Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Technology 6 (2012) 1004 – 1010

Procedia Technology

2nd International Conference on Communication, Computing & Security [ICCCS-2012]

A Real-time Model for Multiple Human Face Tracking from Low-resolution Surveillance Videos

Rajib Sarkar[a], Sambit Bakshi[b], Pankaj K Sa[c]

[a]sarkar.rajibrcciit@gmail.com, [b]sambitbaksi@gmail.com, [c]pankajksa@gmail.com
[a,b,c] Department of Computer Science & Engineering, National Institute of Technology Rourkela, Orissa, India

Abstract

This article discusses a novel approach of multiple-face tracking from low-resolution surveillance videos. There has been significant research in the field of face detection using neural-network based training. Neural network based face detection methods are highly accurate, albeit computationally intensive. Hence neural network based approaches are not suitable for real-time applications. The proposed approach approximately detects faces in an image solely using the color information. It detects skin region in an image and finds existence of eye and mouth region in the skin region. If it finds so, it marks the skin region as a face and fits an oriented rectangle to the face. The approach requires low computation and hence can be applied on subsequent frames from a video. The proposed approach is tested on FERET face database images, on different images containing multiple faces captured in unconstrained environments, and on frames extracted from IP surveillance camera.

Keywords: real-time face tracking; skin detection; eye detection; mouth detection;

1. Introduction

Face detection from an image has been a challenging area of research due to the variability of human face with demographic change in skin-color (pinkish, yellowish, brownish etc), expression (smiling, crying, angry etc), illumination (outdoor or indoor lighting conditions, day and night time of capture etc), occlusion (partial covered face by long hair or sunglass etc), pose (close, distant position; frontal, profile orientation etc). Tracking a human face from a surveillance video is furthermore challenging because of two main reasons: 1. Resolution of surveillance video is low, 2. the face tracking algorithm should be low-computational enough to run in real-time environment to process the captured frames on-the-fly. There has been significant research on learning-based face detection methods as well as detection of face

doi:10.1016/j.protcy.2012.10.122

through color information. The article discusses the related approaches on face detection through skin color detection in Section 2 and illustrates how the proposed approach works better than the previous approaches in Section 3 by justifying the proposed approach with experimental results in Section 4.

2. Related Works

Through the last decade, the researchers have explored different genres of soft computing and image processing for detection of human face in an image. Wang et al. [1] have proposed a shape information based face detection algorithm. Craw et al. have further generated a shape template for finding faces in images [2]. Authors in [3] also describe an approach in the similar direction. In contrast, color components of a color image have been explored by Gomez et al. [4] for detecting skin region in an image. A severe problem in skin detection is the variance of illumination conditions in different images. An attempt towards detecting skin color in varying illumination has been made in [5]. Sun et al. have proposed a novel method of integrating color and local symmetry information for detecting face [6]. From the last two works mentioned, authors in [7] have proposed a method of prediction of color distribution for detecting skin under varying illumination condition. Kjeldsen et al. have integrated skin color information along with size and shape of skin to detect faces [8]. The issue of detection of face in presence of complex background has been discussed in [9,10]. Authors of [11,12] have also extracted facial features from facial images segmented in an image. Few more landmarks in the research of detecting face in color images can be found in [8,13-18]. Authors in [9,19-22] have explicitly dealt in multiple face detection in an image.

Subsequently the research has been directed towards tracking faces in a video foreseeing the application of this research in automated surveillance. [23] quotes one of the pioneer progress in this domain. Authors in [24,25] have discussed an algorithm towards a face tracking system working in real-time. [26,27] have further stepped forward to describe the application-oriented face tracker in to be deployed in videophone application and active vision system.

However none of these works mutually deals with detection of multiple faces from real-time video. The method proposed in this paper mutually deals with: 1. Detection of multiple faces, 2. Effective detection of faces in low-resolution, 3. Real-time low-computational algorithm for suitability for employing the algorithm for a surveillance video.

3. Proposed Work

The proposed color-based approach of face detection from color images works sequentially in three steps: a. detecting and segmenting skin regions in an image (explained in Algorithm. 1), b. checking each skin component whether it is a face (explained in Algorithm. 4) [by verifying existence of mouth (explained in Algorithm. 2) and eyes (explained in Algorithm. 3) in the skin components], c. fitting a oriented bounding box to the detected skin-region on the face depending on anthropometric shape of human face (explained in Algorithm. 5). Fig. 1 illustrates these steps with corresponding outputs as a flow-diagram.

This face detection algorithm checks for existence of a face in every skin component found in an image, and hence also capable to localize multiple faces from an image. As the face detection algorithm is computationally non-intensive, hence the algorithm can be deployed for frames captured from a video and face tracking can be achieved in real-time.

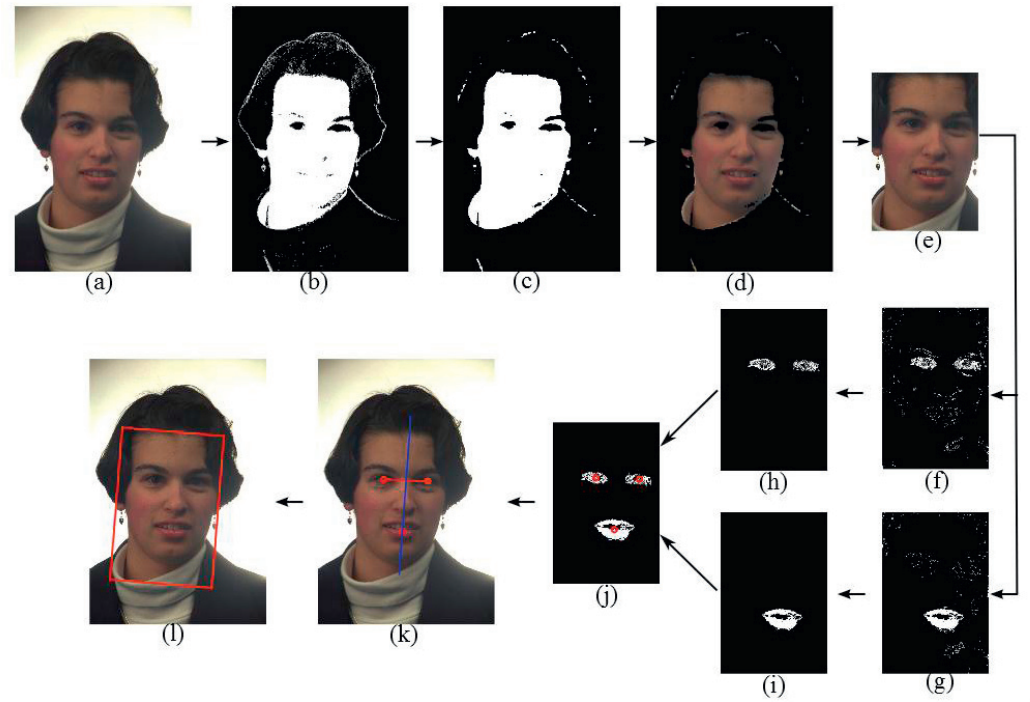

Fig. 1. (a) Input face image; (b) Noisy skin map; (c) Skin map after error correction by filtering; (d) Background subtracted face image; (e)Cropped skin region; (f) Detected eye region; (g) Detected mouth region; (h) Eye segmentation by finding largest connected component in (f); (i) Mouth segmentation by finding largest connected component in (g); (j) Finding midpoint of eyes and mouth; (k) finding orientation of face by connecting a line between midpoint of line connecting the eyes and the midpoint of mouth (blue line); (l) face image with oriented bounding box.

Algorithm 1: Skin_detection

Input: RGB image I of size $m \times n$

Output: Binary image S indicating skin-map

Step 1: Convert RGB image I to YC_bC_r color space

Step 2: Compute the average luminance value of image I as ${}_I Y_{avg} = \sum_{i=1}^{m} \sum_{j=1}^{n} {}_I Y_{i,j}$

Step 3: Normalize ${}_I Y_{i,j}$ to [0,255]

Step 4: Find edge-map using Canny operator from the input RGB image

Step 5: Brightness compensated image ${}_I C'$ is obtained as ${}_I C'_{i,j} = \{ {}_I R'_{i,j}, {}_I G'_{i,j}, {}_I B_{i,j} \}$

where ${}_I R'_{i,j} = ({}_I R_{i,j})^{\tau}$ and ${}_I G'_{i,j} = ({}_I G_{i,j})^{\tau}$ and $\tau = \begin{cases} 1.5 & if \quad {}_I Y_{avg} < 64 \\ 0.7 & if \quad {}_I Y_{avg} > 190 \\ 1 & otherwise \end{cases}$

Step 6: The skin map in the ${}_I C'$ is detected as $S_{i,j} = \begin{cases} 0 & if \quad \frac{R+1}{G+1} > 1.08 \quad \& \quad \frac{R+1}{B+1} > 1.08 \quad \& \quad G > 30 \quad \& \quad G < 140 \\ 1 & otherwise \end{cases}$

where $S_{i,j} = 0$ indicates skin region, and $S_{i,j} = 1$ non-skin region

---

Algorihm 2: Mouth_detection
Input: S
Output: MM
Step 1: For each connected component MC_i in S, repeat Step 2
Step 2: For each pixel in MC_i repeat Steps 3 and 4
Step 3: $\theta = \cos^{-1}\left(\frac{0.5(2R'-G'-B)}{\sqrt{(R'-G')^2+(R'-B)(G'-B)}}\right)$
Step 4: $MB_i = \begin{cases} 0, & \theta < 90 \\ 1, & otherwise \end{cases}$
where $MB_i = 0$ indicates mouth region, and $MB_i = 1$ non-mouth region
Step5: MM ={1} [size of the array MM is equal to that of MC]
Step 6: If all pixels in MB_i is 1 for a particular i, MM_i=0

Algorihm 3: Eyes_detection
Input: S
Output: EM
Step 1: For each connected component MC_i in S, repeat Step 2
Step 2: For each pixel in MC_i repeat Step 3
Step 3: $EB_i = \begin{cases} 0, & 65 < Y < 80 \\ 1, & otherwise \end{cases}$
where $EB_i = 0$ indicates eye region, and $EB_i = 1$ non-eye region
Step4: EM = {1} [size of the array is equal to that of MC]
Step 5: If all pixels in EB_i is 1 for a particular i, EM_i=0

Algorithm 4: Face_Detection
Input: S
Output: FM
Step 1: FM=S
Step 2: MM=Mouth_detection(S)
Step 3: EM=Eyes_detection(S)
Step4: for each connected component in S, repeat Steps 5 and 6
Step 5: if $MM_i = 0$ and $EE_i = 0$
Step 6: $FM = FM - i^{\text{th}}$ component in FM

Algorithm 5: Face_Bound
Input: FM, EM, MM
Output: Boundary to faces
Step1: For all connected components in FM_i
Step2: $(p_1,q_1)\leftarrow$ centre of left eye, $(p_2,q_2)\leftarrow$ centre of right eye
Step 3: $(p_3,q_3)\leftarrow$ centre of mouth
Step 4: $\alpha = \tan^{-1}\left(\frac{q_3 - \frac{q_1+q_2}{2}}{p_3 - \frac{p_1+p_2}{2}}\right)$
Step5: Draw rectangle enclosing FM_i with α alignment with the horizontal axis

4. Experimental Results

The proposed face detection algorithm is tested on 1000 randomly chosen 24-bit 512×768 color images of frontal faces from FERET face database. The dataset contains images of subjects with pinkish, yellowish and brown skin-color as the subjects belong to European, Asian and African regions. Conventional color-based skin detection algorithms working on pink skin fails on brown skin, and vice-versa. Fitting a threshold on Y value of pixels for detecting skin is tested by many researchers. However the same threshold does not work equally well for pink and brown skins. The proposed approaches of detecting faces from skin regions show an accuracy of 96.2% while detecting the face region accurately from 1000 randomly chosen FERET images. Fig. 2 depicts some sample localization of faces from FERET database.

The input face images are also resized to lower resolution using interpolation and the proposed algorithm is also applied to these low-resolution images of size 128×192. The face detection algorithm is also found working efficiently when the image is of low-resolution. This marks the suitability of the proposed algorithm for applying in images containing more than one faces captured from relatively more distance. Fig. 3 shows a sample output of multiple-face detection via proposed approach.

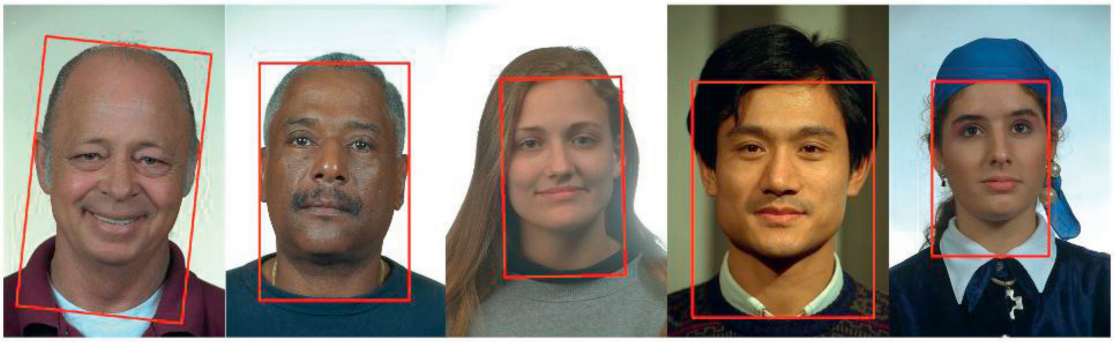

Fig. 2. Sample localization of faces from FERET database

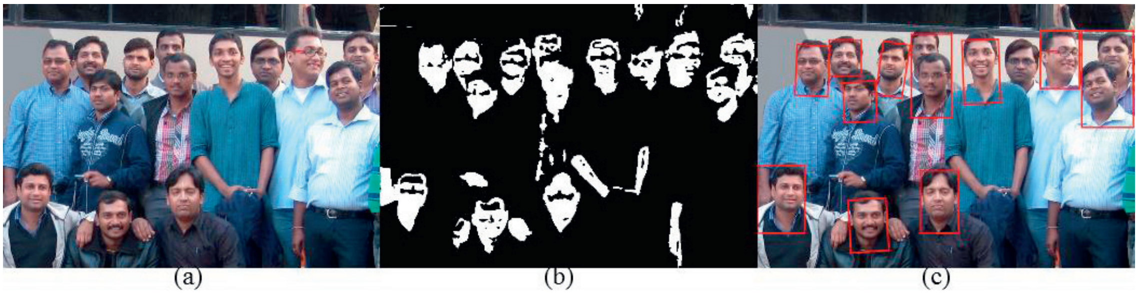

Fig. 3. (a) Input image consisting multiple faces; (b) corresponding skin map; (c) localized faces in the image

Furthermore, the proposed model of low-cost face detection is experimented on some low-resolution videos captured in unconstrained outdoor environments. The algorithm is capable of detecting multiple faces in a frame in real-time when there are more than one faces in the frame. However the algorithm does not detect very small faces that are far from the camera. Fig. 4 displays tracked faces from frames (resolution 315×315) captured with a time gap of 0.5 seconds of a video taken through surveillance IP camera.

5. Conclusion

The proposed model of real-time multiple face tracking considers tracking of multiple faces from color video. Color information of an image is used for detecting faces. The face detection is not based on transforming the image into any multidimensional space. Hence the process proposed is low-computational. The underlying face detection algorithm uses color information of the images, and hence cannot detect faces from grayscale images. The detected skins are filtered from noises by removing very small components detected as skin. Hence the model is unable to detect very small face images that are away from the camera. However the model is capable of detecting faces from low-resolution images or frames from a surveillance video. The model also finds the alignment of the faces in the images. This serves the purpose of image registration if the tracked faces are enrolled for further use.

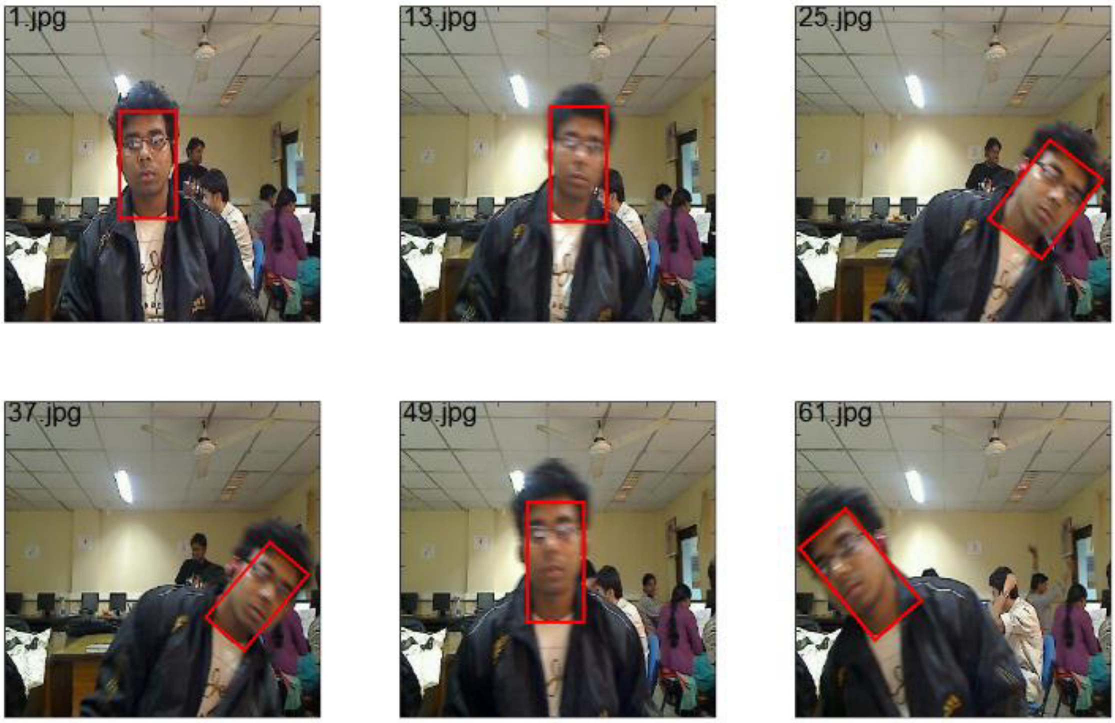

Fig. 4. Tracked faces from frames of a video taken through surveillance IP camera

Acknowledgements

The authors would like to thank all the co-researchers at Image Processing Laboratory at Department of Computer Science, National Institute of Technology Rourkela for their active and hidden cooperation towards manifestation of this research.

References

[1] Wang J, Tan T. A new face detection method based on shape information. *Pattern Recognition Letters* 2000;**21**(6–7), p. 463-471. doi:10.1016/S0167-8655(00)00008-8.

[2] Craw I, Tock D, Bennett A. Finding face features. Proc. 2nd European Conference on Computer Vision 1992. p.92 – 96.

[3] Burl MC, Leung TK, Perona. Face localization via shape statistics. *Proc. 1st International Workshop on Automatic Face and Gesture Recognition* 1995, p. 154-159.

[4] Gomez G. On selecting colour components for skin detection. *Proc. 16th International Conference on Pattern Recognition* 2002;**2**, p. 961- 964. doi: 10.1109/ICPR.2002.1048465.

[5] Storring M, Andersen HJ, Granum E. Skin color detection under changing lighting conditions. *Proc. 7th Symposium on Intelligent Robotics Systems* 1999, p.187-195.

[6] Sun QB, Huang WM, Wu JK. Face detection based on color and local symmetry information. *Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition* 1998, p. 130-135. doi: 10.1109/AFGR.1998.670937.

[7] Sigal L, Sclaroff S, Athitsos V. Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 2000;**2**, p. 152-159. doi: 10.1109/CVPR.2000.854764.

[8] Kjeldsen R, Kender J. Finding skin in color images. *Proc. 2nd International Conference on Automatic Face and Gesture Recognition* 1996, p. 312-317. doi: 10.1109/AFGR.1996.557283.

[9] Dai Y, Nakano Y. Extraction for facial images from complex background using color information and SGLD matrices. *Proc. 1st International Workshop on Automatic Face and Gesture Recognition* 1995, p. 238-242.

[10] Yang G, Huang TS. Human face detection in a complex background. *Pattern Recognition* 1994;**27**(1), p. 53-63. doi: 10.1016/0031-3203(94)90017-5.

[11] Garcia C, Tziritas G. Face detection using quantized skin color regions merging and wavelet packet analysis. *IEEE Transactions on Multimedia*1999;**1**(3), p. 264-277. doi: 10.1109/6046.784465.

[12] Wu H, Yokoyama T, Pramadihanto D, Yachida M. Face and facial feature extraction from color image. *Proc. 2nd International Conference on Automatic Face and Gesture Recognition* 1996, p. 345-350. doi: 10.1109/AFGR.1996.557289.

[13] Yang MH, Ahuja N. Detecting human faces in color images. *Proc. International Conference on Image Processing* 1998;**1**, p. 127-130. doi: 10.1109/ICIP.1998.723442.

[14] Yang J, Lu W, Waibel A. Skin-Color Modeling and Adaptation. *Proc. 3rd Asian Conference on Computer Vision*. Springer-Verlag. 1998, p. 687-694.

[15] Hsu RL. Face detection and modeling for recognition. *PhD Thesis*. Michigan State Uviversity, United States of America, 2002.

[16] Hsu RL, Abdel-Mottaleb M, Jain AK. Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2002;**24**(5), p. 696-706. doi: 10.1109/34.1000242.

[17] Jones MJ, Rehg JM. Statistical color models with application to skin detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1999;**1**(2), p. 274-280.doi: 10.1109/CVPR.1999.786951.

[18] Kim SH, Kim NK, Ahn SC, Kim HG. Object oriented face detection using range and color information. *Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition* 1998, p. 76-81. doi: 10.1109/AFGR.1998.670928.

[19] Menser B, Wien M. Segmentation and tracking of facial regions in color image sequences. *Proc.Visual Communication and Image Processing (SPIE-VCIP)* 2000, p. 731-740.

[20] Cai J, Goshtasby A, Yu C. Detecting human faces in color images. *Proc. International Workshop on Multi-Media Database Management Systems* 1998, p. 124-131. doi: 10.1109/MMDBMS.1998.709526.

[21] Govindaraju V. Locating human faces in photographs. *International Journal of Computer Vision* 1996;**19**(2), p. 129-146. Doi:10.1007/BF00055801.

[22] Govindaraju V, Sher DB, Srihari RK, Srihari SN. Locating human faces in newspaper photographs. *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1989, p. 549-554. doi: 10.1109/CVPR.1989.37900.

[23] Saxe D, Foulds R. Toward robust skin identification in video images. *Proc. 2nd International Conference on Automatic Face and Gesture Recognition* 1996, p. 379-384. doi: 10.1109/AFGR.1996.557295.

[24] Qian RJ, Sezan MI, Matthews KE. A robust real-time face tracking algorithm. *Proc. International Conference on Image Processing* 1998;1, p. 131-135. doi: 10.1109/ICIP.1998.723443.

[25] Yang J, Waibal A. A real-time face tracker. *Proc. 3rd Workshop Applications of Computer Vision* 1996, p. 142-147.

[26] Chai D, Ngan KN. Face segmentation using skin-color map in videophone applications. *IEEE Transactions on Circuits and Systems for Video Technology* 1999;**9**(4), p. 551-564. doi: 10.1109/76.767122.

[27] Scassellati B. Eye finding via face detection for a foevated, active vision system. *Proc. 15th National Conference on Artificial Intelligence* 1998, p. 1146-1151.