CONNECTED COMPONENT BASED TECHNIQUE FOR AUTOMATIC EAR DETECTION

Surya Prakash, Umarani Jayaraman and Phalguni Gupta

Department of Computer Science and Engineering,
Indian Institute of Technology Kanpur, Kanpur-208016, India.
E-mail: {*psurya,umarani,pg*}*@cse.iitk.ac.in*

ABSTRACT

This paper presents an efficient technique for automatic ear detection from side face images. The proposed technique detects ear by exploiting its inherent structural details and is rotation, scale and shape invariant. It can detect ear without any training or assuming prior knowledge of the input image. The technique is based on connected component analysis of a graph constructed using the edge map of the image and is evaluated on a data set consisting of 2361 side face images collected at IIT Kanpur. Ear detection results are found to be very good and speak for the efficiency and robustness of the technique. To show the accuracy of the detection, detected ears are used for recognition and results are compared with the same obtained when ear cropping in done manually.

Index Terms— Biometrics, Skin-segmentation, Connected Graph, Ear detection.

1. INTRODUCTION

Among the various physiological biometric traits such as face, ear and fingerprint etc., ear has received much attention in recent years as it has been found to be a reliable biometrics for human recognition [1]. The reasons behind the growth of ear's prominence as a ear biometric trait, in recent years, are its properties to be remarkably consistent and invariant to different facial expressions and aging. To make the ear based recognition systems automatic, ear need to be localized and cropped automatically from the input test image. In literature, there are very few techniques available for automatic ear detection. Further, these techniques are either semi-automatic and need some user intervention or fail to detect ears in many situations. Burge and Burger [1] have used deformable contours for ear detection which works well only when a small background is present. Alvarez et al. [2] have proposed an ear localization method using ovoid and active contours which requires an initial approximate ear contour as input and cannot be used in a fully automated ear recognition process. Ansari and Gupta [3] have presented an ear detection approach based on edges of outer ear helices. This technique solely relies on the parallelism between the outer helix curves, and hence may fail if the helix edges are not proper. Yuan and Mu [4] have

proposed an ear detection technique based on skin-color and contour information. It assumes the ear shape to be elliptical which may not always be true and may not help in detecting the ear in general. For example, boundary of an ear of triangular shape cannot be detected in this approach. In [5, 6], two template based approaches have been proposed for which ear template has to be resized (using size of the side face) to make its size suitable for the current test image. These approaches may fail when side face image contains some non-facial skin pixels. In [7], another attempt has been made for ear localization where hierarchical clustering of the edges is used. This technique works well only when scale of the side face images does not vary much.

This paper presents an efficient rotation, scale and shape invariant technique for automatic ear detection from side face images. The proposed technique is based on connected component analysis of a graph constructed using the edge map of the side face image. Organization of the paper is as follows. Sec. 2 presents the proposed technique. Sec. 3 discusses the experimental results. Paper is concluded in Sec. 4.
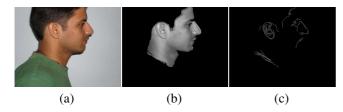
2. PROPOSED TECHNIQUE

Proposed technique is based on the fact that in the side face image, ear is the only part which contains maximum variation in the pixel intensities, as a result of which, this place is rich in edges. Proposed technique also relies on the fact that edges belonging to ear are curved. Following subsections provide details about the technique.

2.1. Preprocessing

2.1.1. *Skin Segmentation and Edge Detection*

Main objective of detecting skin regions in an image is to reduce the search space for the ear. Since, an ear exist in the skin region, there is no point looking for them in non-skin regions. Skin color model presented in [8] can be used for skin segmentation. A sample image, its skin segmented image and its edge map are shown in Fig. 1(a), Fig. 1(b) and Fig. 1(c) respectively.

Edge detection is performed on skin segmented image using Canny edge operator and a list of all the edges is obtained
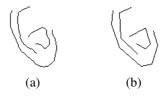
Fig. 1. Skin segmentation: (a) Input image, (b) Skin-segmented image (gray scale), (c) Edge map

Fig. 2. (a) Original edge image, (b) Edge image after approximating edges with line segments

by connecting points together into a lists of coordinate pairs. Wherever an edge junction is encountered the list is terminated and a separate list is generated for each of the branches. Let χ be the set containing all the edges present in the edge list. The spurious edges arose due to noise in the image are generally short in length and can be easily pruned out using edge length based criterion. Length threshold for pruning is automatically chosen proportional to the width of the skin area of the face. Let χ_l be the new set containing all edges obtained after length based edge pruning.

2.1.2. Approximation of Edges using Line Segments

All pixels present in an edge (belonging to set χ_l) may not be equally important and may not be necessary to represent the edge. So to remove the redundant pixels and to get compact representation, line segments are fitted to the edge. This eliminates all pixels from the edge which are not necessary and breaks every edge into a set of line segments. To fit line segments, procedure takes each array of edge points of set χ_l and finds the size and position of the maximum deviation from the line that joins the endpoints. If the maximum deviation exceeds the allowable tolerance, the edge is shortened to the point of maximum deviation and the test is repeated. In this manner each edge is broken down into line segments, each of which adhere to the original data with the specified tolerance. Fig. 2(b) shows an example of edge approximation by line segments for the edge image shown in Fig. 2(a). Let χ_{ls} be the set obtained after line segments fitting.

Curvature based Edge Pruning: In set χ_{ls}, each edge is represented using very few points. These points define the endpoints of the line segments. The edges which are linear (or almost linear) need only two points for their representation after line segment fitting. Since all the edges belonging to the ear contain some curvature, they would need more than two points for their representation. So the edges having two points cannot be a part of the ear and hence can be removed from the set χ_{ls}. This results a new edge set χ_c which can be defined as $\chi_c = \{e \mid e \in \chi_{ls} \text{ and } \gamma(e) > 2\}$ where $\gamma(e)$ is the total points used to approximate edge e using line segments.

2.2. Ear Localization

Ear localization first builds the edge connectivity graph, and subsequently uses it in finding the connected components in the graph for ear localization purpose.
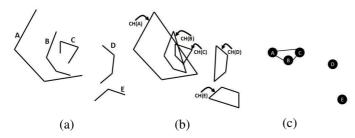
2.2.1. Building the Edge Connectivity Graph

The set χ_c can be used to define the edge map of the side face image. Let there be n edges in χ_c. Define the ith edge e_i in χ_c by a point p_i. Thus χ_c can be represented by a set P of points $p_1, p_2, .., p_n$. Against each edge e_i, define a convex hull[1] $CH(e_i)$. If two convex hulls $CH(e_i)$ and $CH(e_j)$ intersect, then it is assumed that point p_i and p_j of set P are connected through an edge of a newly defined graph $G = (V, E)$, where the set of vertices, V, and the set of edges, E, are given by
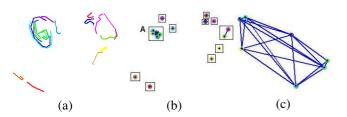
$$
\begin{aligned}
V &= \{p_i \mid p_i \in P\} \\
E &= \{(p_i, p_j) \mid CH(e_i) \text{ intersects } CH(e_j)\}
\end{aligned}
$$

V and E are the set of vertices and edges in graph G. In the ear part, mostly edges are convex and if we move from outer part of the ear towards inside, mostly outer ear edges contain inner ear edges. Because of this nature, convex hulls of the outer edges intersect the convex hulls of the inner edges. So this guarantees that the convex hull of an edge belonging to the ear cuts at least one other convex hull of the edge belonging to the ear. Convex hull based criteria to define connection between vertices (points) in a graph is expected to connect (directly or through another vertex) all the vertices belonging to the ear part with one another. Moreover, this criteria is able to define the connectivity between the vertices irrespective of the scale, which makes the proposed algorithm scale invariant. Property of one edge containing another is not true for the edges belonging to other parts of the side face image. So vertices corresponding to these edges mostly remain isolated in the graph. Fig. 3(a) and Fig. 3(b) show an example of synthetic edge map and edges convex hulls respectively. It can be easily observed from Fig. 3(b) that convex hulls of edges A, B and C intersect with each other. So the vertices corresponding to these edges are being connected to each other in

---

[1]Convex hull for an edge is a tightest convex polygon which includes all edge points.

Fig. 3. (a) Synthetic edge image, (b) Convex hulls of edges, (c) Graph for edge image shown in (a)

Fig. 4. (a) Edge map (colors used to differentiate one edge from another), (b) Graph obtained for the edge map with connected components labeling, (c) Magnified view of component A shown in (b)

the graph shown in Fig. 3(c). Points D and E are isolated in Fig. 3(c) since convex hulls of edges D and E in Fig. 3(b) do not intersect the convex hull of any other edge.

A simple criteria based on some distance metrics can be adopted to define the connectivity among the vertices. But the choice of distance makes the proposed technique scale dependent as the threshold for distance required to define the connectivity among the vertices may vary for the images of different scales.

2.2.2. *Ear Localization using Connected Components*

In an undirected graph, a connected component is defined as a maximal connected subgraph. After defining the graph for the edge map, connected components are computed for it. Graph components having a single vertex can be simply removed from the graph as they cannot represent the ear. In the remaining graph components, component having largest size is found out. Bounding box of the edges corresponding to this component is claimed to be the ear boundary. The reason behind considering the largest component as the ear is as follows. As discussed in the previous section, because of the nature of outer ear edges containing inner ones, the edges representing ear keep good connectivity with each other in the graph. This is not true for other edges and hence component representing the ear turns into the largest.

Fig. 4 shows an example of edge map and its corresponding graph. Fig. 4(a) shows the edge image obtained after

Fig. 5. Ear detection results for Data Set 1(row 1), Data Set 2(rows 2-3) and Data Set 3(row 4)

approximation of edges with line segments of the edge image shown in Fig. 1(c). A graph, shown in Fig. 4(b) is constructed for this edge image. Connected components of this graph are computed and shown by rectangles in Fig. 4(b). Magnified view of the component A is shown in Fig. 4(c). It can be observed that except the component representing ear, all other components are very small in size (mostly contain one vertex).

3. EXPERIMENTAL RESULTS

The proposed technique is validated using three data sets collected at IIT Kanpur. Data Set 1 consists of 490 side face images containing frontal ear views collected from 168 individuals. Data Set 2 consists of 801 side face images collected from 89 individuals, 9 images per subject for various rotations and scales. Images in this data set contain frontal ear views taken at 3 positions, (1) when person is looking straight, (2)&(3) when person is looking approximately 20^0 down and 20^0 up respectively. At all these 3 positions, images are captured at 3 different scales. Data Set 3 is a tedious set where 1070 images are captured at vertical rotations of -40^0, -20^0, 0^0, 20^0 and 40^0 of the side face from 107 individuals. Two images per pose (angle) are obtained, producing 10 images per subject.

Fig. 5 shows the ear detection results for all three data sets. Data Set 1 contains normal frontal ear images. Data Set 2 is used to demonstrate the rotation and the scale invariance of the proposed technique. Ears of different scales and rotations are efficiently detected from the side face images in this case without any user intervention and the change of parameters. The proposed technique has also detected ears successfully for the images of Data Set 3 even for the extreme poses $(-40^0$ and $+40^0)$ where ear localization is very challenging. Since the proposed technique uses edge property of the ear and not the ear shape, it is shape invariant.
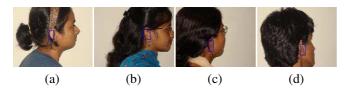
(a)   (b)   (c)   (d)

Fig. 6. Partially detected ears

Data Set	# of Test Images	%age accuracy	
		Exact ear boundary	15% neighboring pixels allowed
Set 1	490	94.01	95.88
Set 2	801	93.20	94.73
Set 3	1070	90.52	91.11

Table 1. Percentage accuracy for various data sets

Table 1 shows the percentage accuracy of ear detection for the proposed technique. Values in the third column present the accuracies when a close-fitting rectangular boundary is considered as the correct detection whereas fourth column shows the same when ear localization is allowed to include 15% neighboring pixels inside the detected ear boundary. For Data Set 1, accuracy is the highest as it contains the frontal ear images. Accuracy for the Data Set 2 is comparable with that of Data Set 1, though it contains scale and rotation variations. The reason behind this is that the proposed technique uses structural details of the ear which do not change with rotation and scale. Data Set 3 gives less accuracy because of the fact that when face rotates with respect to vertical axis, availability of the structural details of the ear degrades as camera moves away from the frontal view. Localization method could partially detect or could not detect the ears in some cases, specially for the images where ear was occluded due to hair (Figs. 6(a), 6(b), 6(c)) or ear structure became invisible due to more vertical rotation (Fig. 6(d)). To see the correctness of the detection, automatically cropped ear images of Data Set 3 are used for verification using the technique presented in [9]. It is found that the verification performance for automatic ear detection is better than the same obtained for manual ear cropping because manual cropping contains some human error in correctly localizing the ear boundary. This is evident from the ROC curves shown in Fig. 7.

4. CONCLUSION

This paper proposes an efficient technique for automatic ear detection using structural details of the ear. The technique detects the ear in the side face image without any user interaction and can be deployed in an automatic ear based biometric system. It is able to detect ears of different scales, rotations, and shapes efficiently without any user intervention. The proposed technique is tested on a database containing 2361 side
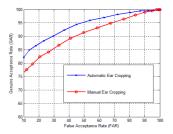

Fig. 7. ROC curves for Data Set 3

face images collected at IIT Kanpur. To show the correctness of the detection, detected ears are used for recognition. It is found that the recognition performance for automatically detected ears is better than the same obtained for manually cropped ears because manual localization contains some human error in localizing ear boundary correctly. The technique can be easily extended for the detection of multiple ears.

5. REFERENCES

[1] M. Burge and W. Burger, "Ear biometrics in computer vision," in *Proc. of ICPR' 00*, 2000, vol. 2, pp. 822–826.

[2] L. Alvarez, E. Gonzalez, and L. Mazorra, "Fitting ear contour using an ovoid model," in *Proc. of ICCST' 05*, 2005, pp. 145–148.

[3] S. Ansari and P. Gupta, "Localization of ear using outer helix curve of the ear," in *Proc. of ICCTA' 07*, 2007, pp. 688–692.

[4] Li Yuan and Zhi-Chun Mu, "Ear detection based on skin-color and contour information," in *Proc. of ICMLC' 07*, 2007, vol. 4, pp. 2213–2217.

[5] Surya Prakash, J. Umarani, and P. Gupta, "Ear localization from side face images using distance transform and template matching," in *Proc. of IEEE Int'l Workshop on Image Proc. Theory, Tools and App., IPTA' 08*, Sousse, Tunisia, 2008, pp. 1–8, IEEE Xplore.

[6] Surya Prakash, J. Umarani, and P. Gupta, "Ear localization using hierarchical clustering," in *SPIE Int'l Defence Security and Sensing Conference, vol. 7306*, Orlando, Florida, 2009, SPIE.

[7] Surya Prakash, J. Umarani, and P. Gupta, "A skin-color and template based technique for automatic ear detection," in *Proc. of ICAPR' 09*, India, 2009, pp. 213–216.

[8] J. Cai and A. Goshtasby, "Detecting human faces in color images," *Image and Vision Computing*, vol. 18, no. 1, pp. 63–75, 1999.

[9] G. S. Badrinath and P. Gupta, "Feature level-fused ear biometric system," in *Proc. of ICAPR' 09*, Kolkata, India, 2009, pp. 197–200.