

Oval Shape Detection and Support Vector Machine Based Approach for Human Ear Detection from 2D Profile Face Image

Kinjal V. Joshi

*G H Patel College of Engineering and Technology
Vallabh Vidyanagar, Gujarat, India
kinjalvjoshi@gmail.com*

Abstract

Ear is a new class of relatively stable biometrics which is not affected by facial expressions, cosmetics, eye glasses and aging effects. Ear detection is the first step of an ear recognition system, to use ear biometrics for human identification. In this paper, an oval shape detection based approach is presented for ear detection from 2D profile face images. The correctness of the detected ear is verified using support vector machine tool. The experimental results prove the effectiveness of the presented method.

Keywords: *ear detection, ear verification, oval shape detection, support vector machine*

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. 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 an image is a side face image, only the ear is unique feature from which a 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. Ear recognition is useful for person identification when only an image of a side face is available.

The number of recent researches [1-4] show that face recognition is possible and effective for side faces by detecting and recognizing components such as ears.

The rest of the paper is organized as follows. The proposed method for ear detection is described in Section 2. Ear verification by SVM tool is described in Section 3. The implementation and verification results are shown in Section 4. Conclusion is discussed in the Section 5.

2. Proposed Method

A block diagram for the proposed method is shown in the Figure 1. As shown in block diagram of figure 1, first skin portion is segmented from input image then nose tip is detected. After that the region of face which has probability to contain ear is identified. Then oval shape is detected from that region because ear shape is almost similar to oval shape. A rectangle is drawn around the detected oval shape that shows the ear.

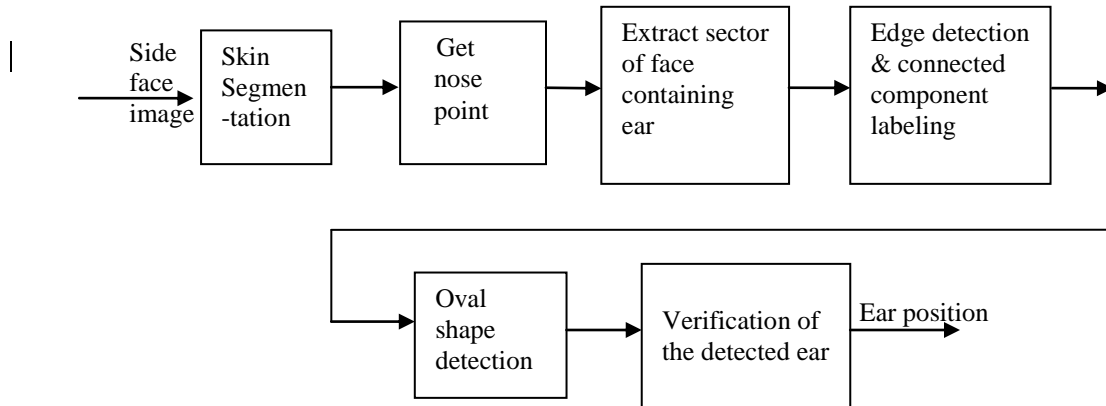


Figure 1. Block Diagram of the Proposed Method

2.1 Skin Segmentation

2.1.1. Color Space Selection: Here the input image is 2D side face color image. From side face image skin portion is separated out using the method suggested in [5, 6]. The goal is to remove the maximum number of non-face pixels from the images in order to narrow the focus to the remaining predominantly skin-colored regions. For this purpose we need to select appropriate color space from the wide variety of choices such as RGB, HSV, CMYK, YCbCr etc. From these, RGB (red-green-blue), HSV (hue-saturation-value) and YCbCr are widely used [7]. In the RGB model, each of the three components may exhibit substantial variation under different lighting environments. The results of YCbCr and HSV are more robust to lighting variations because in both the color spaces, color classification is done using only pixel chrominance. It is expected that skin segmentation may become more robust to lighting variations if pixel luminance is discarded and it is also verified by results. Here HSV color space is preferred for color classification because of its similarities to the way human tends to perceive color. It decouples the chrominance information from the luminance information. Thus we can only focus on the hue and the saturation component.

2.1.2. Setting Threshold and Binary Image Creation: After choosing the suitable color space, the next step is to separate the skin colored region from the given input image. For this, the best technique is to apply threshold. To find the appropriate values for threshold, the many face images in HSV color space are examined and found out some specific ranges of the components for skin color.

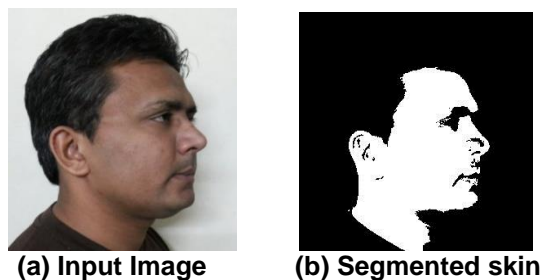


Figure 2. Skin Segmentation

When these threshold values are applied to the input image, the new binary image is formed in which the portions satisfying the conditions is made white and the remaining portion is made black. Figure 2 shows result of skin segmentation. This is a binary image created from a RGB input image.

2.2 Nose Tip Detection

Nose tip is detected by taking first white pixel from segmented skin image. The skin segmentation might not perfect and so any first non-black pixel in the binary image might not always represent a skin. So in order to obtain the nose point, the first non-black pixel is noted as a skin point only if it is surrounded by non-black pixels as well. In this way the first non-black pixel in each column is noted.

Once a vector of the pixel locations is available, the minimum position is noted. Again, the surrounding pixels are examined in order to ensure that the identified point is a skin pixel. Figure 3 shows the identified nose point.



Figure 3. Nose Tip Detection

2.3. Extraction of Sector of Face Containing Ear

Once the nose point is identified, a rectangular region is detected from the face with the width and height ratio 1:2 based on distance estimation between nose tip and ear. The probability of finding the ear in this region is high. Once this region is detected, the ear can be located by looking for oval shape in the region. Figure 4 shows the detected region.



Figure 4. Region Extraction

2.4. Edge Detection and Connected Component Labeling

Here edge is detected from input side face image. First of all, color image is converted to gray scale image. For edge detection Sobel, Robert, Prewitt, Canny, Laplacian of Gaussian, Zero cross *etc.*, methods are available [8]. The Canny method finds edges by looking for local maxima of the gradient of intensity image. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges and

includes the weak edges in the output only if they are connected to strong edges. Therefore this method is less likely when the input image is to be fooled by noise, and more likely to detect true weak edges.

The experimental results using canny edge detector are shown in Figure 5. The edges of ear have been detected in extracted region which has higher probability to contain ear.



Figure 5. Edge Detection

After edge detection, connected component labeling is applied, so extra unconnected small edges are removed. Here an 8-connected neighborhood is used to label a pixel. The result of connected component labeling is shown in Figure 6.



Figure 6. Connected Component Labeling

2.5. Oval Shape Detection

By applying connected component labeling small edges are removed. Now on output of connected component labeling MajorAxisLength and MinorAxislength properties of labeled region are used to detect oval shape. Figure 7 shows output of it. Oval shape in probable area of ear definitely contains ear edges. Extra edges except ear can be removed using this operation. The rectangle size of 50×32 is drawn around detected oval shape in probable area of ear. Figure 8 shows final output.

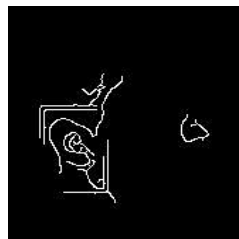


Figure 7. Oval Shape Detection



Figure 8. Final Output

3. Ear Verification Using SVM Tool

A Support Vector Machine (SVM) [9, 10] is one such method that can perform pattern recognition; its use though is not limited to this one application. While most classifiers work on designing rules that will place decision boundaries between data belonging to different classes, SVM goes a step ahead and designs “Support Vectors” such that the data belonging to different classes is now separated by a region rather than just a hyper plane. Thus, the distinction between classes is made more obvious, in an intuitive sense. SVM tries to minimize the upper bound on the expected risk. The generalization ability of SVM is more compared to other classifiers. Moreover, the time required for developing a SVM model is much lower because of the need for fewer data in training. Here prediction of unseen data is done by considering only the Support Vectors and hence presence of outliers in the training set may not influence the generalization accuracy of SVM. Recent work [10] suggests that the structure of SVM can be pre-decided using training data analysis. Hence, developing a SVM based model of a component is much simpler.

The choice of the appropriate Kernel parameter of SVM for a specific application is again problem dependent and often a difficult task. Generally it is expected that if the data is known to be not linearly separable, a non-linear kernel performs better than the one based on a linear kernel.

In this work libsvm-2.91 tool [11] is used for ear verification by classification. The model is created by features which are generated using detection results. The different kernels like Linear, Polynomial, Radial Basis Function, Sigmoid and accordingly parameters i.e. degree, gamma, cost co-efficient are adjusted manually to improve classification (verification) accuracy.

4. Experimental Results

4.1 Data Acquisition

In this work CVL (CVL is library for image and data processing using graphics processing units (GPUs)) dataset [12] is used, which contains total 114 persons with 7 images of each. Resolution of each image is 640 x 480. All the Images are in JPEG format captured by Sony Digital Mavica under uniform illumination and with projection screen in background. The other dataset of 50 images of side faces with dynamic lighting condition and screen resolution of 1656 × 1596 is also produced.

4.2. Ear Detection Results

Figure 9 shows examples of true detection for CVL dataset and Figure 10 shows examples of true detection for general dataset.



Figure 9. True Detection for CVL Dataset



Figure 10. True Detection for General Dataset

This method has failed in some cases, especially for the images where face is oriented with some angle because in that image nose tip cannot be detected correctly. For some people the distance between nose tip and ear is different than the generally estimated distance so there are chances for false detections or partial true detections. Figure 11 shows examples of false detection for CVL dataset and Figure 12 shows examples of false detection for general dataset. Here ear is detected partially because of variation of distance between nose tip and ear comparatively generally estimated distance.

The accuracy is calculated as: $(\text{genuine localization} / \text{total sample}) \times 100$. Here accuracy is found to be 94% for the CVL dataset and 96% for general dataset. The average time to detect an ear from a side face image using this method is 1.8 seconds with Matlab environment.



Figure 11. False Detection for CVL Dataset



Figure 12. False Detection for General Dataset

4.3 Ear Verification Results

Here for ear verification libsvm-2.91 tool is used. The rectangle drawn to show detection is size of 50×32 . So, 1600 pixel intensity values are in one feature vector. The feature vectors are created from the result of the method and model is created from these feature vectors.

Table 1. Effect of SVM Tool Parameters on Accuracy

Kernel type	Co efficient	Gamma	Degree	Cost	Accuracy (%)
Linear	-	-	-	-	84.48
Polynomial	-	-	3	0.01	81.03
				0.1	
				10	
				100	
Polynomial	-	-	2	0.01	82.76
				0.1	
				10	
				100	
Radial basis function	0.01	0.01	-	-	77.59
	0.01	Default			
	10	0.01			
	0.5	Default			
Sigmoid	-	-	-	-	75.80

The different parameters are set for accuracy measurement. Again here the feature vectors used in testing dataset are not used in training dataset. The parameter and accordingly

accuracy is shown in Table 1. The maximum accuracy is found 84.48% using linear kernel. In table 1 we can see that value of degree parameter for polynomial kernel affect accuracy while different values of Gamma and co efficient for RBF kernel do not affect accuracy.

5. Conclusion

In the proposed method the nose tip detection is very important because in this method ear detection is based on distance estimation between nose tip and ear. So skin segmentation must be done properly. For this method accuracy of skin segmentation is 100% because of HSV color space. For some people the distance between nose tip and ear differs than generally estimated distance, so there may be cases for partial ear detection. Again, this method also fails if ears are heavily occluded due to hair. This method sometimes fails, if side face is oriented with some angle. Accuracy of ear detection for CVL dataset and general dataset is 94% and 96% respectively. Verification of results using SVM gives 84.48% accuracy with linear kernel. Experimental results on real side face images demonstrate the effectiveness of the proposed approach.

Reference

- [1] B. V. Srinivasan "Ear Extraction From the Image of a Human Face", University of Maryland, College Park, Available at: http://www.umiacs.umd.edu/~balajiv/Project/enee630_2.pdf.
- [2] H. Chen and B. Bhanu, "Hman Ear Detection From 3D Side Face Range Images", 3D Imaging for Safety and Security, vol.35, Springer (2007), pp.133-155.
- [3] 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).
- [4] S. Prakash, U. Jayaraman and P. Gupta, "A Skin-Color and Template Based Technique For Automatic Ear Detection", Proc. Int'l Conf. Advances in Pattern Recognition, ICAPR' 09, Feruary (2009), pp. 213-216.
- [5] S. L. Phung, A. Bouzerdoun , D.Chai, " Skin Segmentation Using Color Pixel Classification: Analysis and Comparison", IEEE trans. on Pattern Analysis and Machine Intelligence, vol.27,no.1, January (2005), pp.148-154.
- [6] V. Vezhnevets, V. Sazonov, A. Andreeva: "A Survey on Pixel-Based Skin Color Detection Techniques", International Journal of Expert Systems With Applications, vol.3, April (2009), pp 4497-4507.
- [7] J. D. Foley, A. Dam, S. K. Feiner, J. F .Hughes "Computer Graphics Principles and Practice", Pearson Education (2006).
- [8] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", Pearson Education (2002).
- [9] N. C. Chauhan, Y. K. Roy, Arun Kumar, A. Mittal, and M. V. Kartikeyan " SVM-PSO Based Modeling and Optimization of Microwave Components", Frequenz 62 (2008).
- [10] J. A. Suykens, J. Vandewalle, and B. D. Moor, "Optimal Control by Least Squares Support Vector Machines", International Journal of Neural Networks, vol.14, no.1, Jan (2001), pp. 23-35.
- [11] C. C. Chang and C. J. Lin, "LIBSVM – A Library for Support Vector Machines", (2001). Software available at: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [12] <http://www.lrv.fri.uni-lj.si/~peterp/FaceDB.zip>.

Author



Kinjal V. Joshi, she is working as an Assistant Professor in Computer Engineering department of G H Patel College of Engineering & Technology, Vallabh Vidyanagar, Gujarat, India since 2007. She has done M.E. in Computer Engineering from Sardar Patel University. Her area of interest is Image Processing. She has published various research papers in area of Image Processing.