International Journal of Electronics Signals and Systems

Volume 3 | Issue 1 Article 6

July 2013

A NOVEL FACE DETECTION AND TRACKING ALGORITHM IN REAL- TIME VIDEO SEQUENCES

K. KAMATH NAGAPRIYA
Dept. of Electronics and Communication, Canara Engineering College, Mangalore, Karnataka, nagapriyakamath@gmail.com

A NOVEL FACE DETECTION AND TRACKING ALGORITHM IN REAL- TIME VIDEO SEQUENCES

NAGAPRIYA KAMATH K[1], ASHWINI HOLLA[2] & SUBRAMANYA BHAT[3]

[1,2&3]Dept. of Electronics and Communication, Canara Engineering College, Mangalore, Karnataka
E-mail : nagapriyakamath@gmail.com

Abstract - Face detection is a image processing technology that determines the location and size of human faces in digital images or video. This module precedes face recognition systems that plays an important role in applications such as video surveillance, human computer interaction and so on. This proposed work focuses mainly on multiple face detection technique, taking into account the variations in digital images or video such as face pose, appearances and illumination. The work is based on skin color model in YCbCr and HSV color space. First stage of this proposed method is to develop a skin color model and then applying the skin color segmentation in order to specify all skin regions in an image. Secondly, a template matching is done to assure that the segmented image does not contain any non-facial part. This algorithm works to be robust and efficient.

Keywords - *Skin color model, segmentation, template matching.*

I. INTRODUCTION

The goal of this proposed method is to develop multiple face detection and tracking system in real time scenario based on a fast and efficient face segmentation approach. This method consists of two image processing steps. Separate the human skin regions from non-skin regions thereby locating human faces within the skin regions. The next step is to use template matching to assure the segmented image does not contain any non-facial part. There are two types of face detection systems: face detection in still images and face detection in real time video sequences. Few approaches used in face detection are mainly based on PCA[1],neural network, edge detection based, COM and ROI based[2], feature based[3],color based [3][4] etc. Color based face detection and tracking has the following drawbacks: It gives a coarse face segmentation giving spurious results with a background cluttered with skin colored regions and found to be unsuitable when subjects in the image wear skin colored dress or their attire has patches of skin color. To overcome these drawbacks, some heuristics are developed that includes parametric skin color modeling, adaptive thresholding and dynamic template matching for face region detection and tracking. In this proposed work, a reliable skin color model is developed. Further, skin color segmentation, adaptive and optimal thresholding is done. Face detection is achieved by Euler number, heuristics and template matching. The balance of this paper is organized as follows. Methodology is discussed in section II, skin color modeling is proposed in section III, segmentation is proposed in section IV and face detection and tracking algorithm is discussed in section V. Experimental results are presented in section VI. Conclusions are made in section VII.

II. METHODOLOGY

This proposed method uses a skin color model approach in which skin color region is separated from non skin color region. To obtain skin color region, RGB color image is converted into YCbCr and HSV color space. The chromacity is fitted over a Gaussian distribution and applied over the image to obtain a skin likelihood image which is then converted into binary format by applying adaptive and optimal thresholding respectively. In order to obtain the face region from this binary image, the number of holes from the connected components in the binary image using Euler's number is determined. If the number of holes in a given connected component is greater than '1', then the centroid and orientation of skin region w.r.t. the center is determined. The height to width ratio of the oriented image is also determined and compared by developing the heuristics model. If height to width ratio is within the range of 0.8 to 1.6, then template matching is done using a template face. Template matching is performed by finding out the cross-correlation between the skin region and template face images. If cross-correlation exceeds 0.8 then the skin region is identified as a human face.

III. SKIN COLOR MODEL

In order to segment the human skin regions from non-skin regions based on color, a reliable skin color model [1] that is adaptable to people of different skin colors and to different lighting conditions is necessary. In this work, a skin color model in chromatic color space is used for segmenting the skin region. The common RGB representation of color images is not suitable for characterizing skin-color as, the triple component represents not only color but also luminance. Luminance may vary across a person's face due to the ambient lighting and thus is not a reliable measure in separating skin from non-skin region. Luminance can be removed from the color representation in the chromatic color space. Chromatic colors in the absence of luminance, is defined by a normalization process [2] as shown in Eqn (1).

r = R/(R+G+B) ; g = G/(R+G+B); b = B/(R+G+B) (1)

Although skin colors of different people appear to vary over a wide range, they differ much less in color than in brightness.ie: skin colors of different people are very close, but differ mainly in intensities. Hence, YCbCr and HSV color model are often used to implement the skin color region. In the RGB domain, each component of the picture has a different brightness. However, in the YCbCr domain all information about the brightness is given by the Y-component, since the Cb (blue) and Cr (red) components are independent from the luminosity. The following conversions are used to segment the RGB image into Y, Cb and Cr components are given by the Eqn (2).

Y = 0.257 * R+0.504 *G +0.098* B +16
Cb = 0.148 * R -0.291*G +0.439* B +128 (2)
Cr = 0.439 * R -0.368*G -0.071*B +128

The Cb and Cr components give a good indication on whether a pixel is a part of the skin or not. HSV color space represents colors in terms of Hue, Saturation and Intensity of the given pixel. Hue refers to color type, such as red, blue, or yellow. Saturation refers to the purity of the color. Value component refers to the brightness of the color. A total of 100 color images were used to determine the color distribution of human skin in chromatic color space. These samples were taken from persons of different ethnicities: Asian, American and African. As the skin samples were extracted from these color images, the skin samples were filtered using a low-pass filter to reduce the effect of noise in the samples. The impulse response of the low-pass filter is given by:

$$H(Z) = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

The color histogram revealed that the distributions of skin color of different people are clustered in the chromatic color space and a skin color distribution can be represented by a Gaussian model N (m, C) where

Mean: $m = \frac{1}{N}\sum_{0}^{N-1} x$ Where x = (Cr, Cb) ^{T} (3)

Covariance: $C = \frac{1}{N}\sum_{0}^{N-1}\left[(x-m)(x-m)^T\right]$ (4)

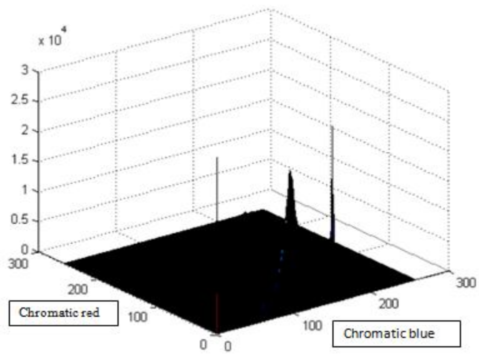


Fig 3.1: Fitting skin color into a Gaussian distribution.

With this Gaussian fitted skin color model, the likelihood of skin for any pixel of an image is obtained. Therefore, if a pixel, having transformed from RGB color space to chromatic color space has a chromatic pair value of (Cr, Cb); the likelihood of skin for this pixel is computed using Eqn (5).

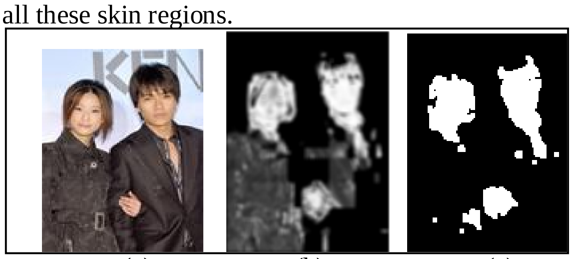
Likelihood = P(r, b) = exp [-0.5(x-m)^T C^{-1}(x-m)] (5)

IV. SKIN COLOR SEGMENTATION

The first stage in segmentation [4] [5] is to transform an image into a skin-likelihood image. This involves transforming every pixel from RGB representation to chroma representation and determining the likelihood value based on the Eqn. (5). The skin-likelihood image is a gray-scale image whose gray values represent the likelihood of the pixel belonging to skin. A sample color image and its resulting skin-likelihood image are shown in Figure 4.1(a) and (b). It is seen that all skin regions like the face, hands and the arms appear brighter than the non-skin region. Since the skin regions are brighter than the other parts of the images, the skin regions need to be segmented from the rest of the image through a thresholding process. Since people with different skins have different likelihoods, an adaptive thresholding process is used to achieve the optimal threshold value for each run.

i) ADAPTIVE AND OPTIMAL THRESHOLDING

The adaptive thresholding means stepping down the threshold value which intuitively increases the segmented region. However, the increase in segmented region will gradually decrease as percentage of skin regions detected approaches 100%, but will increase sharply when the threshold value is considerably too small that other non-skin regions get included. The threshold value at which the minimum increase in region size is observed while stepping down the threshold value gives the optimal threshold. In this proposed work, the threshold value is decremented from 0.85 to 0.05 in steps of 0.1. If the minimum increase occurs when the threshold value changes from 0.45 to 0.35, then the optimal threshold must be taken as 0.4.It is not necessary that all detected skin regions contain faces, many regions often correspond to hands, arms and other exposed parts of the body, while some correspond to objects with colors similar to those of the skin. The skin segmented image of the original color image resulting from this technique is shown in fig.4.1(c). Hence the second stage of face detection algorithm employs facial features to locate the face in all these skin regions.



(a) (b) (c)

Fig. 4.2 (a) original color image, (b) skin likelihood image, (c) skin segmented image.

To determine which regions possibly determine a human face, the number of skin regions in an image need to be determined. A skin region is defined as a closed region in the image, which has a 0, 1 or more holes in it. A skin region can be thought of as a set of connected components within an image. All holes in a binary image have pixel value of zero (black).The process of determining how many regions are present in a binary image is by labelling such regions. A label is an integer value. A 8-connected neighborhood (i.e., all the neighbors of a pixel) is often used to determine the labeling of a pixel. If any of the neighbors had a label, then label the current pixel with that label. If not, a new label is used. Finally, we count the number of labels and this gives the number of regions in the segmented image. To separate each of the regions, a new image is being created that have ones in the positions where the label occurs. The others are set to zero. After this, we iterate through each of the regions found in order to determine if the region contains a human face or not.

V. FACE DETECTION AND TRACKING

After experimenting with several images, it is found that a skin region should have at least one hole inside that region. Regions that have no holes are neglected. The number of holes inside a region is compute using the Euler number of the region given in Eqn. (6)

$$E = C - H \quad (6)$$

Where C: number of connected components and H is the number of holes in a region. Since we are considering one skin region at a time, we set the number of connected components (i.e. the skin region) to 1. Therefore, numbers of holes,

$$H = 1 - E \quad (7)$$

a) Center of the mass

The center of area in binary images is the same as the center of the mass and it is computed using Eqn (8) and Eqn (9).

$$\bar{x} = \frac{1}{A}\sum_{i=1}^{n}\sum_{j=1}^{m} jB(i, j) \quad (8)$$

$$\bar{y} = \frac{1}{A}\sum_{i=1}^{n}\sum_{j=1}^{m} iB(i, j) \quad (9)$$

Where B is the matrix of size [n x m] that represents the region and A is the area in pixels of the region.

b) Orientation

Most of the faces considered in this work are vertically oriented while some of them have a little inclination. A unique orientation is possible by elongating the object. The orientation of the axis of elongation determines the orientation of the region. The axis is computed by finding the line for which the sum of the squared distances between region points and the line is minimum. The least-squares of a line are computed to the region points in the image. Angle of inclination (theta) is given by Eqn. (10):

$$\theta = \frac{1}{2}\tan^{-1}\frac{b}{a-c} \quad (10)$$

Where: $a = \sum_{i=1}^{n}\sum_{j=1}^{m} (x'_{ij})^2 B[i, j]$

$$b = 2\sum_{i=1}^{n}\sum_{j=1}^{m} x'_{ij} y'_{ij} B[i, j]$$

$$c = \sum_{i=1}^{n}\sum_{j=1}^{m} (y'_{ij})^2 B[i, j]$$

And $x' = x - \bar{x}$; $y' = y - \bar{y}$

c) Height-to-width ratio

To determine the width and height of the region, it is necessary to resize the template face so that it has the same width and height of skin region. First, we fill out the holes that the region have. Since the image is rotated by an angle theta, we need to rotate our region by theta degrees so that it becomes completely vertical. Height and width is determined by moving 4 pointers: one from the left, right, top and bottom of the image. If we find a pixel value different from 0, we stop and this marks the coordinate of a boundary. When we have the 4 values, we compute the height by subtracting the bottom and top values and the width by subtracting the right and the left values.

d) Template matching

For an image corresponding to the skin region, we first close the holes in the region and multiply this image by the original one. In template matching[4][5][6], the template face must be positioned and rotated in the same coordinates as the skin region image. Initially, the template frontal face as shown in Fig.5.1 (a) is resized according to the height and width of the region computed and the resized template face is then rotated according to angle theta so that the template face is aligned in the same direction as the skin region. The center of the rotated template face is computed as shown in the previous section. This process creates the grayscale image that will have the resized and the rotated template face model. The cross-correlation value between the parts of an image corresponding to the skin region is computed and the template face properly processed and centered. It was determined experimentally that a good threshold value for classifying a region as a face is if the resulting cross-correlation value is greater than 0.6.To plot the rectangular box, co-ordinates of the template is found that exactly determines the pattern the template face can have on the resultant matched image shown in Fig 5.1(b). With these coordinates, we draw a rectangle in the original color image shown in Fig.5.1(c) which marks the output of the face detection system.

(a) (b) (c)

Fig 5.1 (a) Original template , (b) Template matched output and (c) face detected output.

VI. RESULT AND ANALYSIS

This algorithm is found to detect the face for different illumination, orientations and poses tested across a wide range of subjects. It is also found to work well with partially occluded faces. It detects only the face region and does not detect any other skin colored objects. This algorithm works effectively for real time implementation. Fig.6.1 (a), (b) and (C) depicts real-time face detection under different illumination and position using MATLAB GUI.

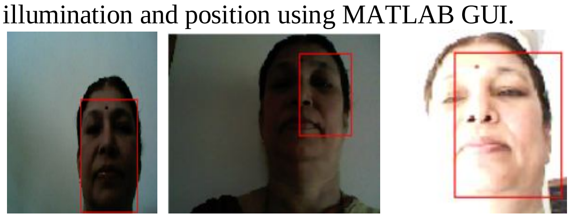


Fig 6.1: (a) (b) (c)

VII. CONCLUSION

The proposed method is face detection and tracking system in real-time video sequences that captures an image sequence from a camera, segments, detects and tracks efficiently a human face. It is an optimal blend of skin color model, heuristics and template face that achieves high performance in real time. This entire algorithm for face segmentation, detection and tracking is implemented in MATLAB R2008b using MATLAB Image Processing toolbox that takes around 20secs per frame in Intel core i3 processor to segment, detect and track multiple faces. We used nearly 100 images to test the performance of this implementation and obtained 80% of accuracy. Real time implementation is done using MATLAB GUI. This proposed method can be further developed to incorporate the real-time face recognition system.

REFERENCES

[1] Liying Lang and Weiwei Gu, “The Face Detection Algorithm Combined Skin Color Segmentation and PCA”, IEEE Trans 2009.

[2] J.Fritsch,S.Lang,G.A Fink, “Improving Adaptive skin color segmentation by incorporating results from face detection”, Proceedings of 2002 IEEE, Workshop on robot & human interactive communication, pp 337-343.

[3] Mohamed A Berbar, “Faces and Facial Features Detection in Color Images”, proceedings of Geometric modelling and imaging, IEEE Trans 2006.

[4] Chen Aiping, Pan Lian, Tong Yaobin, Ning Ning, “Face Detection Technology Based on Skin Color Segmentation and Template Matching”, 2010 Second International Workshop on Education Technology and Computer Science ,IEEE trans 2010,pp 708-711.

[5] Xiaoping Li , Yinxing Li, Yongbiao Yao, Lin Zhang, “Study and Realization of Face Detection Based on Skin Segmentation and Template Matching”, pp 375-378.

[6] Jizeng Wang and Hongmei Yang, “Face Detection Based on Template Matching and 2DPCA Algorithm”, IEEE Trans 2008.

[7] Rafael Gonzalez and Richard E Woods, “Digital Image Processing”, 2nd edition, Pearson Education, 2001.

[8] Ani.K.Jain, “Fundamentals of digital image processing”, 2nd edition, Pearson Education, PHI, 2001.

◈◈◈