

An Approach of Extracting Facial Components for Facial Expression Detection using Fiducial Point Detection

Ashim Saha

Department of Computer Science and Engineering
National Institute of Technology,
Agartala, India

Hillol Das

Department of Computer Science and Engineering
National Institute of Technology,
Agartala, India

Nirmalya Kar

Department of Computer Science and Engineering
National Institute of Technology,
Agartala, India

M C Pal, Ph.D

Department of Mathematics
National Institute of Technology,
Agartala, India

ABSTRACT

Facial Expression detection or Emotion Recognition is one of the rising fields of research on intelligent systems. Emotion plays a significant role in non-verbal communication. An efficient face and facial feature detection algorithms are required to detect emotion at that particular moment. The paper presents a new approach to the problem of extracting the facial components from a still facial image which will be used further for Facial Expression detection. Robust extraction of such facial feature locations is a crucial problem which is used in a wide range of applications. Facial features such as lip corners, eye corners and nose tip are critical points in a human face. The paper approach relies on image segmentation based on skin color algorithm and fiducial point detection. First, the face region is extracted from the image by skin-color filter and window growing. Second, eyes and mouth are approximately located by projection of fiducial point. It is been observed for subjects not wearing glasses, the extraction of eyes could be correctly located in 94% of the images and in over 90% of these images they could be accurately extracted, while for subjects wearing glasses, the success rate is somewhat lower. Also, in certain moods or depending on the facial expression, the detection of lips in case of mouth opening and closing varies which can also be extracted by our method.

Keywords

Feature extraction, fiducial point, facial expression.

1. INTRODUCTION

It has been observed in recent years that detecting human faces and facial features has become an important task in computer vision with numerous potential applications including human computer interaction, video surveillance, face tracking, and face recognition. Hence the extraction of facial features can be used to handle various practical and artificial problems. In the case of facial feature localization, the goal is to detect the presence and location of features. But the challenges associated with face and facial feature detection can be attributed to the following factors [1].

- (1) Intensity. There are three types of intensity: color, gray and binary.
- (2) Pose. Face images vary due to the relative camera-face pose (frontal, 45°, profile), and some facial features such as an eye

may become partially or wholly occluded.

- (3) Structural components. Facial features such as beards, mustaches, and glasses may or may not be presented.
- (4) Image rotation. Face images directly vary for different rotations.
- (5) Poor quality. Image intensity in poor-quality images, for instances, blurry images, distorted images, and images with noise, becomes unusual.
- (6) Unnatural intensity. Cartoon faces and rendered faces from 3D model have unnatural intensity.
- (7) Occlusion. Faces may be partially occluded by other objects such as hand, scarf, etc.
- (8) Illumination. Face images vary due to the position of light source.

The existing techniques are not applicable for all images with any of the above factors or their combinations. In addition, the accuracy of those techniques is based on the usage of intensity information as input.

It is well known that viewing a person's eyes, mouth is essential to grasping the information and emotions they convey and face components like eyes, nose, eyebrows, lips, etc. together describe the overall shape of the face. Due to great difference between face photo and face-sketch it is not possible to compare directly the features obtained from original face photo and the features obtained from face sketch. In this context extraction of individual facial features can proved to be very helpful. A lot of works on facial recognition based on facial feature extraction have been reported in [1–4].

In this method, effort has been made towards efficient extraction of facial features from randomly selected face and the experimental results has shown comparatively higher accuracy rate.

2. LITERATURE SURVEY

Based on the existing surveys conducted so far, five categories have been identified for face and facial feature detection.

—Geometry-based methods. These methods utilize geometrical information [5]. Each feature is demonstrated as a geometrical shape. They can accurately detect face and facial features, but cannot handle large variations of the face images such as images with some occluded facial features and images with noise.

—Color-based approaches. These approaches face difficulties in robustly detecting skin colors in the presence of complex background and different illuminations [12, 13]. These algorithms are applicable only for color images.

—Appearance-based methods. These methods use the models learned from a set of training images [15–17]. Gray value(intensity) is the most important parameter for the detection. They are not able to perfectly detect face images with poor quality in intensity, some occlusions, and unnatural intensity.

—Motion-based methods. Face and facial features are detected from the image sequence [14]. Using such methods, facial features cannot be detected using only a single still image from one view.

—Edge-based methods. In the last class of methods [19], faces are detected without information due to intensity and motion. The edge information is used as input. These methods can handle large variations of the face images. However, these methods detect only face or facial features [19].

3. SYSTEM OVERVIEW

For face localization relying on a combination of skin color and template based approach described in [8].The paper extends it by supporting three experimentally determined scaling factors for the template in order to localize faces with different size. The localization determines the position and the scaling factor of the matched template. After localize the face, pupils are used for face tracking since they can be easily determined with high accuracy. Since the positions of a face different only a few pixels between subsequent images, also the search regions to be analyzed for tracking of the pupils are very small i.e. search regions of 24X24 pixels. Hence after a face has been successfully localized, for subsequent images at first detection of fiducial points for the pupils, and depending on these for the nose tip and the corners of the mouth are determined. The proposed scheme contains four parts: facial area extraction in an entire image, estimation of the vertical positions of the facial features, such as the eyes and mouth, extraction of control points of facial features, and facial model adaptation.

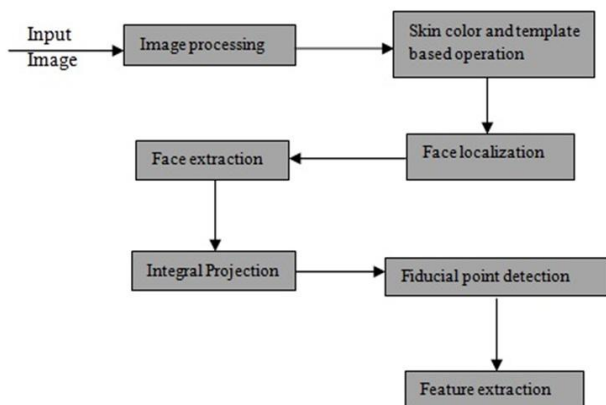


Fig. 1. Flow chart of the proposed scheme

4. FACIAL AREA EXTRACTION

The first step is to find the face location and separate it from background. Before locating the face area, a preprocessing procedure is necessary to reduce noise and get a good result from

edge detection. Histogram equalization used to produce an image with higher contrast, it can enhance the edge continuity of an image. Skin color detection algorithm is performed to get a binarized edge image. Using the binarized edge image and template based approach, separated the head from the background using following method. First, for each horizontal line, the two edge points which are close to both borders of the image are determined. By using the location information of these edge pixels and finding the brighter pixels, the area which includes the head and shoulder can be generated. Because the feature searching scheme focuses on the facial part, so the hair part must be removed. In general, the gray level of hair is lower than that of a face, and the height of the head area is about two-thirds the height of the total area, which includes the head and shoulder. The two lowermost side pixels whose gray levels are larger than the threshold values are detected because the face is assumed to be brighter. According to the positions of these two pixels, a bottom-up searching process that finds the side pixels of the facial part is performed. Finally, only the edge pixels in this new area are retained, and others are discarded. Resultant face image are shown in Fig 2.



Fig. 2. Extracted face image

4.1 Fiducial Point detection

After successful detection of face region, for fully automatic detection of 20 facial feature points plus the both eyes and the center of the mouth, lips in a face image, Gabor-feature based classifiers as proposed in [6, 18] is used. The method uses the input image is a face region. The input face region is then divided into 20 regions of interest (ROIs), each one corresponding to one facial point to be detected.

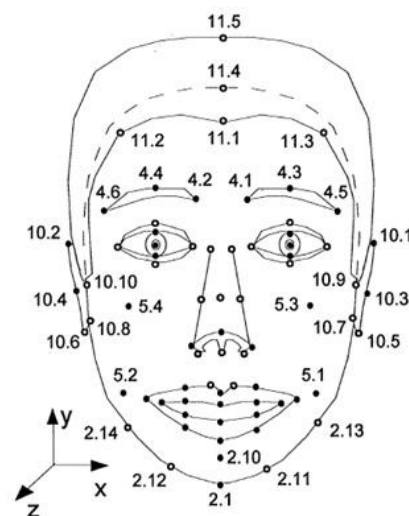


Fig. 3. Template of Fiducial point [6]

Initially the irises and the medial point of the mouth are detected. A combination of heuristic techniques based on the analysis of the vertical and horizontal histograms of the upper and the lower half of the face-region image achieves this. Subsequently, the detected positions of the irises and the mouth are used to localize 20 ROIs. Figure 4 shows the detection of various fiducial points.

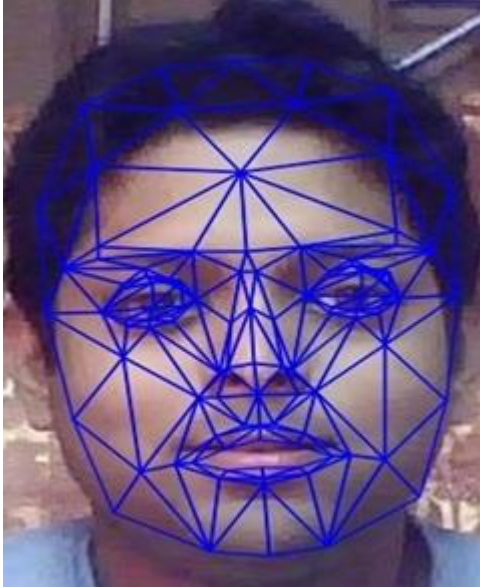


Fig. 4. Detection of Fiducial points

5. FEATURE EXTRACTION AND EXPERIMENTAL RESULTS

A mathematical statistical method proposed by M. Chuang et al [7] used to estimate the approximate vertical positions of the facial features, such as the eyes and mouth. An edge map image $I_{edge}(x, y)$ is divided into areas of equal size in the vertical direction. Then, the vertical integral projection of the edge map image in a rectangular area of $1 \leq x \leq IW$ and $i * step + 1 \leq y \leq (i + 1) * step$ is estimated by

$$v(i, x) = \sum_{y=i*step+1}^{(i+1)*step} I_{edge}(x, y)$$

where $v(i, x)$ indicates that the number of the edge pixels which have the same x coordinate in the i^{th} rectangle, W is the width of the input image, and $step$ is the rectangle height, where the image height is divided by N . N is the number of equal size rectangles.

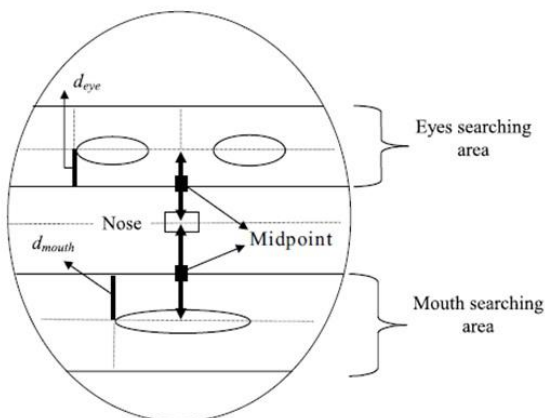


Fig. 5. Template of facial feature area [7]



Fig. 6. Fiducial point detection for point of interest

5.1 Extraction of Eye search area

After detecting fiducial point of face it is been observed a sharp change in color in fiducial point no 3.1 to 3.14 considered as a key to localize the eye region. For example, while marking the pixel values from the forehead region, the eyebrow region has a distinct color than the forehead region. In this approach, first the column wise sum value is computed for each of the columns. Eliminated the columns which have a sum less than 50% of the maximum column sums are eliminated out. This narrows down the search region by redefining the minCol and maxCol values. To determine the sharp color changes, the row sums enclosed by the column boundaries (say sum_1 to sum_h) are taken. Then the gradient between two consecutive row sums i.e difference between sums of i^{th} and $(i + 1)^{th}$ row are calculated. Taking the ten maximum gradients and their corresponding row values, it has been noticed that the least row value among these gives the row where the eyebrows are located. In case, the left and right eyebrows are not aligned in the same row, the row value of the upper eyebrow are used.

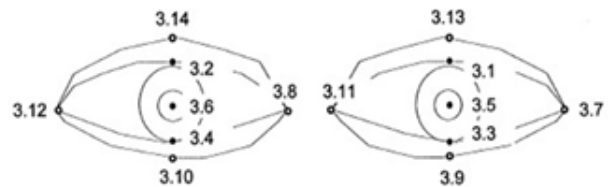


Fig. 7. Fiducial points of right and left eye regions

Once the eyebrows are located, the upper limit of the search area is defined. The lower boundary of the search area will be limited by upper limit plus half of the column window width ($maxCol - minCol$). Fig 8 shows the extraction of right and left eyes.



Fig. 8. (a) Right eye extraction (b) Left eye extraction

5.2 Extraction of Mouth area

To estimate the mouth area, first localize fiducial point from 8.2 to 8.10 as shown in Fig. 9. Then a conversion from rgb to lab color space is undertaken. The k-means clustering algorithm is applied next on this image to get two clusters.

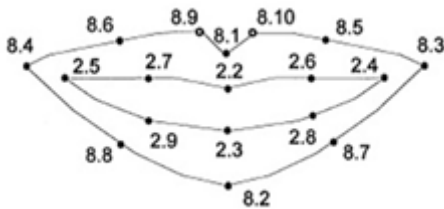


Fig. 9. Fiducial points of mouth regions

The two clusters are: skin cluster, lip cluster. The lip cluster (from Fig. 6) is selected to determine the mouth opening. To obtain mouth opening, the average intensity of three primary pixel colors are taken, and plot the row-average of such value contributed by each pixel against the row number. Fig 10 shows the extraction of mouth region.



Fig. 10. Extraction of mouth region

Table 1. The result for processed frames

| FrameType | No.ofTest Frame | No.of errorframesforfeature extraction | No.ofcorrectframesforfeature extraction |
|------------------------------------|-----------------|--|---|
| Faceimagewithoutglass | 25 | 2 | 23(92%) |
| Faceimagewithglass | 25 | 4 | 21(84%) |
| Faceimagewith h<25 degreetocamera | 25 | 1 | 24(96%) |
| Faceimagewith h25-45degreetocamera | 25 | 4 | 21(84%) |

6. CONCLUSION

In this study the proposed scheme used fiducial point of a facial image to extract the facial components which will further used in facial expression detection. Skin color and template based algorithm is used to localize facial area. After successful localization of face, detection of fiducial points for the pupils and depending on these both eye and the corners of the mouth are determined. Proposed scheme has limitations like background color and facial objects, also poses some advances like the user's face does not need to appear at a fixed position in the input image, the area of the person's face does not need to be fixed. In other words, the distance from the user to the camera can be varied. The geometry of the eye pair region and that of the mouth region is not limited.

As extensions of the proposed work, detection Left Eye Opening (EOL), Right Eye Opening (EOR), Eyebrow Raise Left (EBL), Eyebrow Raise Right (EBR), Mouth Opening Width (MOW), and the Mouth Opening Height (MOH) can be determined which are most important component for emotion detection.

7. ACKNOWLEDGMENTS

This research work has been carried out in National Institute of Technology, Agartala, India sponsored by All India Council for Technical Education (AICTE), Govt. of India under Research Promotional Scheme (RPS)(Ref No: 8023/RID/RPS-23/(NER)/2011-12)

8. REFERENCES

- [1] M.Yang,D.J.Kriegman, N.Ahuja, Detecting faces inimages:asurvey, IEEE Transactions on Pattern Analysisand Machine Intelligence.24(1)(2002)3458.
- [2] WuH,ChenQ,YachidaM(1995),Anapplication offuzzy theory:facetedetection.In: Proceedings of IWAIFGR95 ,pp314319.
- [3] Rowley HA, BalujaS, Kanade (1998) Neural network-based face detection.IEEE Trans Pattern Anal MachIntell20(1):2338.
- [4] BhatiaN,KumarR,MenonS(2007)FIDA: face recognition usin gdescriptive inputsemantics,December14.
- [5] Z.Liu,J.Yang,N.S.Peng, Anefficient face segmentation algorithm based on binary partition tree,Signal Processing: Image Communication20(4)(2005)295314.
- [6] A.M.Tekalp, J.Ostermann Face and 2-Dmesh animation in MPEG-4, Signal Processing:Image Communication 15 (2000)387-421.
- [7] M.Chuang,R.Chang and Y.Huang Automatic Facial Feature Extractionin Model-based Coding Journal of Information Science and Engineering16,447-458(2000).
- [8] P.Ekman, Facial Expression and Emotion,"in America Pshychologist, vol.48,no.4,pp.384-392,1993.
- [9] D. Vukadinovic and M. Pantic, Fully automatic facial feature point detection using gabor feature based boosted classiers,In SMC05,pp.1692-1698,2005
- [10] F.Fleuret, D.Geman, Coarse-to-fine face detection, International Journal of Computer Vision41 (12)(2001)85107.

- [11] L.Huang, A.Shimizu, Y.Hagihara, H.Kobatake, Gradient feature extraction for classification-based face detection, *Pattern Recognition*36(11)(2003)25012511.
- [12] F.Tsalakanidou, S.Malassiotis ,M.G.Strintzis, Face localization and authentication using color and depth images, *IEEE Transactions on Image Processing*14(2)(2005)152168.
- [13] M.Soriano, B.Martinkauppi, S.Huovinen, M.Laaksonen, Adaptive skin color model using the skin locus for selecting training pixels, *Pattern Recognition*36 (3)(2003)681690.
- [14] R.L.Hsu, M .Abdel-Mottaleb, A.K.Jain, Face detection in color images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*24(5)(2002)696706.
- [15] R.Xiao, M.Li, H.Zhang, Robust multipose face detection in images, *IEEE Transactions on Circuits and Systems for Video Technology*14(1)(2004)3141.
- [16] C.Liu, A Bayesian discriminating features method for face detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*25(6)(2003)725740.
- [17] Y.Li, S. Gong, J.Sherrah ,H.Liddell, Support vector machine based multi-view face detection and recognition, *Image and Vision Computing* 22(5) (2004)413427.
- [18] P.Viola, M.J.Jones, Robust real-time face detection, *International Journal of Computer Vision* 57 (2) (2004) 137154.
- [19] S.Phimoltares, C.Lursinsap, K.Chamnongthai, Locating essential facial features using neural visual model, in: *Proceedings of the First IEEE International Conference on Machine Learning and Cybernetics*, 2002, pp.19141919.