International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-3, Issue-1)

Research Article

January 2014

Automatic Detection of Ear from Side Face Image Using Template Matching

Ms. Aradhana Singh*
CSE, CSIT Durg
India

Mr. Vikas Sahu
Mechanical Engg., CSIT Durg
India

Mr. Rohit Raja
CSE, SSCET, Bhilai
India

Abstract-:

This paper proposes an efficient and template based technique for automatic ear detection in a side face image. The technique first separates skin regions from nonskin regions and then searches for the ear within skin regions. Ear detection process involves three major steps. First, Skin Segmentation to eliminate all non-skin pixels from the image, second Ear Localization to perform ear detection using template matching approach, and third Ear Verification to validate the ear detection using the Zernike moments based shape descriptor. To handle the detection of ears of various shapes and sizes, an ear template is created considering the ears of various shapes and resized automatically to a size suitable for the detection.

***Keywords**: Skin-segmentation, distance transform, Edge Detection, shape descriptor, Zernike moments*

I. INTRODUCTION

Biometric systems have become very essential components in almost all security aspects. These systems perform the recognition of a human being based on physiological and behavioural characteristics. Physiological characteristics are related to the shape of the body. Biometric traits such as face, fingerprint, iris, hand geometry fall under this category. Behavioural characteristics are related to behaviour of a person. Signature, voice, character strokes etc. are some of the biometric traits which fall under this category. Among the various physiological traits, ear has gained much attention in recent years as it has been found to be a good and reliable biometrics for human verification and identification [1]. Reason behind the ear biometrics gaining popularity is that ears are remarkably consistent. Unlike faces, they do not change shape with different expressions or age, and remain fixed in the middle of the side of the head against a predictable background [2]. To automate ear based recognition process, automatic ear localization is necessary but at the same time detection of the ear from an arbitrary side face image is a challenging problem. This is because of the fact that ear image can varying appearance under different viewing and illumination conditions. To detect ears of different size, they maintain template of various sizes. Ear Templates approach is created manually by cropping the ears of different size from the side face images and template matching is performed for matching purpose. In real scenario, ear occurs in various sizes and the pre-estimated templates are not sufficient to handle all the situations. Further, detection of ear using templates of various sizes and then selecting best detection is a very computation intensive task.

II. PRELIMINARIES

This section discusses some of the basic techniques which are required in developing the proposed ear localization model[3]. Firstly a skin segmentation is used for separating skin region and non skin region [4][5][6]. Next the details of distance transform used for getting the ear template in their proper shape [7]. Last discusses Zernike moments based shape descriptor which is used for ear verification in the proposed technique [8].

Color based skin segmentation and edge detection: The technique first separates skin regions from non-skin regions and then searches for the ear within skin regions. 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 can be used for skin segmentation. A sample image, its skin segmented image and its edge map are shown in Fig.2.1 (a), Fig. 2.1(b) and Fig. 2.1(c) respectively. Edge detection [12] is performed on skin segmented image using canny edge operator and a list of all the edges is obtained 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.

(a) (b) (c)

Fig. 2.1. Skin segmentation: (a) Input image, (b) Skin segmented image (gray scale), (c) Edge map

All pixels present in an edge 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 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 adheres to the original data with the specified tolerance. Fig. 2.2(b) shows an example of edge approximation by line segments for the edge image shown in Fig. 2.2(a).

(a) **(b)**

Fig. 2.2. (a) Original edge image, (b) Edge image after approximating edges with line segments

A. ***Distance Transform:***

Distance transform [7] is computed normally for a binary image where each pixel in the binary image is assigned a number that is the distance between that pixel and the nearest nonzero pixel of binary image. Formally, for a binary image I, its distance transform image I^d is given by the following equation:

$$I^d(x, y) = \begin{cases} 0, & \text{if } (x;\ y) \text{ is a non zero pixel} \\ d, & \text{otherwise} \end{cases} \quad \text{.....Eq(1)}$$

Where, *d* is the Euclidian distance of pixel (*x, y*) to the nearest non-zero pixel. Fig. 2.3(b) shows an example of distance transform image for the edge image shown in Fig. 2.3(a).

(a) **(b)**

Fig. 2.3.(a) Edge map, (b) distance transform of (a)Zernike Moments:

Among the family of moment invariants, Zernike moments [9][10][11] are one of the most commonly used feature extractor and have been used in variety of applications. In the proposed technique, Zernike moments based shape features are used for ear verification. Zernike moments provide non-redundant shape representation because of their orthogonal basis.

III. PROPOSED METHODOLOGY

In this section, proposed ear localization technique is discussed in detail [7]. First pre-processing steps are discussed followed by ear localization and ear verification. In pre-processing step, skin areas of the input side face image are segmented and processed for edge computation. Ear localization step creates a suitable size of ear template and performs the ear localization. Localized ears are verified using Zernike moment based shape descriptor at ear verification step. Fig. 3.1 shows complete flow chart of the proposed technique.

A. ***Preprocessing:***

1) Ear Skin Segmentation: The first step of the proposed technique is skin segmentation and detection of skin regions in the image. Main objective of detecting skin regions is to reduce the search space for the ear. Naturally, ears are in skin region and there is no point looking for them in non-skin regions.

2) Ear Edge Detection: After skin segmentation, edge map of the skin regions is obtained using canny edge detector [12] and used for further processing. This edge map contains many spurious edges which may not belong to the ear. These edges are pruned using a criterion based on the edge length and curvature.

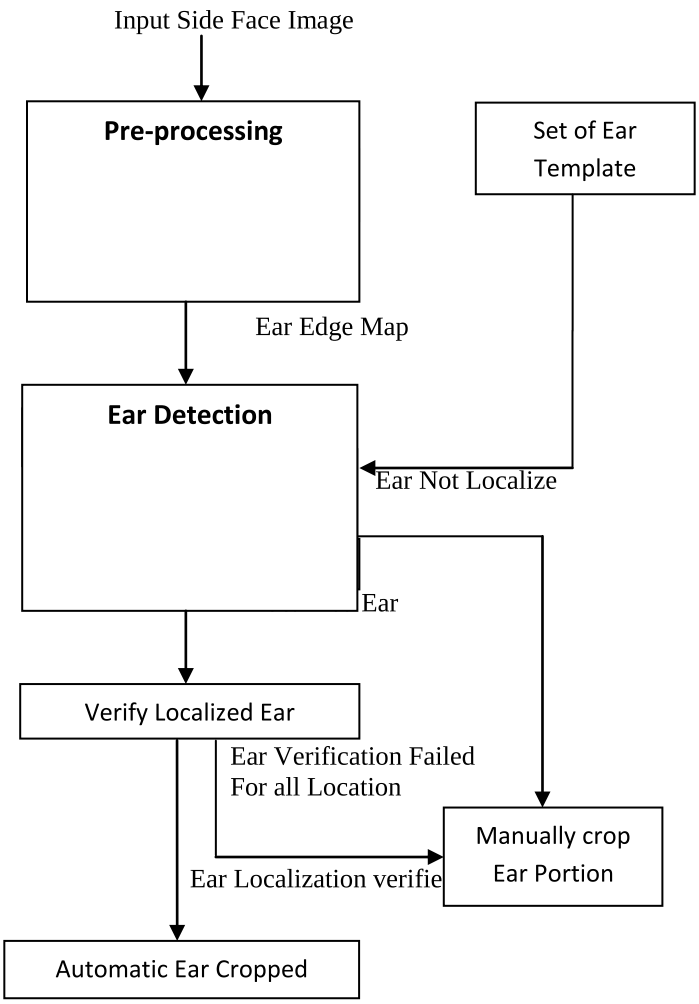

Fig. 3.1 Flow chart of the proposed technique

A. ***Ear Detection:***

After computing the edge map of the skin regions is ear localization. Ear localization involves four steps: ear template creation, resizing of the ear template, distance transform computation and localization of the ear. Ear template creation is an off-line process and carried out beforehand. Resizing of the off-line created ear template is done in accordance with the size of each input image and resized template is used for ear localization.

1) Ear Template Creation: For any template based approach, it is very much necessary to obtain a template which is a good representative of the data. In the proposed technique, ear template is created by averaging the intensities of a set of ear images. The set includes ear images of men and women. These images are selected from the ear database. For the creation of ear template in the proposed technique, all types of the ears are considered to obtain a good template. Ears from the selected images are cropped and resized to same size. The ear template T is estimated by taking the mean of the pixel values of all ear images and formally defined as follows:

$$T(i; j) = \frac{1}{N}\sum_{k=0}^{N} E_k(i,j) \qquad ......Eq(2)$$

Where N is the number of ear images used for ear template creation and E_k is the k^{th} ear image. $E_k(i,j)$ and T(I, j) represent the pixel values of the $(I, j)^{th}$ pixel of E_k and T respectively.

2) Ear Template Resizing: Resizing of ear template is an important step. As we know, ear in a side face image may occur in different size depending on the distance of the camera from the person. Not only this, some people may have small or large ear too. So to handle the detection of ears of various sizes, ear template need to be resized to make its size appropriate for the detection of ear from an image.

3) Compute Distance Transform: In place of using directly the edge information or the gray level images of face and ear template, distance transform is used in localization as it contains optimum meaningful information about the neighborhood of a pixel. Distance transform has been successfully used in many computer vision applications

[13][14][15][16]. After getting the distance transforms of the edge maps of the face and ear template, cross-correlation based searching is applied to locate the ear from the side face image.

4) *Localization Ear Portion:* Once the distance transform (*Td*) of resized ear template edge image is created, it is used for the localization of the ear from the side face image. To search an ear in the image, *Td* is moved over the distance transform of the face skin edge image (*Id*) and Normalized Cross correlation Coefficient (NCC) is computed [17][18] at every pixel. NCC at point (*x; y*) is defined as follows:

$$\text{NCC(x,y)} = \frac{\sum_{u,v}[I^d(u,v)-\bar{I}^d_{x,y}][T^d(u-x,v-y)-\bar{T}^d]}{\sqrt{\sum_{u,v}[I^d(u,v)-\bar{I}^d(x,y)]^2 \sum_{u,v}[T^d(u-x,v-y)-\bar{T}^d]^2}} \quad \text{.......Eq(3)}$$

Where sum is performed over *u; v* under the window containing the template *Td* positioned at (*x,y*). $\bar{I}^d_{x,y}$ and $\bar{T}^d$ are the average of brightness values of the portion of the target image under the template and template image respectively. Normalized cross-correlation coefficient estimates the degree of linear dependence between the corresponding pixel brightness values being compared. Since the cross-correlation coefficients lie between -1.0 and 1.0, match ratings also lie between -1.0 and 1.0. When the match rating is typically above a preestimated threshold we accept the hypothesis that an ear exists in the region. Otherwise, we reject the hypothesis. Value of normalized cross-correlation coefficient closer to 1 indicates a better match. This step accepts all the points having NCC value greater than the threshold as probable locations of the ear. These locations are sorted in non-increasing order according to their NCC values. Ear verification is applied to all these detection to validate the existence of the ear.

B. *Verify Localized Ear:*

To determine whether a detected ear is actually an ear or not, shape based ear verification is performed. Since small set of lower order Zernike moments can characterize the global shape of an object effectively [18], these moments are used for ear shape representation in the proposed technique. As Zernike moments are complex number, only their magnitudes are considered for the shape representation. Similarity between the two sets of Zernike moments (one for template and another for detected ear) is estimated to validate the claim. To measure the similarity, Euclidian distance between the two sets of Zernike moments is estimated as follows:

$$\text{Distance} = \sqrt{\sum_{i=1}^{L}(\,|\,M_i^T\,| - |\,M_i^E\,|\,)^2} \quad \text{.....Eq(4)}$$

where $\sum_{i=1}^{L}(M_i^T)$ and $\sum_{i=1}^{L}(M_i^E)$ are the L Zernike moments used to represent the shape of ear template and detected ear respectively.

IV. RESULT

Ear Templates approach are created manually by cropping the ears of different size from the side face images and template matching will performed for matching purpose. In real scenario, ear occurs in various sizes and the pre-estimated templates are not sufficient to handle all the situations. Further, detection of ear using templates of various sizes and then selecting best detection is a very computation intensive task.

V. CONCLUSION

This paper presents a technique for automatic ear localization from side face images using distance transform and template matching approach. The technique is able to detect ears of different size and shape automatically from the side face image. It first segments the skin and non-skin pixels in the image and then computes the edge map of the skin areas. Edge map is done to remove the spurious edges which can not belong to ear for this canny edge detection method can be adopted. Distance transform images of the edge maps of face skin region and ear template can use in the ear localization process. For ear verification Zernike moments based shape descriptor technique is can adopted.

Lower order Zernike moments of the distance transforms of the edge images of the ear template and the detected ear will be estimated and similarity distance between them will be calculated using Eqn. 4. To get the distance transform image of the detected ear at point (x, y), template sized image will be cropped from the input distance transform image of side face image keeping the point (x, y) at the centre of the template. If the value of distance is less than a pre-estimated threshold, detection is accepted, otherwise it is rejected.

REFERENCES

[1] B. Bhanu and H. Chen, *Human Ear Recognition by Computer*. Springer, 2008.

[2] Ping Yan and Kevin W. “Bowyer, Biometric Recognition Using 3D Ear Shape”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 29, NO. 8, AUGUST 2007.

[3] Surya Prakash, UmaraniJayaraman and Phalguni Gupta, “Ear Localization from Side Face Images using Distance Transform and Template Matching”,Proceedings of IEEE Int'l Workshop on Image Processing Theory, Tools and Applications, IPTA 2008, Sousse, Tunisia, pp. 1-8, Nov 2008.

[4] Takashi Ida, *Member, IEEE*, and Yoko Sambonsugi, "Image Segmentation and Contour Detection Using Fractal Coding", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 8, NO. 8, DECEMBER 1998.

[5] J. Cai, A. Goshtasby, and C. Yu, "Detecting human faces in color images," in *International Workshop on Multi-Media Database Management Systems, IWMDMS' 98*, pp. 124–131, 1998.

[6] J. Cai and A. Goshtasby, "Detecting human faces in color images," vol. 18, pp. 63–75, December 1999.

[7] H. Breu, J. Gil, D. Kirkpatrick, and M. Werman, "Linear time Euclidean distance algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 529–533, 1995.

[8] JrromeRevaud, Guillaume Lavoué and AtillaBaskurt, "Improving Zernike Moments Comparison for Optimal Similarity and Rotation Angle Retrieval", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. XX, NO. XX, APRIL 2008

[9] K. W.-Y. and K. Y.-S. *Signal Processing: Image Communication*, vol. 16, no. 1, pp. 95–102, 2000.

[10] S. Belkasim, E. Hassan, and T. Obeidi, "Explicit invariance of Cartesian zernike moments," *Pattern Recognition Letters*, vol. 28, no. 15, pp. 1969–1980, 2007.

[11] M. Liu, Y. He, and B. Ye, "Image zernike moments shape feature evaluation based on image reconstruction," *Geo-Spatial Information Science*, vol. 10, no. 3, pp. 191–195, 2007.

[12] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[13] D. W. Paglieroni, G. E. Ford, and E. M. Tsujimoto, "The positionorientation masking approach to parametric search for template matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 740–747, 1994.

[14] P. Tsang, P. Yuen, and F. Lam, "Classification of partially occluded objects using 3-point matching and distance transformation," *Pattern Recognition*, vol. 27, no. 1, pp. 27–40, 1994.

[15] M. Sanjay, S. Das, and B. Yegnanarayana, "Robust template matching for noisy bitmap images invariant to translation and rotation," in *Indian Conference on Computer Vision, Graphics and Image Processing, December 21-23, 1998*, (New Delhi, INDIA), pp. 82–88, 1998.

[16] D.-J. Lee, J. Archibald, X. Xu, and P. Zhan, "Using distance transform to solve real-time machine vision inspection problems," *Machine Vision and Applications*, vol. 18, no. 2, pp. 85–93, 2007.

[17] J. K. Aggarwal, L. S. Davis, and W. N. Martin, "Correspondence processes in dynamic scene analysis," in *Proceedings of the IEEE*, vol. 69, pp. 562–571, 1981.

[18] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," vol. 10, no. 4, pp. 496–513, 1988.

BIOGRAPHY

[1] Ms. ARADHANA SINGH
BE (Computer Science Engineering) from CSIT Durg, M.Tech(Computer Science Engineering) from SSCET, Bhilai Assistant Prof., Dept. of Computer Science Engineering, CSIT Durg
aradhana2907@gmail.com

[2] Mr. VIKAS SAHU
BE (Hons.) (Mechanical Engineering) from BCE, RGPV Bhopal Assistant Prof., Dept. of Mechanical Engineering, CSIT Durg
vikassahu87@yahoo.co.in

[3] Mr. ROHIT RAJA
BE, M.Tech (Computer Science Engineering), Ph.D (Pursuing) Senior Assistant Prof., Dept. of Computer Science Engg, SSCET Bhilai
rohitraja4u@gmail.com