,pages 2213–2217, 2007.

[9] A. Sana, P. Gupta, and R. Purkait. Ear biometric: A new approach. In Proc. of ICAPR, 2007, pages 46–50, 2007.

[10] Liu, H., and Liu, D.: 'Ear Segmentation Using Histogram Based K-means Clustering and Hough Transformation under CVL Dataset ', Proc. SPIE 7495,74952N, 2009.

[11] ZHOU, J., CADAVID, S., AND ABDELMOTTALEB, M. 2010.Histograms of categorized shapes for 3D ear detection. In Proc. of the Biometrics: Theory, Applications, and Systems BTAS. Washington-DC, USA.

[12] Balaji Vasan Srinivasan,"Ear Extraction from The Image of a Human Face".

[13] Michael Kass, Andrew WitkinadDemetriTerzopoulos, "Snakes: Active Contour Models",International Journal of Computer Vision, vol.1,no. 4, pages 321-331, 1987.

[14] Surya Prakash, UmaraniJayaraman and Phalguni Gupta, "A Skin-Color and Template Based Technique for Automatic Ear Detection".Advances in Pattern Recognition, ICAPR'09.

[15] S. Prakash, U. Jayaraman and P. Gupta," Ear Localization from Side Face Images using Distance Transform and Template

Matching" Proc. of IEEE Int'l Workshop on Image Processing Theory, Tools and Applications, IPTA 2008, pp. 1-8.

[16] S. M. S. Islam, M. Bennamoun and R. Davies,"Fast and Fully Automatic Ear Detection Using Cascaded AdaBoost"Proc. of IEEE Workshop on Application of Computer. 2008

[17] BanafsheArbab-Zavar and Mark S. Nixon, "Ear detection using method based on reduced Hough transform".

[18] H. Chen and B. Bhanu,"Hman Ear Detection From 3D Side Face Range Images".3D Imaging for Safety and Security, Springer-2007.vol.35,pages 133-155.