Face Detection using Fusion of Skin Detection and Viola-Jones Face Detection

Hla Myat Maw, K Zin Lin, Myat Thida Mon
University of Information Technology
Yangon, Myanmar
hmyatmaw@uit.edu.mm, kzinlin@uit.edu.mm, myattmon@uit.edu.mm

Abstract

Face Detection is the first step of any automated face and facial expression recognition system. It is applied in many applications such as computer vision and biometric applications like face recognition, expression recognition and classification, security access systems, video surveillance, and intelligent human computer interaction. A false detection rate and slow detection speed remain major problems in face detection. In this paper, we present a method to avoid these problems, by combining skin detection and Viola and Jones Face detection. Skin detection is used to reject false positive. After determining all possible face candidate regions, the Viola Jones detector is applied to detect face. Our method has high detection rate and accuracy than an individual face detector. It can perform quiet well in complex background and varying lighting conditions when it is applied images with complex backgrounds. Experimental results show the effectiveness of the proposed method.

Keywords-Face detection, Skin detection, Viola and Jones face detector, Color space model.

1. Introduction

Face detection is an essential step and usually in various computer vision and biometric applications such as face recognition, criminal investigation, and security access system, video surveillance, and human computer-interaction. Numerous researches were performed in the field of face detection and generally can be classified into four categories [1]: feature invariant approaches, template matching, knowledge based and appearance based methods. In the following, a brief review of these four categories is given. (1) Knowledge based methods are rule-based methods, which encode human knowledge about what a face is. For example, in our human mind, symmetric eyes, ears, nose and mouth are key face feature. Developing these methods in different situations is sometimes difficult because not all states are countable (2) Feature invariant approaches are grouping methods with aim to find robust structural features which are invariant to pose, lighting, etc. This method is one of the most important methods for face detection and features use low-level features such as intensity, color, edge, shape, and texture to locate facial features, and further, find out the face location. (3) Template matching methods compute the correlation between patterns of a face and an input image in order to detection. In these methods, the correlation of several patterns of face in different poses and the input images are stored to be a criterion for face validation. It is scale-dependent, rotation-dependent and computational expensive. (4)Appearance-based methods use models learned from training sets to represent the variability of facial appearance. Actually in these methods the templates are learned from face image samples. Generally, appearance-based methods utilize statistical analysis and machine learning to find characteristics related to face or non-face images.

Among feature-based face detection methods, skin color has been used and proven to be an effective feature to increase detection rate [3]. It is often considered to be a useful and discriminating image feature for facial area and usually employed as a preliminary step in face detection. It provides computationally effective yet robust to variation in scale, orientation and partial occlusion. Skin detection is also a challenging task since the skin color is sensitive to various factors including illumination, ethnicity, individual characteristics, subject appearances and camera characteristics. [4].

There are three primary steps for color-based skin detection in an image: representing the image pixels in a suitable color space, modeling the skin and non-skin pixels using an appropriate distribution and classifying the modeled distributions. A skin detection system utilizing skin color as a feature and HSV, YCbCr, YCgCr and YDbDr are the most appropriate color spaces for skin color detection.

The rest of this paper is organized as follows: Section 2 gives an overview of the related work. Section 3 presents background theory. The proposed system is presented in Section 4. Section5presents experimental result. Section 6 closes with a conclusion and future work.

2. Related work

A literature of some research work is presented to identify the best color model for a specific task. Fatma et al. [5] proposed technique is based on finding the maximum energy of histogram signal for skin which is limited to the ranges for each component of the color space under study. Different parameters such as energy of the histogram of each component of the color space, the limit of skin range in each color space and the maximum energy of the color spaces are used to evaluate. The result indicates that YCbCr provide better performance compared to RGB, YUV, HSV and CMYK color model. A detection rate of 97.51% was obtained using (Psychological Image Collection at Stirling) PICS database.

In [6] presents a human face detection scheme by combining a novel hybrid color models and Viola-Jones face detector. A hybrid skin color model RGB-CbCrCg was proposed for classifying skin and non-skin pixels. Afterward the segmented face regions are identified using Viola-Jones algorithm. A detection rate of 83.91% was obtained using (Edith Cowan University) ECU database.

In [7] describes a machine learning approach for visual object detection which is capable of processing images extremely rapidly and achieving high detection rates. Viola and Jones face detector has become the standard to build successful face detection in real time, however, it produces a high false positive (detecting a face when there is none) and false negative rate (not detecting a face that's present) when directly applied to the input image. To deal with this problem, a various improvements have been proposed, such as using skin color filters (whether pre- filtering or post-filtering) to provide complementary information in color images. Though many experimental results have demonstrated the feasibility of combining skin color detection with Viola-Jones face detector to reduce false positive, both methods suffer from high false negative rate as some face regions may be ignored by detector. A detection rate of 76.1% was obtained using (Edith Cowan University) MIT + CMU database.

3. Background theory

In this section, various techniques which have been used in the proposed algorithm will be explained.

3.1. Haar-like features

Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real time face detector. Historically, working with only image intensities (i.e., the RGB pixel values at each pixel of image) lead to expensive computation in feature calculation. The idea of Haar-like feature is to consider the adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and compute the difference between these sums. This difference is then used as a feature response to categorize subsections of an image.

Figure 1 shows three kinds of rectangle Haar-like features. The value of a two-rectangle feature is the difference between the sums of the pixels within two rectangular regions. The regions have the same size and shapes are horizontally or vertically adjacent. A three rectangle feature computes the sum within two outside rectangles subtracted from the sum in a center rectangle. Finally a four-rectangle feature computes the difference between diagonal pairs of rectangles.

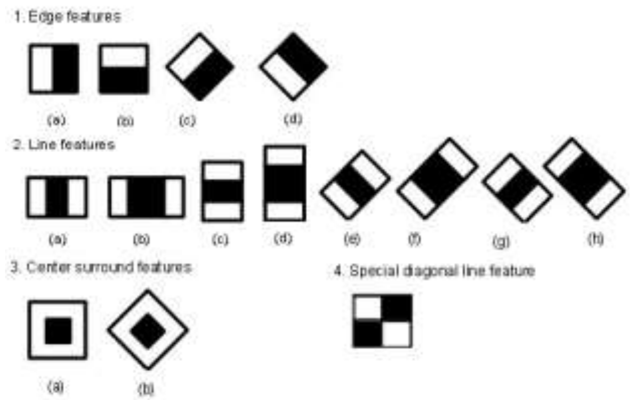

Figure1. Example rectangle features. The sums of the pixels which lie within the white rectangles are subtracted from the sum of pixels in the grey rectangles. Two-rectangle features are shown in (1), three-rectangle features are shown in (2), center surround feature are shown in (3) and a four rectangle features is shown in (4).

3.2. Integral image

Rectangle features can be computed very rapidly using the integral image [8]. The integral image at location x, y contains the sum of the pixels above and to the left of x, y Inclusive:

$$ii(x,y) = \sum_{x' \le x, y' \le y} i(x', y') \quad (1)$$

Where $ii(x,y)$ is the integral image and $i(x,y)$ is theoriginal image.Using the following pair of recurrences:

$$s(x,y) = s(x,y-1) + i(x,y) \quad (2)$$

$$ii(x,y) = ii(x-1,y) + s(x,y) \quad (3)$$

Where s(x, y) is the cumulative row sum, s(x,−1) = 0, and ii (−1, y) = 0, the integral image can be computed in one pass over the original image.

Using the integral image any rectangular sum can be computed in four array references (see Figure 2). We can easily see that the difference between two rectangular sums can be computed in eight references. Since the two-rectangle features defined above involve adjacent rectangular sums they can be computed in six array references, eight in the case of the three-rectangle features, and nine for four-rectangle features.

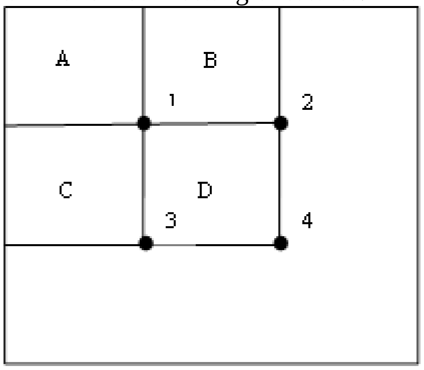

Figure2. The sum of the pixels within rectangle D can be computed with four array references. The value of the integral image at location 1 is the sum of the pixels in rectangle A. The value at location 2 is A+B, at location 3 is A+C, and at location 4 is A+B+C+D. The sum within D can be computed as 4+1-(2+3).

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.

In continue the Adaboost classifier was used for feature classifying. This classifier has two functions. At the first, the appropriate feature has been selected for every template then these features will be learned. Generally, the purpose of the Adaboost algorithm is improving and increasing the efficiency of classified simple algorithms. The importance of the Adaboost algorithm is selecting a few numbers of these features according to the template structure. During the training process, these features are extracted for different training templates. Then those features with no change for all training data are selected. Some of the selected features, which create the most distinction between the two classes, are extracted as final features.

3.3. Skin color model

A color space is a geometrical representation of colors in a space and allows specification of colors through three components whose arithmetical values define an exact color. The native and most available representation of color images is the RGB color format describing the world view in three color matrices. Any other color space can be derived from a linear or nonlinear transformation from the RGB channels. Moreover, different skin color spaces may be combined to get improved. The existing color-space combination methods can be categorized as mixture of collections. Color space selection is the primary process in skin color modeling and further for classification. One or more color spaces can give an optimal threshold value for detection of pixels of skin in a given image. The choice or components are hybrid/fusion of color spaces.

Figure 3 shows YCbCr (Luminance, Chrominance) Color Model. YCbCr is an encoded non-linear RGB signal, commonly used by European television studios and for image compression work. As shown in fig. 3 color is represented by luma (which is luminance computed from non linear RGB) constructed as a weighted sum of RGB values [4]. YCbCr is a commonly used color space in digital video domain. Because the representation makes it easy to get rid of some redundant color information, it is used in image and video compression standards like JPEG, MPEG1, MPEG2 and MPEG4.Thetransformationsimplicity and explicit separation of luminance and chrominance components makes YCbCr color space [3]. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference between the red component and a reference value. YCbCr values can be obtained from RGB color space according to eq. 4 to eq. 6.uses YCbCr space for skin detection.

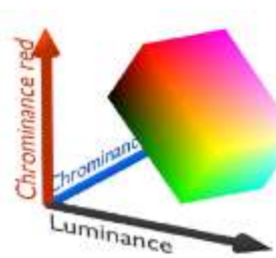

Figure 3. YCbCr Color model [12]

$$Y=0.299R+0.287G+0.11B \quad (4)$$

$$Cr=R-Y \quad (5)$$

$$Cb=B-Y \quad (6)$$

3.4. Skin color detection

Skin detection is the process of finding skin-colored pixels and regions in an image or a video. This process is typically used as a preprocessing step to find regions that potentially have human faces and limbs in images. Skin detection techniques can be broadly classified as pixel-based techniques or region-based techniques. In the pixel-based skin detection, each pixel is classified as either skin or non-skin pixel individually depending on certain conditions. The skin detection based on color values is pixel-based. In region-based skin detection technique, spatial relationship of pixels is considered to define some area from given image as skin region.

The most important parameter which is used to detect skin pixels in the image is color. According to the importance of color in skin detection, the used color space is also important. HSV, YCbCr, HIS, HUV, and YIQ color spaces are persistent against light and luminance intensity changes. Eliminating the y component in these spaces reduces input space in addition to reduce dependency to light intensity. In color segmentation scheme needs a proper representation of color spaces to interpret image information in many cases. In this paper, we use YCrCb color space for detecting skin in color images.

4. Proposed system

The proposed system is fused skin detection based on YCbCr color model and face detection based on Viola and Jones algorithm for face detection from color images. Figure 4 shows the block diagram of the proposed system.

4.1. Skin detection

Skin color methodologies have been widely used in many applications. Skin color is an affective key for face detection since it provides computationally effective, robust information against rotation, scaling and partial occlusions. In general, the final goal of skin color detection is to build a decision rule, which will discriminate between skin and non-skin pixels. Skin color detection may fall into two main categories [9]: pixel-based skin detection methods and region-based skin detection methods. Pixel-based skin detection methods classify each pixel as skin or non-skin individually, independently from its neighbors. On the other hand, region- based skin detection methods try to take the spatial arrangement of skin pixels into account during the detection stage to enhance the methods performance.

In our proposed method, especially for real-time implementation, the pixel-based skin detection method is chosen for fast processing. We also use a color compensation step prior to skin detection, to reduce the effects of lighting. A pixel with color components (R, G, B) is detected as skin if the conditions given in (7) below hold. The second line in (7) ensures that RGB components must not be close together, which ensures grayness elimination. The third line in (7) ensures that R and G components must not be close together, which must be true for fair complexion.

(a) $R > 95 \ \& \ G > 40 \ \& \ B > 20$ and (7)
(b) $\max\{R,G,B\} - \min\{R,G,B\} > 15$ and
(c) $|R - G| > 15 \ \& \ R > G \ \& \ R > B$

For YCbCr color system, a pixel (Y, Cb, Cr) is classified as skin if:

(a) $60<Y<255$ and (8)
(b) $100<Cb<125$ and
(c) $135<Cr<170$

4.2. Viola and Jones face detection

To classify the areas into faces or non faces, the Viola and Jones face detector is performed locally on all selected bounding boxes around connected pixel regions on an image. This allows help to decrease the false positives in face detection. Viola & Jones [7] have presented a face detection method based on an over-complete set of Haar-like features in (Figure.1) 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.

Thus, if the image passes through all the stages then it is a face. If it fails even in any one of the stages then it is not a face. This gives an overall description of the algorithm.

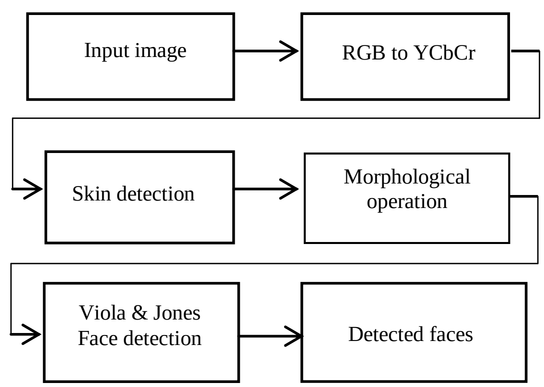

Figure 4. Block diagram of the proposed system

The steps of implemented algorithm are:

Step1: Read the image, and capture the dimensions.
Figure 5 shows the original image.

Figure5. Original Image

Step2: Lighting compensation

Lighting compensation is used for reducing the effects of lighting. Figure 6 shows the results of lighting compensation.

Figure6. Lighting Compensated Result

Step3: Skin region extraction

Applying the color segmentation to the image shown in Figure 6 using inequalities and the result is shown in Figure 7. The output image is a binary image that contains ones in the skin regions and zeros in non-skin regions. The detected skin regions may be discontinuous; this discontinuity may be due to lighting effects that leads to missed skin pixels or due to the presence of non-skin face features like the eyes and brows.

Figure7. Extracted Skin Result

Step4: Noise removing

In this step, morphological operation is applied to reduce false positives. Figure 8 shows the output after noise removing.

Figure8. De-noised Skin Result

Step5: Find skin color block

Figure 9 shows the result of skin color block.

Figure9. Skin Color Block Result

Step6:.Face Detection

The Viola and Jones face detector is used to detect face from skin region. The algorithm is to scan a sub window capable of detecting faces across the skin region using features consisting of two or more rectangles. The result is shown in figure 10.

Figure10. Detected Face(s) Result

5. Experimental results

In order to evaluate the performance of the proposed method, many experiments have been carried out using a total of 30 images of varying lighting conditions and complex background.

We also compare our proposed method with Viola-Jones face detector as an individual method. The comparison is based on the accuracy, the false positives rate, and the false negatives rate, which are defined below Table1.

The results presented in Table 1 were obtained using a variety of face dataset. The analysis of these results show clearly, that the accuracy of the Viola & Jones detector increases when a skin color detection is used as a prior stages which means that this face detector gives low false negatives rate when it is not applied directly to the entire input image. The proposed system improves significantly the accuracy by overcoming the problem of false detection. The faces of those people wearing caps, glasses are also detected.

Table. 1. Comparisons of our results with other face detector method

	Accuracy	FNR	FPR
Viola & Jones [07]	79.46%	20.54%	34%
Proposed system	86.55%	0.89%	0.15%

- False Positives Rate (FPR) = the ratio of the number of detected false positives to the total number of faces.
- False Negatives Rate (FNR) = the ratio of the number of false negatives to the total number of faces.
- Accuracy = 1-(FPR+FNR)/2

6. Conclusion and future work

In this paper, we proposed system using a fusion of skin detection and Viola and Jones face detector to improve the accuracy of face detection in varying illumination and complex background. Skin detection is used before face detection to reduce false negative rate and overcome variation in pose and illumination. The experimental results presented illustrate the effectiveness of this method, compared to some other method proposed in the literature, especially the well- known Viola &Jones face detector. The proposed method works for non-frontal face and frontal face and different lighting condition. But it takes times for complex background.

In future, more research work should continually focus on human face detection for people of different races and computation time should be further saved for real time applications.

7. References

[1] M.-H. Yang, D. J. Kriegman, and N. Ahuja, ‟Detecting Faces in Images: A Survey”, IEEE Transaction on Pattern Analysis and Machine Learning, vol. 24, no.1, 2002, pp. 34-58.

[2] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, “Face Detection in Color Images”, IEEE Transaction on Pattern Analysis and Machine Learning, vol. 24, no.5, 2002, pp. 696-706.

[3] 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”, International Journal of Computer Science and Network Security, vol. 9, no.5, 2009, pp. 181-185.

[4] Y. Wang, B. Yuan, “A Novel Approach for Human Face Detection from Color Images under Complex Background”, Pattern Recognition, vol. 34, no.10, 2001, pp. 983-1992.

[5] FatmaSusilawatiMohamad, Abdulganiyu Abdu Yusuf, ZahraddeenSufyanu, “Evaluation of Suitable Color Model for Human Face Detection”, International Conference on Electrical Engineering and Electronics Communication System (ICEEECS), 2015.

[6] Guan-Chunluh, “Face Detection Using Combination of Skin Color Pixel Detection and Viola-Jones Face Detector”, International Conference on Machine Learning and Cybernetics, Lanzhou, 13-16 July, IEEE , 2014.

[7] Viola Paul, and Michael Jones. "Rapid Object Detection using A Boosted Cascade of Simple Features." Proceedings of the 2001 IEEE Computer Society Conference on. Computer Vision and Pattern Recognition, CVPR 2001.

[8] P. Viola and M. J. Jones, “Robust Real Time Face Detection”, International Journal of Computer Vision, vol. 57, no. 2, 2004,pp. 137-154.

[9] Deepak Ghimire and Joonwhoan Lee, “A Robust Face Detection Method Based on Skin Color and Edges”, J. Inf. Process Syst., vol.9, no.1, March 2013.

[10] 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.

[11] Amit Pal, “Multicues Face Detection in Complex Background for Frontal Faces”, International Machine Vision and Image Processing Conference, IEEE 2008.

[12] https://en.wikipedia.org/wiki/Talk:YCbCr