A Hybrid Face Detection System using combination of Appearance-based and Feature-based methods

Zahra Sadri Tabatabaie[1], Rahmita Wirza Rahmat[2], Nur Izura Binti Udzir[3] and Esmaeil Kheirkhah[4]

Faculty of Computer Science and Information Technology, University Putra Malaysia[1,2,3]
Faculty of Information Science and Technology, University Kebangsaan Malaysia[4]

Summary
Human face detection is preliminary required step of face recognition systems as well as a very important task in many applications, such as security access control systems, video surveillance, human computer interface and image database management. This paper intends to combine Viola and Jones face detection method with a color-based method to propose an improved face detection method. Experimental results show that our method efficiently decreased false positive rate and subsequently increased accuracy of the face detection system especially in complex background images. Also our proposed method considerably increased face detection speed.

Key words:
Face detection, Feature-based, Appearance-based, Skin color classification.

1. Introduction

Face detection and tracking has been the topics of an extensive research for the several past decades. The development of robust face detection systems is quite essential in a variety of applications such as robotics, security systems, intelligent human-computer interfaces, etc. A number of face detection algorithms such as those using Eigen-faces [5] and neural networks [7], for instance, have been developed. In these algorithms, however, a large amount of numerical computation is required, making the processing extremely time-consuming.

There are various algorithms including skin color based algorithms for detecting faces within an image. Color is an important feature of human faces. Using skin-color as a feature for tracking a face can help to increase detection rate because color processing is usually much faster than processing other facial features. However, color-based face detection algorithms have several problems like the color representation of a face obtained by a camera is influenced by many factors (ambient light, object movement, etc.), different cameras produce significantly different color values even for the same person under the same lighting conditions and skin color differs from person to person [11].

In color-based face detection systems we are usually facing to three main problems, that is, what color-space is better to choose, how exactly the skin color distribution should be modeled, and finally, what will be the way of processing of color segmentation results for face detection [12].
This paper discuss a pixel-based skin detection methods, that classify each pixel as skin or non-skin individually, independently from its neighbors and combine it with Viola and Jones based face detection to improve performance of face detection systems in terms of increasing the face detection speed and decreasing false positive rate.

2. Appearance-based face detection

In general, appearance-based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face and non-face images. The learned characteristics are in the form of distribution models or discriminant functions that are consequently used for face detection. Meanwhile, dimensionality reduction is usually carried out for the sake of computation efficiency and detection efficacy. For Example, face detection techniques include the AdaBoost algorithm [13][14], the S-AdaBoost algorithm [2] , the FloatBoost algorithm [4], neural networks [7][1], Support Vector Machines (SVM) [6][10], and the Bayes classifier [8][9].

Viola and Jones [13][14] propose a robust AdaBoost face detection algorithm, which can detect faces in a rapid and robust manner with a high detection rate. Li et al. [4] propose the FloatBoost algorithm, an improved version of the AdaBoost algorithm, for learning a boosted classifier with minimum error rate. The FloatBoost algorithm uses a backtracking mechanism to improve the face detection rate after each iteration of the AdaBoost procedure. However this method is computationally more intensive than the AdaBoost algorithm. Jiang and Loe [2] propose the S-AdaBoost algorithm, a variant of the AdaBoost algorithm, for handling outliers in pattern detection and classification. Since the S-AdaBoost algorithm uses different classifiers

Manuscript received May 5, 2009
Manuscript revised May 20, 2009

in different phases of computation, it suffers from computational inefficiency and lack of accuracy.

Viola and Jones based face detection: The basic principle of the Viola-Jones algorithm is to scan a sub-window capable of detecting faces across a given input image. The standard image processing approach would be to rescale the input image to different sizes and then run the fixed size detector through these images. This approach turns out to be rather time consuming due to the calculation of the different size images. Contrary to the standard approach Viola-Jones rescale the detector instead of the input image and run the detector many times through the image – each time with a different size. At first one might suspect both approaches to be equally time consuming, but Viola-Jones have devised a scale invariant detector that requires the same number of calculations whatever the size. This detector is constructed using a so-called integral image and some simple rectangular features reminiscent of Haar wavelets. The next section elaborates on this detector.

The first step of the Viola-Jones face detection algorithm is to turn the input image into an integral image. This is done by making each pixel equal to the entire sum of all pixels above and to the left of the concerned pixel. This allows for the calculation of the sum of all pixels inside any given rectangle using only four values. These values are the pixels in the integral image that coincide with the corners of the rectangle in the input image. This is demonstrated in Figure 1.

Sum of grey rectangle = D - (B + C) + A

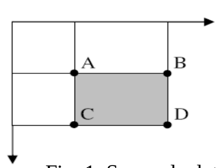


Fig. 1 Sum calculation.

The Viola-Jones face detector analyzes a given sub-window using features consisting of two or more rectangles. The different types of features are shown in Figure 2.

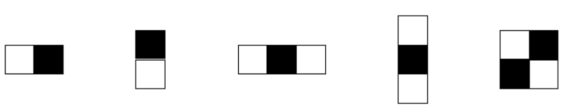


Fig. 2 Different types of features.

Each feature results in a single value which is calculated by subtracting the sum of the white rectangle(s) from the sum of the black rectangle(s). Viola and Jones have empirically found that a detector with a base resolution of 24*24 pixels gives satisfactory results.

Viola and Jones used a simple and efficient classifier built from computationally efficient features using AdaBoost for feature selection. AdaBoost is a machine learning boosting algorithm capable of constructing a strong classifier through a weighted combination of weak classifiers. (A weak classifier classifies correctly in only a little bit more than half the cases.) To match this terminology to the presented theory each feature is considered to be a potential weak classifier. A weak classifier is mathematically described as:

$$h(x, f, p, \theta) = \begin{cases} 1 & \text{if } pf(x) > p\theta \\ 0 & \text{otherwise} \end{cases}$$

Where x is a 24*24 pixel sub-window, f is the applied feature, p the polarity and θ the threshold that decides whether x should be classified as a positive (a face) or a negative (a non-face).

The basic principle of the Viola-Jones face detection algorithm is to scan the detector many times through the same image – each time with a new size. Even if an image should contain one or more faces it is obvious that an excessive large amount of the evaluated sub-windows would still be negatives (non-faces). This realization leads to a different formulation of the problem:

Instead of finding faces, the algorithm should discard non-faces.

The thought behind this statement is that it is faster to discard a non-face than to find a face. With this in mind a detector consisting of only one (strong) classifier suddenly seems inefficient since the evaluation time is constant no matter the input. Hence the need for a cascaded classifier arises. The cascaded classifier is composed of stages each containing a strong classifier.

The job of each stage is to determine whether a given sub-window is definitely not a face or maybe a face. When a sub-window is classified to be a non-face by a given stage it is immediately discarded. Conversely a sub-window classified as a maybe-face is passed on to the next stage in the cascade. It follows that the more stages a given sub-window passes, the higher the chance the sub-window actually contains a face. The concept is illustrated with two stages in Figure 3.

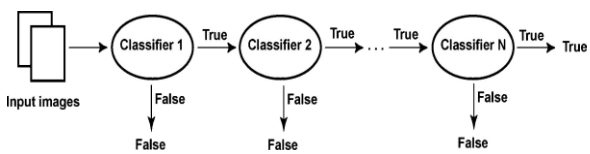


Fig. 3 the cascade classifier

3. Feature-based Face Detection

In feature-based approaches researchers have been trying to find invariant features of faces for detection. The underlying assumption is based on the observation that humans can effortlessly detect faces and objects in different poses and lighting conditions and, so, there must exist properties or features (such as eyebrows, eyes, nose, mouth, and skin color) which are invariant over these variabilities. Numerous methods have been proposed to first detect facial features and then to infer the presence of a face. Based on the extracted features, a statistical model is built to describe their relationships and to verify the existence of a face. In this paper skin color feature will be discussed and used.

Skin Color Classification: The study on skin color classification has gained increasing attention in recent years due to the active research in content-based image representation. For instance, the ability to locate image object as a face can be exploited for image coding, editing, indexing or other user interactivity purposes. Moreover, face localization also provides a good stepping stone in facial expression studies.

It would be fair to say that the most popular algorithm to face localization is the use of color information, whereby estimating areas with skin color is often the first vital step of such strategy. Hence, skin color classification has become an important task. Much of the research in skin color based face localization and detection is based on RGB, YCbCr and HSI color spaces. This paper uses RGB color space for skin color classification.

The RGB colorspace is a three-dimensional color space whose components are the red, green, and blue intensities that make up a given color. The colors red, green, and blue were chosen because each one corresponds roughly with one of the three types of color-sensitive cones in the human eye. It is one of the most widely used colorspaces for processing and storing of digital image data.

4. Hybrid Face Detection

Our proposed approach for improving face detection systems is combination of an appearance-based and feature-based face detection system. Due to high speed and accuracy in detecting faces in an image, we chose Viola and Jones proposed face detection system. In order to improve the face detection speed rate and decrease false positive rate in this approach we combined it with a (skin) color-based method. This approach is very simple but it works well. Figure 4 demonstrates overall schema of the proposed face detection system.

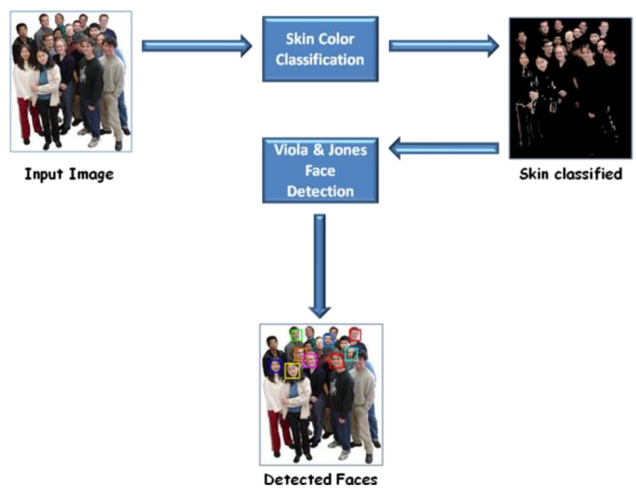


Fig. 4 – Overall schema of our method

One method to build a skin classifier is to define explicitly (through a number of rules) the boundaries skin cluster in some colorspace. For example in [3]:

(R,G,B) is classified as skin if:

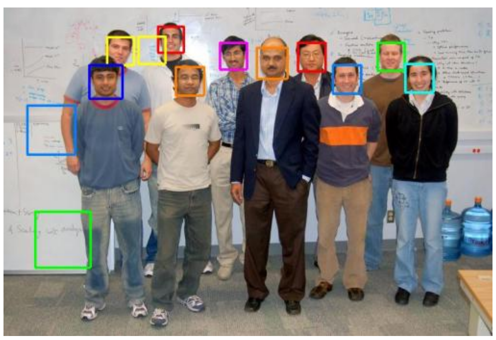
$R > 95$ and $G > 40$ and $B > 20$ and

$Max\{R, G, B\} - Min\{R, G, B\} > 15$ and

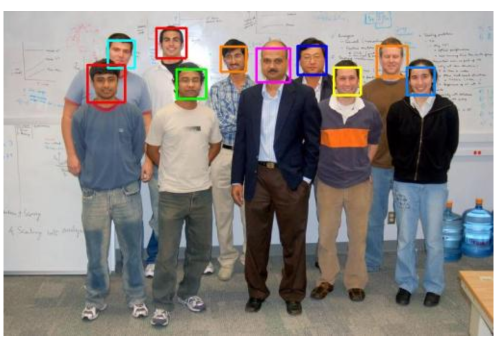
$|R-G| > 15$ and $R > G$ and $R > B$

In this approach, first skin regions in the input image are identified using above mentioned method and then Viola and Jones algorithm is applied for detecting faces. After applying skin color classifier, all non-skin regions replace with black, whereas skin regions remain stationary. This helps face detection algorithm to quickly identify non-faces which include majority pixels of each image. Also this method efficiently reduces false positive rate. Figure 5 shows the results of face detection before and after applying our hybrid algorithm.

(a) Original image



(b) Faces detected by V&J face detector



(c) Faces detected by our hybrid approach

Fig. 5 Results of face detection with V&J and our approach

In this approach, first skin regions in the input image are identified using above mentioned method and then Viola and Jones algorithm is applied for detecting faces. After applying skin color classifier, all non-skin regions replace with black, whereas skin regions remain stationary. This helps face detection algorithm to quickly identify non-faces which include majority pixels of each image. Also this method efficiently reduces false positive rate. Figure 5 shows the results of face detection before and after applying our hybrid algorithm.

5. Experimental Results

In this section a detailed experimental comparison of the above stated approach has been presented. We have used 100 color images with complex background obtained from internet and applied Viola and Jones based face detection and our hybrid approach on these images. Then we compare the results from two aspects, that is accuracy and face detection speed. The accuracy is obtained by using the following equation:

% Accuracy = 100 – (False positive Rate + False negative Rate)

The results of applying Viola and Jones face detection system and the proposed method are shown in Table 1.

Table 1: V&J face detector results

Criteria	Viola & Jones	Proposed method
False positive rate	13.45 %	5.44 %
False negative rate	15.61 %	17.42 %
Accuracy	70.94 %	77.14 %
Face detection time (Avg)	270.7ms	162.21ms

Clearly proposed method improve viola and Jones face detector in terms of both accuracy and speed. However, false negative rate in our method is a little bit increased. In figure 6 these two methods have been compared in terms of False Positive Rate, False Negative Rate, and Accuracy.

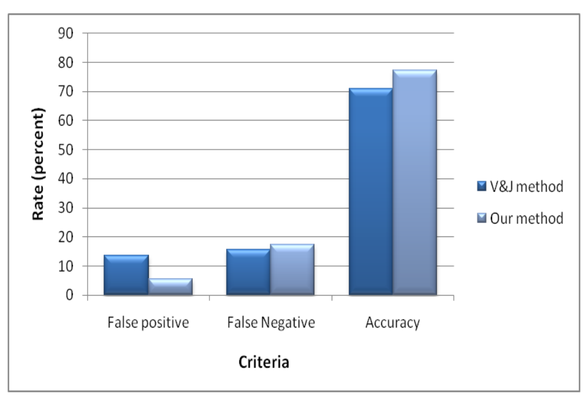


Fig. 6 Comparison of two Methods

Figure 7 demonstrates detection time among 100 images in comparison of our method and Viola and Jones Method.

Fig. 7 Comparison of detection time between two methods

6. Conclusion and Future work

In this paper we briefly described Viola and Jones face detector. Skin color classification has also been discussed and the combination of these two methods is proposed as a hybrid face detection system. The results show that our proposed method efficiently increases face detection speed as well as decreases false positive rate.

For the future work, in order to increase efficiency of our method other colorspaces and methods for skin color classification will be studied and examined.

References

[1] Curran, K., X. Li, et al. (2005). "Neural network face detection." Imaging Science Journal, the **53**(2): 105-115.

[2] Jiang, J. L. and K. F. Loe (2003). S-AdaBoost and pattern detection in complex environment, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops.

[3] Kovac, J., Peer, P. and Solina, F. (2003). "Human skin color clustering for face detection." *EUROCON 2003. Computer as a Tool. The IEEE Region 8* 2.

[4] Li, S. Z. and Z. Q. Zhang (2004). "FloatBoost Learning and Statistical Face Detection." IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE: 26 (9) 1112-1123.

[5] Liu, C. and Wechsler, H. (2002). "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition." *Image Processing, IEEE Transactions on* 11: 467-476.

[6] Osuna, E., Freund, R. and Girosit, F. (1997). "Training support vector machines: an application to face detection." 130-136.

[7] Rowley, H. A., Baluja, S. and Kanade, T. (1998). "Neural network-based face detection." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 20: 23-38.

[8] Schneiderman, H. and T. Kanade (1998). Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp 45-51.

[9] Schneiderman, H. and T. Kanade (2000). A Statistical Approach to 3D Object Detection Applied to Faces and Cars, Ph.D. Dissertation, Carnegie Mellon University, the Robotics Institute.

[10] Shih, P. and C. Liu (2004). Face Detection Using Discriminating Feature Analysis and Support Vector Machine in Video, in: Proc. 17th Intl. Conf. Pattern Recognition, IEEE Computer Society Washington, DC, USA.

[11] Singh, S. K., Chauhan, D. S., Vatsa, M. and Singh, R. (2003). "A Robust Skin Color Based Face Detection Algorithm." *TAMKANG JOURNAL OF SCIENCE AND ENGINEERING* 6: 227-234.

[12] Vezhnevets, V., Sazonov, V. and Andreeva, A. (2003). "A Survey on Pixel-Based Skin Color Detection Techniques."

[13] Viola, P. and M. J. Jones (2004). "Robust Real-Time Face Detection." International Journal of Computer Vision **57**(2): 137-154.

[14] Viola, P. and M. Jones (2001). Rapid Object Detection Using a Boosted Cascade of Simple Features, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition.