COMBINING HAAR FEATURE AND SKIN COLOR BASED CLASSIFIERS FOR FACE DETECTION

C. E. Erdem *, S. Ulukaya, A.Karaali*

Bahcesehir University
Dept. Electrical and Electronics Engineering
Istanbul, Turkey
{cigdem.eroglu, sezer.ulukaya, ali.karaali}@bahcesehir.edu.tr

A. T. Erdem

Ozyegin University
Dept. Electrical and Electronics Engineering
Istanbul, Turkey
Tanju.Erdem@ozyegin.edu.tr

ABSTRACT

This paper presents a hybrid method for face detection in color images. The well known Haar feature-based face detector developed by Viola and Jones (VJ), that has been designed for gray-scale images is combined with a skin-color filter, which provides complementary information in color images. The image is first passed through a Haar-Feature based face detector, which is adjusted such that it is operating at a point on its ROC curve that has a low number of missed faces but a high number of false detections. Then, using the proposed skin color post-filtering method many of these false detections can be eliminated easily. We also use a color compensation algorithm to reduce the effects of lighting. Our experimental results on the Bao color face database show that the proposed method is superior to the original VJ algorithm and also to other skin color based pre-filtering methods in the literature in terms of precision.

Index Terms— face detection, skin color detection, adaboost, haar features

1. INTRODUCTION

The problem of face detection refers to determining whether or not there are any faces in a given image and to estimate the location and size of any detected faces [1]. Face detection is a trivial task for humans, however it is not very easy for computers due to geometric (scale, in-plane rotation, pose, facial expressions, occlusion etc.) and photometric variations. In the next subsection, face detection algorithms in the literature will be briefly reviewed.

1.1. Previous Work

Methods in the literature on face detection can be grouped as knowledge-based, feature-based, template-based and appearance-based methods [1, 2, 3]. Face detection is an expensive search problem. In general, a sliding window is scanned through an image at various scales to classify the window as face or non-face. Therefore, many background windows need to be processed as well as actual face regions. The ratio of the number of non-face windows to face windows can be as high as 100000:1. Hence, a well trained classifier is necessary that will produce a low number of false positives.

Face detection methods based on learning algorithms have shown good results [4]. Viola and Jones (VJ) [4] proposed a frontal face detection system in gray-scale images based on the Adaboost learning algorithm. More details about this algorithm will be provided in section 2.

The number of false detections in VJ algorithm increases when a high true detection rate is desired. For example, for a database containing 507 faces, there are over 150 false positives to achieve a true detection rate of about 93% [4]. This false positive rate is too high for some applications. The VJ face detector has been reported to fail if the face is tilted beyond about ± 15 degrees in plane, rotated beyond about ± 45 degrees out of plane, towards a profile view [4]. The work of Viola and Jones (VJ) have been extended to handle multi-pose (frontal to profile) faces [5] and in-plane rotation [6].

Skin-color is an effective cue for face detection since it is highly invariant to geometric variations of the face (pose, facial expressions) and fast to process [7]. Skin-color has been shown to be useful for face detection [8, 9]. There are some approaches in the literature that try to combine the VJ face detector with skin-color detector, which are mainly based on pre-filtering [9, 10, 11]. Instead of making an exhaustive search on the whole image, VJ algorithm is applied around probable face regions, where skin-color pixels are highly populated. This results in an improvement in speed and a decrease in false positive rate. In [12], the authors use a pre-filtering approach to detect candidate face regions and then use a hybrid set of features consisting of Haar-like and Gabor Features to train various classifiers for faces in multiple poses.

1.2. Contributions and Outline of the Paper

In this paper, we propose a method that utilizes skin color detection to decrease the high false positive rate of the VJ face detector [4]. The VJ algorithm uses only the brightness information in a search window, resulting in a high false acceptance rate due to face-like brightness patterns in the background. Therefore, skin-color is a complementary channel of information, and it is very fast to process.

In order to achieve a low false detection rate while keeping a high true detection rate, we propose a skin-color based post-filtering method for color images. The windows that are detected as face by the VJ algorithm are verified if the window contains sufficient number of skin pixels. To maximize the overall true detection rate, we adjust the parameters of the VJ algorithm such that the number of misses is low, and the number of false detections is high. Most of the false detections are easily eliminated by the proposed skin-color based post-filtering method.

In order to reduce the effects of illumination, we also use a color compensation method before the skin-color detection step to improve the effectiveness of skin color detection, which was not present in previous pre-filtering based approaches [9, 10].

The organization of the paper is as follows. In Section 2, background information is provided. In Section 3, the proposed method for face detection is presented. Then in Section 4, experimental re-

*This work was supported by Turkish Scientific and Technical Research Council (TUBITAK) under project EEAG-110E056 and Bahcesehir University Research Fund.

sults are provided, which is followed by conclusions in Section 5.

2. BACKGROUND

2.1. Adaboost Based Face Detection using Haar Like Features

Viola & Jones [4] have presented a face detection method based on an over-complete set of Haar-like features which are calculated in scaled analysis windows. The rectangular Haar-like features are sensitive to edges, bars and other similar structures in the image and they are computed using an efficient method based on the integral image concept. After calculation of a huge number of features for each analysis window, the AdaBoost algorithm is used for combining a small number of these features to form an effective classifier. For example, for an analysis window of size 24×24, there are approximately $160,000$ features, far more than the number of pixels. A variant of AdaBoost is used both to select the best features and to train the final classifier [13].

2.2. Skin Color Detection

The approaches for skin segmentation in the literature have been summarized in several survey papers [14, 15]. Two methods for skin color detection have been tested in this work, which are described below.

Bayesian Classifier with the Histogram Technique: The first skin color detection method that is used is based on a Bayesian classifier with histogram technique in RGB space. This method has been reported to be superior to other methods in terms of accuracy and computational cost for classifying pixels as skin and non-skin [16, 17]. Using the likelihood ratio method, a pixel with a color vector $\underline{c}$ is classified as a skin pixel if:

$$\frac{P(\underline{c}|skin)}{P(\underline{c}|nonskin)} \geq \tau, \tag{1}$$

where $P(\underline{c}|skin)$ and $P(\underline{c}|nonskin)$ are the class conditional pdfs of skin and nonskin color distributions, respectively. The threshold τ is theoretically defined by the ratio $\tau = P(nonskin)/P(skin)$, where the prior probabilities can be estimated from the training set. In practice, the threshold τ is determined empirically, giving a trade-off between the number of false positives and false negatives. A value around $\tau = 0.4$ gives good results in our experiments.

Explicitly Defined Skin Color Detector: The second skin color detection method that we test is based on a set of rules on R, G, B color components of a pixel. In order to detect the skin colors for a fair complexion under uniform daylight illumination, the following set of rules have been found to be superior to other models under some constraints [18]. A pixel with color components (R, G, B) is detected as skin if the conditions given in (2) below hold. The second line in (2) ensures that RGB components must not be close together, which ensures greyness elimination. The third line in (2) ensures that R and G components must not be close together, which must be true for fair complexion [18].

$$\begin{gathered} R > 95 \ \& \ G > 40 \ \& \ B > 20 \ \& \\ max\{R, G, B\} - min\{R, G, B\} > 15 \ \& \\ |R - G| > 15 \ \& \ R > G \ \& \ R > B. \end{gathered} \tag{2}$$

3. COMBINING HAAR FEATURES AND SKIN COLOR BASED FACE DETECTORS

Our motivation in this work is to decrease the false positive rate of the VJ face detector [4]. The flowchart in Figure 1 shows the steps followed to decrease the false positive rate. Given a color image possibly containing a number of faces, the first step is to apply an illumination compensation algorithm with the goal of reducing the effects of lighting. Then, the image is passed through the VJ detector and a skin-pixel detector. In the next step, the analysis windows that are detected as face by the VJ algorithm are verified by a skin-color based method. Below each step shown in Figure 1 will be explained in more detail.

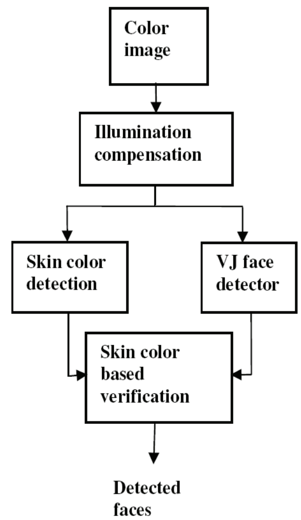


Fig. 1. The block diagram of the proposed face detection method.

3.1. Illumination Compensation

Illumination compensation is important for eliminating the effects of non standard illumination for skin color detectors. In this work, we use a color compensation based on the Gray World method [19], which is fast and simple to implement in RGB color space. This method depends on the assumption that the average surface color on the image, which is reflected from the surfaces corresponds to the illumination. The algorithm consists of the following steps: i) Calculate the averages of each color channel R, G, B for the whole image to get R_{ave}, G_{ave}, and B_{ave}. ii) Calculate a linear scaling factor for each color component $S_R = R_{std}/R_{ave}$, $S_G = G_{std}/G_{ave}$, $S_B = B_{std}/B_{ave}$, where $(R_{std}, G_{std}, B_{std})$ denotes the standard gray value.

3.2. Skin Color Based Verification

In order to detect the skin-colored pixels in the illumination compensated image, we implemented the two methods described in Section 2.2: Bayesian classifier with the histogram technique and the explicitly defined skin color detector.

The Bayesian classifier with the histogram technique requires a training step in order to estimate the class conditional pdfs $P(\underline{c}|skin)$ and $P(\underline{c}|nonskin)$. We estimated these pdfs with the RGB color histograms of skin and nonskin pixels using the Compaq skin database [16]. This database contains wide variability in lighting (indoor/outdoor), background and skin types (white/yellow/brown skins).

In order to estimate the class conditional pdfs of skin and nonskin pixels using histograms, let $\underline{c}$ denote the color vector of a pixel, i.e. $\underline{c} = [r, g, b]$. First we find $H_{skin}(\underline{c})$ and $H_{nonskin}(\underline{c})$, which denote the color histograms of the pixels labeled as skin and nonskin in the training set. Then the histograms are normalized by dividing by the total number of skin and nonskin pixels. Finally, the skin pixels are detected by applying (1).

Using either the Bayesian classifier or the explicit detector, binary skin color masks are generated, where skin color pixels are denoted by 1 and nonskin pixels are denoted by 0. Let S_j denote binary the skin color mask for image j, where M is the total number of images and $j = 1, \ldots, M$. Let V_j^i denote the i^{th} detected window by the Viola and Jones method for image j, where $i = 1, \ldots, N$ and N denotes the total number of detected windows claimed as face by VJ method.

Given the detection windows generated by the VJ algorithm and the binary skin pixel masks, the skin color based verification step is carried out as follows:

1. Count the number of skin colored pixels C_j^i in window V_j^i:

$$C_j^i = \sum_{(x,y)\in V_j^i} S_j, \tag{3}$$

2. Verify window V_j^i as face if:

$$\frac{C_j^i}{w_j^i \times h_j^i} \geq \mu, \tag{4}$$

   w_j^i and h_j^i denote the width and height of window V_j^i, and μ is a threshold, which is determined experimentally. A value around $\mu = 0.2$ gives good results during the experiments.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed method, we used the Bao face database [20, 21], which consists of color images containing single and multiple frontal and non-frontal faces with a cluttered background. We used the first 100 images of this database containing a total of 859 faces. Since the original Bao database does not contain the ground truth information for the face locations, we first marked ground-truth faces manually in each image by forming a rectangle using the outer corner of the right eye and the left corner of the mouth (see Figure 2). If this rectangle is completely within the face detection window generated by VJ algorithm, we define that window as a correct detection. The ground truth data is available from [22].

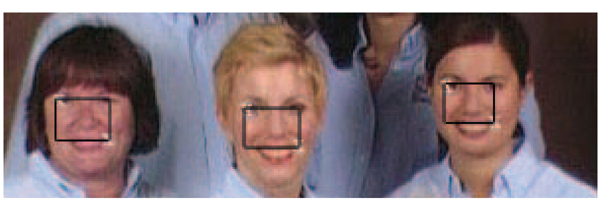


Fig. 2. The manual labeling of the true face locations for the Bao face database. A part of image 22 is shown. A rectangle is formed by manually labeling the outer corner of the right eye and the left corner of the mouth, which are shown by the white x signs.

We used pre-trained implementation of the the Viola and Jones face detector in OpenCV library [23]. In order to combine the Haar-feature based face detector with the skin-color based post-filtering method, we adjusted the parameters of VJ such that the number of missed faces is as small as possible giving a high correct detection rate. This gives us a high number of false positives, but we expect to eliminate them with the skin color based verification step.

Table 1. Performance comparison of seven face detection methods using first 100 images of Bao dataset. The first four rows show the results of the proposed method.

Method	TP	FN	FP	Recall	Precision
Bayesian	812	47	39	94.53	95.42
Bayesian-I	811	48	36	94.41	**95.75**
Explicit	812	47	46	94.53	94.64
Explicit-I	810	49	40	94.30	95.29
VJ [4]	813	46	237	94.65	77.43
VJ-EER [4]	770	89	83	89.64	90.27
Pre-filter [10]	760	99	77	88.48	90.80

In Table 1, we compare the face detection performances of seven methods denoted with the following acronyms:

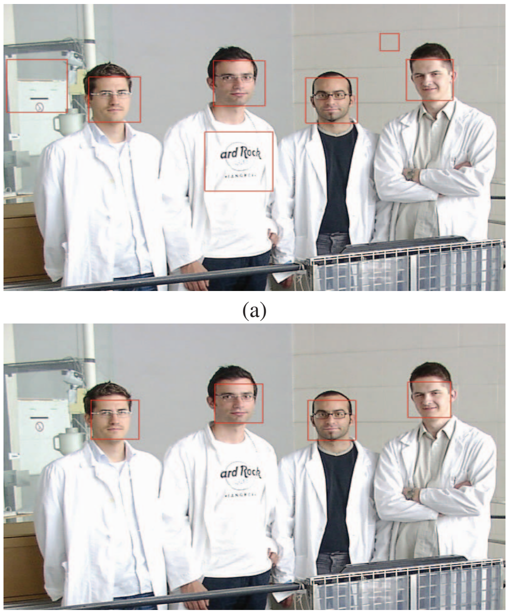
1. **Bayesian**: Our skin color based post-filtering method using a Bayesian skin classifier.
2. **Bayesian-I**: Our skin color based post-filtering method using a Bayesian skin classifier after illumination compensation.
3. **Explicit**: Our skin color based post-filtering method using an explicit skin classifier.
4. **Explicit-I**: Our skin color based post-filtering method using an explicit skin classifier after illumination compensation.
5. **VJ**: The original Viola and Jones algorithm [4].
6. **VJ-EER**: The original Viola and Jones algorithm [4] operating at the equal error rate point of the ROC curve.
7. **Pre-Filter**: The skin-color based pre-filtering method [10].

The acronyms used in the table TP, FP, and FN denote the number of true positives, false positives and false negatives, respectively. The precision and recall are defined as $Precision = TP/(TP + FP)$ and $Recall = TP/(TP + FN)$. We can observe from Table 1 that the highest precision is achieved by the proposed Bayesian-I method, which is the post filtering method using a Bayesian skin classifier. This precision is much higher than that of the VJ [4] and skin pre-filtering [10] methods.

Since the skin color filter is not perfect, it may miss some skin colored pixels. Therefore, the integrity of the face region may not be preserved if the skin color is used as a pre-filter. This causes an increase in the miss (FN) rate of the VJ algorithm which follows the skin color pre-filter, as can be observed in the last row of Table 1. However, this effect is not observed if the skin color is used as a post-filter as proposed in this paper.

If we compare the proposed Bayesian-I and VJ methods given in second and fifth rows of Table 1, we can see that the precision increased from 77.43 % to 95.75 %, while keeping the recall rate almost the same. We can also observe that using an illumination compensation step is also beneficial in terms of increasing precision.

In Figure 3 (a), the face detection results of the VJ algorithm are shown for image 26 of the Bao database, where we can see three false detections. In Figure 3 (b), the result after the proposed skin color based post-filtering is shown, where all false positives have been successfully eliminated.



(a)

(b)

Fig. 3. (a) The face detection results of the VJ [4] algorithm for image 26 of the Bao database are shown with red squares. (b) The face detection results after skin color based post-filtering. All false detections have been eliminated.

5. CONCLUSIONS AND FUTURE WORK

We presented a method for combining the Haar feature based face detector [4] which uses brightness information with a skin-color classifier in a post-processing framework. We compared two methods for skin pixel classification: Bayesian method with the histogram technique and the explicit method. We also used an illumination compensation step prior to skin color detection. The experimental results on the Bao color face image dataset show that the skin-color post-filtering method using the Bayesian classifier is superior to the original VJ [4] algorithm and a pre-filtering method in the literature [10].

We plan to do experiments on more extensive color face databases. We also plan to improve the skin-color based face detection algorithm to further decrease the number of false negatives.

6. REFERENCES

[1] M. H. Yang, D. J. Kriegman, and N. Ahuja, “Detecting faces in images: A survey,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.

[2] E. Hjelmas and B. K. Low, “Face detection: A survey,” *Computer Vision and Image Understanding*, vol. 83, no. 3, pp. 236–274, 2001.

[3] Cha Zhang and Zhengyou Zhang, “A survey of recent advances in face detection,” Tech. Rep., Microsoft Research, 2010.

[4] P. Viola and M. J. Jones, “Robust real-time face detection,” *Int. J. of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[5] J. P. Niyoyita, Z. H. Tang, and L. J. Ping, “Multi-view face detection using six segmented rectangular features,” in *The Sixth International Symposium on Neural Networks (ISNN 2009)*, 2009, vol. 56, pp. 333–342.

[6] B. Wu, H. Z. Ai, C. Huang, and S. H. Lao, *Fast rotation invariant multi-view face detection based on real Adaboost*, Sixth IEEE Int. Conf. on Automatic Face and Gesture Recognition. 2004.

[7] M. Al Haj, A. D. Bagdanov, J. Gonzalez, and X. F. Roca, “Robust and efficient multipose face detection using skin color segmentation,” in *Lecture Notes in Computer Science*, vol. 5524, pp. 152–159. 2009.

[8] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, “Face detection in color images,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696–706, 2002.

[9] L. Shobana, A. K. Yekkala, and S. Eajaz, “Face detection using skin segmentation as pre-filter,” in *Int. Conf. on Advances in Pattern Recognition*, 2007.

[10] Z. S. Tabatabaie, R. W. Rahmat, N. I. B. Udzir, and E. Kheirkhah, “A hybrid face detection system using combination of appearance-based and feature-based methods,” *Int. J. of Computer Science and Network Security*, vol. 9, no. 5, 2009.

[11] B. Kim, S. Ban, and M. Lee, “Improving adaboost based face detection using face-color preferable selective attention,” *Lecture Notes in Computer Science*, vol. 5326, pp. 88–95, 2008.

[12] H. Y. Chen, C. L. Huang, and C. M. Fu, “Hybrid-boost learning for multi-pose face detection and facial expression recognition,” *Pattern Recognition*, vol. 41, no. 3, pp. 1173–1185, 2008.

[13] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” *J. of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[14] V. Vezhnevets, V. Sazonov, and Andreeva A., “A survey on pixel-based skin color detection techniques,” *Proc. Graphicon*, 2003.

[15] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, “A survey of skin-color modeling and detection methods,” *Pattern Recognition*, vol. 40, no. 3, pp. 1106–1122, 2007.

[16] M. J. Jones and J. M. Rehg, “Statistical color models with application to skin detection,” *Int. J. of Computer Vision*, vol. 46, no. 1, pp. 81–96, 2002.

[17] S. L. Phung, A. Bouzerdoum, and D. Chai, “Skin segmentation using color pixel classification: Analysis and comparison,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 148–154, 2005.

[18] F. Solina, P. Peer, B. Batagelj, S. Juvan, and J. Kovac, “Color-based face detection in the 15 seconds of fame art installation,” in *Proceedings of Mirage*, 2003.

[19] B. Funt, K. Barnard, and L. Martin, “Is machine colour constancy good enough?,” in *European Conference on Computer Vision*, 1998, pp. 445–459.

[20] X. Wang, H. Xu, H. Wang, and H. Li, “Robust real-time face detection with skin color detection and the modified census transform,” in *Proc. of IEEE Int. Conf. on Information and Automation*, China, 2008.

[21] R. Frischholz, “Bao face database at the face detection homepage,” in *http://www.facedetection.com*.

[22] S. Ulukaya, “Ground truth faces for bao database,” in *http://staff.eng.bahcesehir.edu.tr/c̃igdemeroglu/groundTruthBAO.rar*.

[23] OpenCV, ,” in *http://sourceforge.net/projects/opencvlibrary/*.