ISSN:0975-9646

Jyoti Rani et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4394-4397

An Interface for Extracting and Cropping of Face from Video Frames

Jyoti Rani
Scholar, M Tech(CSE)
Department of Computer Science and Application
Kurukshetra University, Kurukshetra
Haryana (India)

Dr. Kanwal Garg
Assistant professor
Department of Computer Science and Application
Kurukshetra University, Kurukshetra
Haryana (India)

Abstract:The premise of this paper is to design an interface that can detect human face and obtain the face images which have normalized intensity, uniform size and shape. In this paper a system is developed for face extraction and cropping from consecutive video frames. The developed system can be divided into two main parts. First is face detection and extraction, second is normalization and cropping of faces. Input to the system is the sequence of frames captured using a built-in webcam or external video device. The proposed system is implemented in MATLAB version 8.1.0.604 (R2012a).

Keywords: YCbCr and RGB Color Spaces, Skin Color Model, Face Detection and Extraction.

I. INTRODUCTION

Face is the primary focus of interest in social interactions and plays a major role in conveying identity and emotion of the person. Various methods of face detection are discussed in [l]. These methods make use of different techniques like neural networks, machine learning, pattern matching etc. The neural network [2] and view-based [3] methods involve a huge number of face and non-face training examples, and work with grayscale images. However, most of these methods for face detection are sensitive to different lighting conditions. It is difficult to design a system that work better in all conditions like different illuminations, face colors, sizes and backgrounds. A statistical method was also introduced that extends the face detection by training two separate classifiers for 3d objects [4]. In face detection systems skin color can play a very important role. But in real time systems variation of lighting condition can create more complexity. Different people have dissimilar skin tone, while the difference lies generally in the color intensity not in chrominance value. Studies show that YCbCr color space is a powerful color space in segmenting the skin color accurately. The Y component in YCbCr color space represents the intensity of the light, Cr is red and Cb is the blue chrominance. Here a real time system is proposed that can detect and crop face and is also able to work better with different lightning conditions.

II. RELATED WORKS

Face detection systems frequently adopt the extraction of skin-tones. Ishii et al. [5] uses the YCbCr color space for face extraction using a back-propagation method of neural network. Performance of face detection is improved using skin color parts. Skin-color models are constructed using a large number of skin-tone pixels for detecting possible candidates of faces. A lighting compensation technique is proposed in [6]. Under a wide range of lighting variation compact skin clusters can be acquired in the YCbCr color space [7, 8]. YCbCr color space makes sure that the algorithm still maintains robustness for different skin-tones because the skin-tone color depends on the luminance value. As compared to RGB(red, green, blue) color space better performance can be achieved for varying illumination using YCbCr [9]. Methods for face detection and extraction can be categorized as holistic based and feature based methods. Holistic based method treat the face as a single unit while feature based methods divide the face into small regions of facial features like eyes and mouth. Most approaches to eye and face localization are template based [10]. Feature based method has downsides when dealing with identification of closed eyes or mouth. So in this paper holistic based method is used to overcome this problem.

III. SYSTEM ARCHITECTURE

3.1 Face detection and extraction

The problem of face detection has been studied broadly. It is a challenging task to design a system that work for all illuminations, face colors, sizes backgrounds. The very first task for any vision system is the image acquisition task. Here the task of image acquisition in real time is carried out using live streaming where image frames are received using streaming media. Streaming media can be a built-in webcam or external video device. The device properties like brightness, contrast, sharpness etc. are adjusted according to lightning conditions.

```
obj =imaq.VideoDevice('winvideo', 1, 'YUY2_320x240');
set(obj,'ReturnedColorSpace', 'rgb');
set(obj.DeviceProperties, 'Brightness',20,...
        'Contrast', 10,...
        'Saturation', 100,...
        'Sharpness', 8);
```

The live streaming keep on going until required input image frame is acquired. The acquired image frame contains a single face with the following general features:

- Uniform illumination conditions
- Light color background
- Faces in upright and frontal position

After acquiring the required frame, different vision tasks can be performed to the input frame using different processing methods. On the other hand, if the image has not been acquired accurately then the proposed tasks may not be attainable, even with the help of some form of image improvement techniques.

Next step is face detection from the acquired image frame. It can be considered as a particular case of object-class detection that detect the size and location of specified objects within an input image frame.
Viola-Jones detection algorithm and a trained classification model for detection are used in this system for face detection. The basic principle behind the Viola-Jones algorithm is to scan a sub-window that is able to detect faces in the input image frame [11]. The standard method would be to resize the input image frame to different sizes and then apply the fixed size detector to these images. This method looks to be rather time consuming because of different size images. Opposing to this standard method Viola-Jones resize the detector itself instead of the input image frames and applies the detector with different size several times through the image.
Cascade object detector can be used to detect and track a face in consecutive video frames. Face detection and tracking may be lost when the face angle changes or the person rotate their head. This is because of the type of trained classification model used. A bounding box is then inserted around the detected face

```
faceDetector = vision.CascadeObjectDetector();
frame = step(obj);
bbox = step(faceDetector,frame);
boxInserter=
        vision.ShapeInserter('BorderColor','Custom',...
        'CustomBorderColor',[255 255 0]);
videoOut = step(boxInserter, frame,bbox);
imshow(videoOut,'border','tight');
```

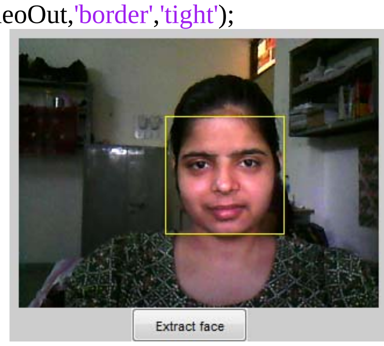


Figure 1: Result of face detection

The rectangular area of box with the detected face is extracted and saved to disk into a specific location in jpg format.

```
faceImage = imcrop(frame,bbox);
imwrite(faceImage,'face.jpg');
```

Figure 2: Result of face extraction

3.2 Face normalization and cropping

The aim of this phase is to achieve images which have normalized intensity, the same size and shape. In this phase only the skin part of face is cropped with removing non skin parts like background, head and hairs using skin color detection. Following four steps are carried out for this purpose.

3.2.1 Lighting compensation

Skin color of every individual can be different and also the skin-tone can be different due to variation in lighting conditions. Therefore the system should be able to handle these modifications in skin tone colors. To improve the performance a light compensation function is used to deal with varying illumination. The function is designed to balance the light in the input image according to [6]. According to the lighting conditions the brightest pixels in the image are set to as reference white then all pixels are linearly scaled to 255.

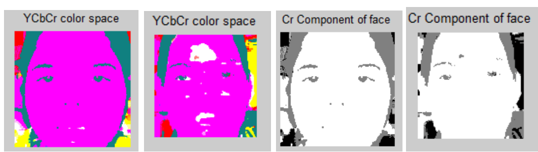
```
I = imread('face.jpg');
C=255*imadjust(I/255,[0.1,0.95],[0,1]);
```

Figure 3: Result of lighting compensation

3.2.2 Color space transformation and Skin color detection

In this phase skin color segmentation is performed to reject non-skin color regions from the input image. Human skin tone color across all regions agrees very much in its chrominance value and differs generally in its luminance value.
An appropriate color space should be used for modeling skin color and identifying a cluster associated with skin color in this space. Based on the comparison of the various color spaces for face detection [12], the YCbCr color space is chosen because it is extensively used in video compression standards. RGB color space of the input image is transformed into YCbCr color space so that the luminance value can be separated from the chrominance value.

```
YCbCr=rgb2ycbcr(C);
Cr=YCbCr(:,:,3);
```



(a) (b)

Figure 4: Result of Color space transformation (a) YCbCr color space of face image
(b) Cr components of face image.

As the skin-tone color differs in luminance value, the author transforms the YCbCr color space to make the skin cluster luma-independent. Only the Cr component of

YCbCr color space are choosed as Cr is red chrominance. This also enables better detection of dark and light skin tone colors. Skin regions and eyes can be extracted using information in the chrominance components Cr. Earlier studies indicate that the chrominance is independent from the luminance, which can be used to separate the skin color[13].

```
[SkinIndexRow,SkinIndexCol] = find(8<Cr & Cr<40);
for i=1:length(SkinIndexRow)
(SkinIndexRow(i),SkinIndexCol(i))=1;
end
```

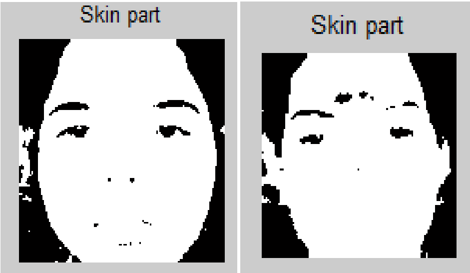


Figure 5: Result of skin color detection

3.2.3 High frequency noisy removing

In this phase noise is removed by using noise removal algorithm. In previous stage we get the skin part of the face with holes at eyes and lips. These holes are filled first using a matlab function for fillings holes in binary images and then high frequency noise removal is applied on the face part.

```
fill = imfill(S,'holes');
noise = medfilt2(fill,[14,14]);
```

2-D median filtering is used for noise removal. It performs filtering of the input image in two dimensions and each output pixel has the median value around the equivalent pixel in the input image. Median filtering is not a linear operation and usually used to reduce noises like salt and pepper in the images. A median filter is very effective to reduce noise and preserve edges.

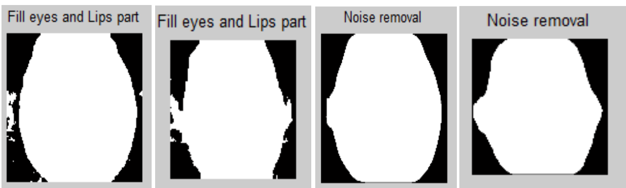
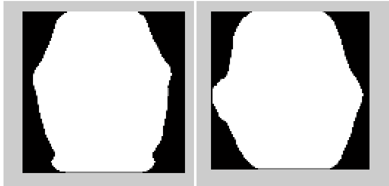


Figure 6: Result of holes filling and noise removal

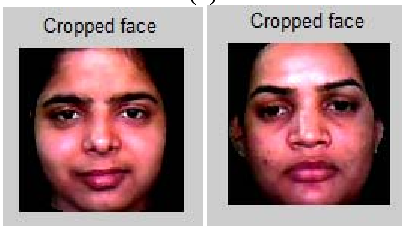
3.2.4 Masking and Resizing

The result of previous stage is a facemask which can be used to remove everything except skin-regions. This is done by applying the facemask to the original image and set the pixels inside the interval of mask as of face image and the rest of the pixels to 0.

```
cropped = bsxfun(@times, uint8(C), uint8(noise));
```



(a)



(b)

Figure 7: (a) Mask to be applied on face image, (b) Resulted cropped face

After that resulted cropped face is obtained and resizing operation is performed to that cropped image so that all images we obtain from this system will be of a specific size of 120 x 120 pixels. These cropped faces can be used in systems like face recognition and emotion detection with improved performance because these only contain the required information.

```
cropped =  imresize(cropped,[120,120]);
imwrite(cropped,[datestr(now,'HH-MM-SS--dd-mmm-yyyy'),'.jpg'],'jpg');
```

Figure 8: Result of face extraction and cropping in different lighting conditions and skin tones.

IV. CONCLUSION AND FUTURE WORK

In this paper a system is designed for extraction and cropping of face from the video frames using skin color detection. A light compensation function is used to improve the performance of the system. The proposed method can deal with range of lighting variation and skin tones. But still there are some conditions like light background of skin color which require more work to be done. Therefore, research in this field will remain under continuous study for many years to come because many problems have to be solved in order to create an ideal user interface.'

REFERENCES

[1]. D. Maio and D. Maltoni, "Real-time face location on gray-scale static images," Pattern Recognition, vo1.33, no. 9, pp. 1525-1539, Sept. 2000.

[2]. H.A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," IEEE Trans. PAMI, vol. 20, pp. 23-38, Jan. 1998.

[3]. K.K. Sung and T. Poggio, "Example-based learning for view-based human face detection," lEEE Trans. PAMI, vol. 20, pp. 39-51, Jan. 1998.

[4]. H. Schneiderman and T. Kanade, "A statistical method for 3D object detection applied to faces and cars," IEEE CVPR, June 2000.

[5]. H. Ishii, M. Fukumi, and N. Akamatsu, "Face detection based on skin color information in visual scenes by neural networks", Proc. of the International Conference on Systems, Man, and Cybernetics, Tokyo, Japan, vol. 5, pp. 350-355, 1999.

[6]. R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color image", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-706, 2002.

[7]. C. Garcia, and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis", IEEE Trans. Multimedia, vol. 1, no. 3, pp. 264-277, 1999.

[8]. D. Chai, and A. Bouzerdoum, "A Bayesian approach to skin color classification in YCbCr color space", Proc. of the IEEE Region 10 Conference (TENCON), Kuala Lumpur, Malaysia, vol. 2, pp. 421-424, 2000.

[9]. Learning-based Face Detection by Adaptive Switching of Skin Color Models and AdaBoost under Varying Illumination,Deng-Yuan Huang et al.http://bit.kuas.edu.tw/~jihmsp/2011/vol2/JIH-MSP-2011-04-001.pdf

[10]. W. Huang, Q. Sun, C.-P. Lam, and J.-K. Wu, "A robust approach to face and eyes detection from images with cluttered background," ICPR, vol. I , pp. 110-1 14, Aug. 1998.

[11]. Paul Viola, Michael J. Jones, "Robust Real Time Face Detection", International Journal of Computer Vision 57(2), 2004.

[12]. J.C. Tenillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," Proc. IEEE Automatic Face and Gesture Recognition, pp. 54-61, 2000.

[13]. Deng-Yuan Huang et al. "Learning-based Face Detection by Adaptive Switching of Skin Color Mode and AdaBoost under Varying Illumination", Journal of Information Hiding and Multimedia Signal Processing. ISSN 2073-4212, Volume2, Number 3, July 2011

AUTHOR'S PROFILE

Jyoti Rani received the B.Tech degree in Information Technology in 2012 and she is currently pursuing the M. Tech degree in computer science and engineering at the Department of Computer Science & Applications, Kurukshetra University, Kurukshetra.

Dr. Kanwal Garg is currently working as an Assistant Professor in the Department of Computer Science & Applications, Kurukshetra University, Kurukshetra with a teaching experience of twelve years. His research interest includes Data Base, Data Mining, Web Mining and Warehousing. He has published more than fifty research papers in international and national journals, conferences and seminars. He has attended many workshops and faculty development programme.