



Real Time Face Detection

M. SHARIF, K. AYUB, D. SATTAR, M. RAZA

Department of Computer Sciences, COMSATS Institute of Information Technology, Wah Cantt-Pakistan

E-mail: kamran.nse@gmail.com, danishsattar@gmail.com, mudassarkazmi@yahoo.com

Received 07th December 2011 and Revised 03th May 2012

Abstract: In this paper, the difficulty of face detection in real time scenario is considered and an appropriate solution is being proposed for this problem. The proposed idea is to detect human face in color images. For this purpose, HSV color space is used in the experiments to detect faces. The proposed algorithm detects face in a real time video in two steps. In the first step, it applies statistical model to get H (Hue) and S (Saturation) ratios for skin region. Secondly, it applies these defined ratios for scene width and height to get approximation of face location in an image with respect to the detected skin region. Finally, an eye template matching algorithm is applied on the previous roughly detected skin region to verify the face from this region. The proposed model has been tested in real time environment with fairly acceptable performance.

Keywords: Face, Detection, RGB, (Red, Green, Blue), HSV, (Hue, Saturation, Value) Skin.

1. INTRODUCTION

Face detection (Feraud *et. al.* 2001, Hjelm *et. al.* 2001), (Hsu *et. al.* 2002, Jones *et. al.* 1999, Rajagopalan *et. al.* 1998, Yang *et. al.* 2002) is the main module for any face recognition (Singh *et. al.* 2002, Takacs *et. al.* 1998) system which outlines and mark boundaries of face regions from messy images. The image data to be processed may be obtained from video sequence or it may be a still image data. Most of the present face recognition systems require a facial image to be processed and it ultimately requires human computer interaction. But in real time scenarios, it is difficult to get facial image to be taken as an input for surveillance or point automation systems. Face detection module should have the capability to find, outline and extract the exact faces from messy images to provide faces as an input to the face recognition module. Face detection module should have the following properties:

(1) It should be capable of outlining exact faces as fast as possible. (2) It should be capable of handling exact borders of malfunction faces. (3) The probability of false face identification should be nil. (4) Constraints of the system in context of human interaction should be low. (5) It should have the capacity of noise and background removal.

For a human brain, it is easy to detect and analyze face but it's not so easy for computer yet because of the reason that human face expressions get changed by different internal and external factors like lighting, pose variation, mustache, bear, glasses background colors etc and computer can't detect these types of changes easily.

Research for face detection problem is in process and a wide variety of approaches had been proposed for solving the detection problems. Face detection is difficult in video sequences as compared to the still images. Video sequences are largely affected by internal and external factor changes and thus maximize the complexities for detection module. Previous research defines a large number of skin ratios on the basis of different races (Yang *et. al.* 1996, Yoo *et. al.* 1999). One can detect skin by making use of these skin ratios. However, it is not always good to use this approach because in messy images faces get close to one another and the skin region gets confused.

This paper proposes an idea of extracting skin region from converting an image from RGB color (Zarit *et. al.* 1999) space to HSV color space. The H and S ratios are defined for extracting skin region. H and S values are proposed and used in our experiments for extracting skin. Some global parameters are defined for face area in an image. By using these parameters the algorithm roughly detects the face region and then it uses eye template matching to verify the face region in an image. The proposed algorithm has been tested and its results have been mentioned in this paper. Comparison of these results with the previous techniques (given in the results section) proves the new proposed algorithm quite satisfactory and efficient.

2. MATERIAL AND METHODS

This section provides the detailed description of proposed algorithm in three steps. It describes the HSV color model and conversion from RGB to HSV color model, skin verification process and face detection.

2.1. Face Detection

Our proposed face detection system is mainly comprised of two major steps. The initial step is to categorize the skin pixels. For classifying skin pixels, the equations are defined which satisfy the presence of skin in an image. The second step is to use the defined global parameters for scene height and width for approximating the face region in an image. Then the eye template matching is used to verify the face in an image.

2.2. HSV Model and RGB to HSV

Low complicated but precise methods such as HSV modeling and color histogram are favored. To section human skin areas from a consistent skin color model that is flexible to various skin colors and to diverse illumination conditions is required. The basic RGB color model is not appropriate for skin color classification because the triplet (R, G, B) represents not only color but also lighting which varies transversely the person's face due to the ambient illumination. HSV color model is a nonlinear renovation of the RGB color space. The HSV model is user-oriented and is based on artist's philosophy of hue, shade, and tone, with self-determining values for Hue, Saturation, and Value, corresponding respectively to wavelength, excitation, and intensity.

Black has HSV = (0, 0, 0). Thus the facial HSV model of Africans is clustered about the source of the coordinate arrangement. Across the African continent, unassuming variations occur which do not leave significantly from this value. Most color pictures are recorded as (R, G, B) triplets. Given a color distinct by (R, G, B) where R, G, and B are normalized to 0.0 to 1.0, an equivalent (H, S, V) color is determined by the following set of formulas. (Sandro et. al.)

$$H = \begin{cases} 0 + \frac{G - B}{MAX - MIN} * 60, & \text{if } R = MAX \\ 2 + \frac{B - R}{MAX - MIN} * 60, & \text{if } G = MAX \\ 4 + \frac{R - G}{MAX - MIN} * 60, & \text{if } B = MAX \end{cases}$$

$$S = \frac{MAX - MIN}{MAX}; V = MAX$$

Considering MAX to be the maximum of the (R, G, B) values and MIN the minimum values, the model is: where H, S and V vary from 0.0-360 and 0.0-1.0 respectively.

2.3. Skin Verification

Following mathematical model is proposed to classify skin pixels in an image. The model uses the values as:

$$\begin{aligned} \text{skin_color_H_min} &= 0.10; \\ \text{skin_color_H_max} &= 0.90; \\ \text{skin_color_S_min} &= 0.30; \\ \text{skin_color_S_max} &= 0.99; \end{aligned}$$

This proposed model has been tested and its results have been mentioned in results section.

$$\Theta_h = \bigcup_{i,j=1}^{m,n} (\alpha_{i,j} < \phi_{h\min} \cup \alpha_{i,j} > \phi_{h\max}) \quad (1)$$

where, $\alpha, \phi_{h\min}, \phi_{h\max}$ represent the hue value, minimum and maximum threshold values for hue respectively.

$$\Theta_s = \bigcup_{i,j=1}^{m,n} (\beta_{i,j} < \phi_{s\min} \cap \beta_{i,j} > \phi_{s\max}) \quad (2)$$

Where, $\beta, \phi_{s\min}, \phi_{s\max}$ represent the saturation value, minimum and maximum threshold values for saturation.

$$\Theta = \Theta_h \cap \Theta_s \quad (3)$$

where, Θ gives the resultant detected skin region.

Along with these HSV ratios the proposed algorithm also used the defined ratios of chromic blue and chromic red which are as under:

$$Cb = (\eta \times R) - (\delta \times G) + (\gamma \times B) + C \quad (4)$$

$$Cr = (\gamma \times R) - (\nu \times G) - (\lambda \times B) + C \quad (5)$$

Where, Cb and Cr represent chromic blue and chromic red respectively. $\eta, \delta, \gamma, \nu, \lambda$ are conversion constants from RGB to CbCr color space.

2.4. Face Detection

In this step the task is to detect face area in a detected skin region. The skin region detected may be rough up to some extent. As it is must to clean the skin region for verifying face in this region, therefore, image cleaning on above results is applied using a defined global parameter 'noise_face_max'. It removes all connected components (objects) from a binary image that have less than a defined number of pixels. Now label the cleaned image. The proposed algorithm uses defined ratios for scene height and width for searching a bounding box in an image. The bounding box must satisfy the conditions for height, width, portal and landscape ratios for a suitable bounding box for face detection. The proposed bounding box used in our approach is of size 100x100. Then the proposed algorithm applies region propping on the basis of labeled image and the characteristic of bounding box. Next it calculates the length of a propped region. The length of a propped region tells the predicted number of faces in an image. The final step is to search all the faces in an image with the help of eye template matching.

Traditional template matching approaches are very slow. Since the proposed algorithm is being designed for real time scenarios, therefore, timing is a crucial factor in it. So, template matching is applied in frequency domain instead of using time domain because in frequency domain both images must have the same size for applying template matching. So a resizing phenomenon is defined here for resizing an eye template according to the parameters of bounding box

of predicted face region. Then fast Fourier transform is applied on both images separately. To analyze the results of matching it must go back again in time domain because correlation does not work in frequency domain. As a result inverse of Fourier transform is applied. Following are the formulae used for FFT (Fast Fourier Transform) and IFFT (Inverse Fast Fourier Transform). (Xioguanglu *et al.*), (Frigo *et al.* 1998)

$$X_k = \sum_{j=1}^N x_j \omega_N^{(j-1)(k-1)}$$

$$x_j = (1/N) \sum_{k=1}^N X_k \omega_N^{-(j-1)(k-1)}$$

where, $\omega_N = e^{(-2\lambda i)/N}$ is an N^{th}

root of unity, N =number of samples.

For template matching a mathematical model used is represented as under:

$$SAD(x, y) = \sum_{i=0}^{T_{\text{rows}}} \sum_{j=0}^{T_{\text{cols}}} \text{Diff}(x+i, y+j, i, j) \quad (6)$$

Save the correlation results and analyze them. If the value of correlation coefficient comes out to be greater than 0.15 then it satisfies to be a face. The above proposed method is quite efficient for real time scenarios and has been tested. The detailed experimental results are provided in the results section.

3. PROCESSING OF ALGORITHM

The algorithm is processed as follows:

Step 1: Following image is taken as test (input) image for applying the proposed algorithm on it.



Fig. 1 Input Image



Fig. 2 HSV Color Space Image

Step 2: Following figure shows the converted image in to HSV color space.

Step 3: Following figures show the results of extraction of chromic red, chromic blue and hue elements respectively.



Fig. 3 Chromic Red Element Detected



Fig. 4 Chromic Blue Element Detected

Step 4 : Following figure shows the result of combining all three regions/elements produced above and it gives the required skin region in binary format.

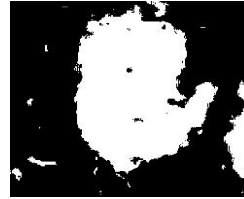


Fig. 5 Hue Element Detected

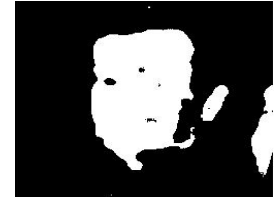


Fig. 6 Detected Skin region in binary

Step 5: Convert binary image back to RGB color space and use morphological filling for better results.

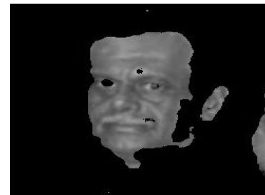


Fig. 7 RGB skin Detected



Fig. 8 Morphological Filled image

Step 6: In this step, eye template matching for face detection is used. If the correlation is higher than 0.65 then it satisfies the face otherwise it is a non face region. The template for eye and the detected face results are as follows:



Fig. 9 Eye Template



Fig. 10 Detected Face Region

4. RESULTS AND DISCUSSION

Following results for skin detection are taken on the basis of true point (T.P) and false point (F.P) respectively. Comparison with different techniques is shown in the following table. True points are calculated on the basis of hit ratio and false points on the basis of miss ratio.

Table 1 Skin detection comparison Existing results taken from (Aznavah *et al.* 2008)

Method	TP	FP
Bayes SPM in RGB [Jones and Rehg 1999]	80%	8.5%
	90%	14.2%
Bayes SPM in RGB [Brand and Mason 2000]	93.4%	19.8%
Maximum Entropy Model in RGB [Jedynak <i>et al.</i> 2002]	80%	8%
Gaussian Mixture Models in RGB [Jones and Rehg 1999]	80%	~9.5%
	90%	~15.5%
SOM in TS [Brown <i>et al.</i> 2001]	78%	32%
Elliptical boundary Model in CIE-xy [Lee and Yoo 2002]	90%	20.9%
Single Gaussian in CbCr [Lee and Yoo 2002]	90%	33.3%
Gaussian Mixture in IQ [Lee and Yoo 2002]	90%	30.0%

Thresholding [BrandandMason 2000]	94.7%	30.2%
A New Color Based Method for Skin Detection [18]	83.3%	~15.6%
Proposed Technique	95.5%	~14.6%

Results for face detection are taken on the basis of hit and false percentage of an algorithm. Following table shows the comparison of results of proposed algorithm with the previously defined techniques. Comparison shows that the proposed algorithm produces better results than the previous ones taken by existing techniques.

Table 2 Face detection comparison

Methods	Hit	False
The Proposed Technique	93.2%	6.8%
M2HD (Sim et. Al., 1999)	52.95%	47.05%
MEHD (Huttenlocher et. Al.1993)	66.3%	33.7%
HD (Saber et. Al.1996)	35.3%	64.7%
YES+HD (Srisuk et. Al.2000)	82.36%	17.64%
HSV+AMHD (Srisuk et. Al.2000)	88.24%	11.76%
AMHD (Xioguanglu et. Al.)	64.71%	35.29%

Following bar chart shows the comparison of hit and miss ratios of proposed technique with the existing techniques like M2HD, MEHD, HD, YES+HD, HSV+AMHD and AMHD. Y-axis shows the ratios in percentage of hit and misses for face detection.



Fig. 11. Face Detection comparison charts

Following shows the results of proposed face detection algorithm on single person images.



Fig. 12. Single Face Detection

5.

CONCLUSION

The HSV color space is used in this paper for face detection. On the basis of above experiments and results taken, it is found that the proposed technique marks the faces well in images as compared to the previous techniques. In future, this technique can be used in face recognition and surveillance systems, where high accuracy is required.

REFERENCES:

- Aznaveh, M.M., H. Mirzaei E. Roshan and M. Saraee (2008) "A new and improves skin detection method using RGB vector space " 5th International Multi-Conference on Systems, Signals and Devices, 2008. IEEE SSD 2008. 1-5.
- Feraud, R., O. J. Bernier, J. Viallet, and M. Collobert. (2001) "A fast and accurate face detector based on neural networks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (23): 42–53.
- Frigo, M., and S. G. Johnson, (1998) "FFTW: An Adaptive Software Architecture for the FFT," *Proceedings of the International Conference on Signal Processing*, Vol. (3): 1381-1384.
- Hjelmas, E. (2001) *Face detection: Survey*, *Computer Vision and Image Understanding*, (83): 236–274.
- Hsu, R. L., M. A. Mottaleb, and A. K. Jain. (2002) "Face detection in color images", *IEEE Trans. Pattern Analysis and Machine Intell.*, (24): 696–706.
- Huttenlocher, D. P. and W. J. Rucklidge, (1993) "Comparing image using the hausdorff distance", *IEEE Trans. on PAMI*, Vol. (15): 850-863.
- Jones, M. J. and J. M. Rehg, (1999) "Statistical color models with application to skin detection", *Computer vision and Pattern Recognition*, 274–280.
- Rajagopalan, A. N., K. Kumar, J. Karlekar, P. Poonacha, and S. Chaudhuri, (1998) "Finding faces in photographs", In *Proceedings sixth IEEE Int'l Conference on Computer Vision*, 640–645.
- Saber, E. (1996) "Face detection and facial feature extraction using color, shape and symmetry-based cost functions", *IEEE Int. Conf.* 654-658.
- Sandro, D. S., Johnson I. A., (2004), "Human face segmentation algorithms for mobile phones," *Proc. South African Telecommunications Networks and Applications Conference (SATNAC)* Vol. (2): Cape Town, South Africa, 373 – 374.
- Singh, R. K. (2002) "Background learning robust face recognition" In *Int'l Conference on Pattern Recognition*. Vol. (3): 305-325.
- Srisuk, S. and W. Kurutach, (2000) "Fast detection of scalable and multiple human faces" *Int. Conf. on Intell. Tech., Bangkok, Thailand*, 351-357.
- Takacs, B. and H. Wechsler (1998) "Fast searching of digital face libraries using binary image metrics", *IEEE Int. Conf. on Pattern. Recog.* 1235-1237.
- Xiaoguang L. (2003) "Image Analysis for Face Recognition – A brief survey," personal notes, May 2003 Dept. of com science and Engg, Michigan state university, East learning, MI 48824.