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

Human Ear localization in 2D profile image

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, eye glasses and aging effects. To use ear biometrics for human identification, ear detection is the first step of an ear recognition system. In this paper, an improved template matching approach is presented for ear detection from 2D profile image. Here extra skin portion like neck is removed and template is moved over the exact side face, so required detection time decreases and accuracy increases. The correctness of the detected ear is verified using euclidean distance and support vector machine tool. The experimental results prove the effectiveness of the presented method.*

Keywords: *Ear detection, skin-segmentation, euclidean distance, 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 eyes, and 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. This paper only deals with ear detection from a side face image, which is the first step towards human detection system using only side face images. Human ear detection is the first task of a human ear recognition system and hence its performance significantly affects the overall quality of the system. Ear recognition is useful for person identification when an image of a side face is available.

Recently, a number of researches [1, 2, 3,4] showed 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 is described in section 2. The implementation results and verification results are shown in section 3. A conclusion is shown at the end.

2. PROPOSED METHOD

Here basic method is based on template matching but pre-processing of input image and verification of detected ears using support vector machine tool libsvm-2.91 is included to improve accuracy. As pre-processing extra skin portion of input image except side face is removed so template matching takes less time to locate ear. Method has following steps.

2.1 Side Face Detection

Side face detection includes following steps.

- Skin Segmentation:

First step for finding side face is skin detection. For that an appropriate color space should be chosen. RGB, HSV and YCbCr are the most widely used color spaces [5-8]. RGB model gives good quality of

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

results but Skin colors are sensitive to the lighting condition. In the RGB space, each of the three components may exhibit substantial variation under different lighting environments. In YCbCr and HSV color spaces, Colors are specified in terms of luminance and chrominance .In this case, color classification is done using only pixel chrominance because 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. For different components threshold is set for skin detection. By seeing results of both color spaces HSV color space is chosen because HSV model gives good result as compared to the YCbCr model and the whole skin colored part is segmented. The result of skin segmentation is shown in figure 1.

(a) (b)

Figure 1: (a) Input image (b) Skin segmentation

- Morphological Processing:

Morphological image processing techniques are useful for extracting image components that are useful in representing and describing region shapes. Skin color segmentation does a good job of rejecting non-skin colors from the input image. However, the resulting image has quite a bit of noise and clutter. A series of morphological operations like morphological opening and hole filling are performed to clean up the image. The goal is to end up with a mask image that can be applied to the input image to yield skin color regions without noise and clutter.

First of all morphological opening is performed to remove very small objects from the image while preserving the shape and size of larger objects in the image. Then whole filling is done to keep the side faces as single connected regions in anticipation of a second much larger morphological opening. Otherwise, the mask image will contain many cavities and holes in the faces. Morphological opening is performed again to remove small to medium objects that are safely below the size of a side face.

- Connected Region Analysis:

Once we have a mask showing potential side faces, we need to split this mask up into regions which can be searched for side faces. After some very basic erosion and hole-filling steps, most of the side faces are nicely contained in a single contiguous set of pixels. These sets can easily be found and labelled by connected component labelling.

- Rejection Based on Geometry:

The image output by connected region analysis contains quite a few non-face regions. Most of these are hands, arms, regions of dress that match skin color and some portions of background. In Rejection based on geometry, image statistics from the training set are used to classify each connected region in the image.

Three cases are defined for geometry rejection.
- Too small: The regions that are having very small area are rejected.
- Narrow and tall: The regions that have a small width but large height.
- Wide and short: The regions that have a small height but large width.

Figure 2 shows the result of rejection based on geometry. The regions which are not rejected with above criteria are considered to be side faces. Figure 3 shows the result of side face detection.

(a) (b)

(a) Before Rejection (b) After Rejection

Figure 2: Result after applying rejection criteria

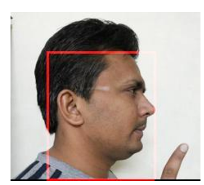


Figure 3: side face detection

After this pre-processing, we have applied template matching method on detected side face.

2.2 *Template Matching*

This technique includes four steps: (1) Template creation (2) Resizing template (3) Ear localization.

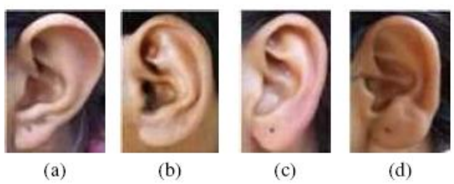
- Template Creation:

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS…..

For any template based approach, it is very much necessary to obtain a template which is a good representative of the data. In this technique, ear template is created by averaging the intensities of a set of ear images. Human ear shape can broadly be categorized into four classes: triangular, round, oval, and rectangular as shown in figure 4. For creation of ear template, all types of the ear are considered to obtain a good representative template.

Intuitively, it seemed reasonable to us that the best template to use would be one derived by somehow averaging the some ear images of the dataset that would likely be in the testing images. We would like to find a good subset of the ears found in the training images that are clear, straight, and representative of typical lighting/environmental conditions.



(a) Triangular, (b) Round, (c) Oval, (d) Rectangular
Figure 4: Shapes of the ear [9]

It is also important that these images be properly aligned and scaled with respect to one another. To this end, considerable time was spent for manually segmenting, selecting, and aligning ear images. At the end 17 ear images were chosen. These cropped images were first converted into gray scale and then the average was found which gives final template. The ear template T is formally defined as,

– Eqn-1

where N is the number of ear images used for ear template creation and E_k is the ear image. $E_k(i, j)$ and $T(i, j)$ represent the pixel values of the pixel of E_k and T respectively.
Thus, our final template for ear detection is a result of averaging together the 17 ear images. The actual template used in the matched filtering is of size 50×32 pixels. The template generated and used in the experimentation is shown in figure 5.

Figure 5: Created template image

- Resizing Template:

It is observed that the size of the ear is proportional to the size of the side face image. This observation is used for resizing the ear template in the proposed technique. To handle the detection of ears of various sizes, ear template need to be resized to make it appropriate for the detection of ear in the image.

— Eqn-2

By keeping the aspect ratio of the ear template same, it is resized to the width obtained in above Equation-2. Where and be the widths of the input face image and the reference face image respectively and and are the widths of input ear image and the reference ear image respectively.

- Localization:

To search an ear in the image I, ear template T is moved over the probable area of ear in the image and normalized cross correlation coefficient (NCC) [4] is computed at every pixel. NCC at point (x, y) is defined in Equation-3 as,

Eqn-3

where sum is performed over u,v under the window containing T positioned at (x,y). $\bar{I}$ x,y and are the average of brightness values of the portion of the target image under the template and template image respectively. Values of NCC lie between -1.0 and 1.0. Where it is found maximum, a rectangle is drawn around it to show detected ear. Value of NCC closer to 1 indicates a better match.

2.3 Ear Verification

To determine whether a detected ear is a true ear or not two methods are used namely verification using Euclidean distance [4] and verification using SVM tool [10-12].

- Ear Verification Using Euclidean Distance :

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

Here to determine whether a detected ear is a true ear or not, shape based ear verification is performed. To measure the similarity, Euclidean distance between the two sets (one for template and another for detected ear) of mean is used, which is estimated as follows:

Distance= Eqn-4

where M^T and M^E are mean of ear template and detected ear respectively.
Edge images of the ear template and the detected ear image are obtained using canny edge detector and the similarity distance between them is calculated using Equation-4. If the value of distance is less than a pre estimated threshold, detection is accepted otherwise it is rejected.

- Ear Verification Using SVM Tool :

A Support Vector Machine (SVM) [10, 11] is a machine learning tool that is becoming popular due to its success in pattern recognition applications; though, is not limited to this one application. SVM tries to minimize the upper bound on the expected risk. It is always guaranteed to find the global minima. Therefore the generalization ability (error in predicting data not present in training set) is more accurate compared to other classifiers. Without proper quality and quantity of data, the generalization achieved by SVM would be better 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.

In this work libsvm-2.91 tool [12] is used for ear verification. The choice of the appropriate kernel for a specific application is again problem dependent and often a difficult task. The different kernel like Linear, Polynomial, Radial Basis Function, Sigmoid and accordingly parameters i.e. degree, gamma, cost co-efficient are adjusted to improve the verification accuracy.

3. EXPERIMENTAL RESULTS

3.1 Data Acquisition

In this work CVL (CVL is library for image and data processing using graphics processing units (GPUs)) dataset [13] 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 is also produced by the authors, having images of 40 side faces with dynamic lighting condition with screen resolution of 2848 × 2144.

3.2 Results of Ear Detection

To create ear template, a set of ear images of 17 people is considered. NCC is used to localize the ear. Points having maximum NCC values are declared as the detected ear.

This experiment is performed on 100 images of CVL dataset and 40 images of general dataset. On the detected ears after experimentation, if the detected region contains part of ear, it is considered to be a positive detection; otherwise it is a false detection. Figure 6 shows examples of some of the positive detections. The proposed technique is also able to detect ear in presence of little occlusion due to hair. The fourth image of fourth row of figure 6 shows such example. Figure 7 shows examples of some of the false detections. Localization method has failed in some cases, especially for the images which are of poor quality or heavily occluded due to hair (Fourth image of second row in figure 7) or face is oriented with some angle (Third image of first row and third image of second row in figure 7).

Accuracy of the localization is defined by (genuine localization/total sample) × 100. It is found to be 88% for the CVL dataset and 90% for general dataset. The average time to detect an ear from a side face image is approx. 3 seconds with Matlab environment.

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

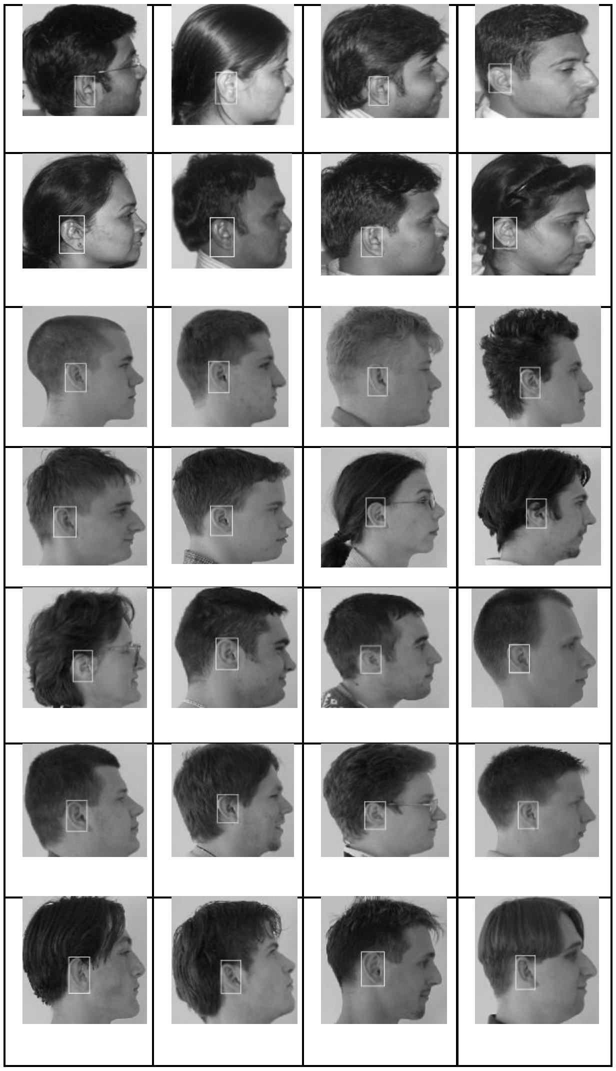


Figure 6: Positive Detection

3.3 Results of Ear Verification

Here for ear verification libsvm-2.91 tool is used. Here template is of size 50×32 and so the detected ear is also of the same size. Therefore, 1600 pixel intensity values are in one feature vector. The different parameters are set for accuracy measurement. The maximum accuracy is found to be 89.29%. The maximum accuracy was obtained with linear and polynomial kernel. However, in our experimentation few other combinations were also found which gave the same accuracy. Radial Basis Function kernel with default gamma (1/no of features =1/1600=0.000625) gives 71.43% accuracy and with gamma 0.001 gives 78.57% accuracy. Sigmoid kernel gives minimum accuracy i.e.53.57 %.

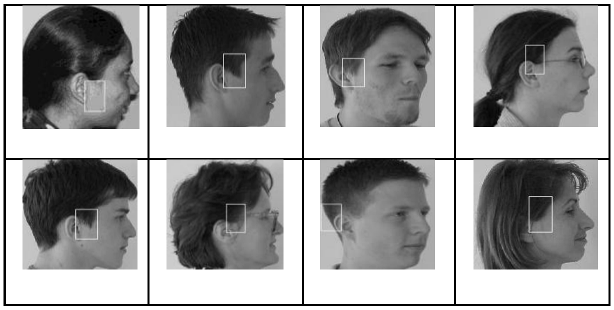


Figure 7: False Detection

4. CONCLUSIONS AND FUTURE WORK

The proposed method does not have any effects of different environments on the detection of ears. The constraint for getting very good result is that the template has to be recreated for different datasets otherwise it degrades the performance of detection. Here by pre-processing extra skin portion except side face i.e. neck is removed so template is moved on only side face so detection time is less and accuracy is high. If the template is moved over the whole skin area then required detection time is approx. 7 seconds while in our proposed method it is approximately 3 seconds only. This method fails if ears are heavily occluded due to hair or face is oriented with some angle. Experimental results on real side face images also demonstrate the effectiveness of the presented approach.

REFERENCES

[1] **B. V. Srinivasan, "Ear Extraction From the Image of a Human Face".**

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

[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", Proceedings of International Conference on Advances in Pattern Recognition,**

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

ICAPR' 09, pp. 213-216, Kolkata, India, Feruary 2009.

[5] SL Phung, A, Bouzerdoum , D, Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison", IEEE Transactions on Pattern Analysis and Machine Intelligence, January 2005, vol.27,no.1, 148-154.

[6] V. Vezhnevets, V. Sazonov, and A. Andreeva: "A Survey on Pixel-Based Skin Color Detection Techniques", International Journal of Expert Systems With Applications, vol.3, pp 4497-4507, April 2009.

[7] N. Rahman, K. Wei, and J. See, "RGB-H-CbCr Skin Color Model for Human Face Detection".

[8] J. D. Foley, A. Dam, S. K. Feiner, and J. F. Hughes "Computer Graphics Principles and Practice", Pearson Education, 2006.

[9] K. Joshi, and N. Chauhan, "A Template Matching and Support Vector Machine based Approach for Human Ear Detection and Varification", Int. Conf. Information, Signals, and Communications (ICISC-2011), A. D. Patel Institute of Technology, Gujarat, India, 5-6 February, 2011.

[10] N. C. Chauhan, Y. K. Roy, A. Kumar, A. Mittal, and M. V. Kartikeyan, "SVM-PSO Based Modeling and Optimization of Microwave Components", Frequenz, Vol. 62, 2008.

[11] 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, pp. 23-35, Jan,2001.

[12] 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.

[13] www.lrv.fri.unilj.si/~peterp/FaceDB.zip