mekatronika – Journal of Intelligent Manufacturing & Mechatronics
VOL. 01, ISSUE 01, 58 – 63
DOI: https://doi.org/10.15282/mekatronika.v1i1.375

ORIGINAL ARTICLE

Automated Face Detection Using Skin Colour Segmentation and Viola–Jones Algorithm

Ahmad Fakhri Ab. Nasir*, Ahmad Shahrizan Abdul Ghani, Muhammad Aizzat Zakaria, Anwar P P Abdul Majeed and Ahmad Najmuddin Ibrahim

Innovative Manufacturing, Mechatronics and Sports Laboratory (iMAMS), Faculty of Manufacturing Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

ABSTRACT – Viola–Jones algorithm is an effective feature-based approach for face detection. This algorithm consists of three main processes, namely, Haar feature, AdaBoost and cascading. These processes require scanning the patterns of a human face to obtain all pixels in the image. An empirical experiment was conducted using this algorithm, wherein nonface regions in the image are detected as face given its similarity with human face features. In addition, this approach only searches for patterns (features) but ignores colour information. Therefore, this research proposes a hybridisation of skin colour segmentation prior to using Viola–Jones algorithm. Several images with different types of environments (light conditions and face orientations) and various ethnicities (Malay, Chinese and Indian) are tested using the improved algorithm and the original Viola–Jones algorithm. Experimental results reveal that combining YCbCr colour and Viola–Jones algorithm is the optimal model (average accuracy is approximately 88%) for human face detection under various conditions. Other colour models, such as hue saturation value (HSV) and standard algorithms, have slightly low detection rates given certain false face detection. Despite this improvement, this research has also determined that the YCbCr and HSV colour models demonstrate certain limitations when handling dark faces and lighting condition because the skin colour becomes slightly out of range from the normal skin colour distribution.

ARTICLE HISTORY
Received: 10 October 2018
Revised: 2 January 2019
Accepted: 9 January 2019

KEYWORDS
Face detection
image processing
Viola–Jones algorithm
skin colour segmentation

Introduction

Face recognition has become a popular research area covering many fields and disciplines. The application of face recognition includes monitoring and surveillance systems and access control, such as smart door entrance identification and attendance system, to predict human behaviour. These applications help improve the daily lives and routines of humans. In general, the three steps in an automated face recognition system are detection, feature extraction and face recognition [1-2] (Figure 1). Figure 1 illustrates that the first step in this system is face detection which is a process of detecting or locating face regions from the entire picture or video. Face detection involves many processes, including segmenting the background, scanning the patterns of a face and noise removal. After locating the face region, certain useful strong features are extracted to distinguish the faces of different people. Many studies are available on face features [3-5]. Finally, these extracted features are trained and tested through classification methods.

As the first step in the system, overcoming the problem of face detection is crucial. Detection must be sufficiently favourable for application to different scenarios that can lead to false face detection. For example, if the system requires finding the face of a student for attendance purposes, then the system will not detect (underdetection) or will overdetect (more detection) a face in the class when face detection algorithm is substandard. Figure 2 depicts this scenario wherein several faces in the front, middle and back positions of the class are undetected.

Methodology

Viola–Jones Algorithm

Numerous face detection approaches have been proposed in recent years. These approaches are divided into the following main categories of face detection: (i) local feature and (ii) global feature-based approaches. Viola–Jones algorithm (local feature), which was developed by Paul Viola and Michael Jones in 2000, was the first algorithm used for real-time face detection systems [7-9]. The three main components of this algorithm are (i) Haar feature, (ii) AdaBoost and (iii) cascading [9]. Haar feature is used to detect human features in images and is typically used as a 24×24 window base size to start evaluating every single spot in the image. Figure 3 demonstrates the four types of Haar feature for evaluation. These

*CORRESPONDING AUTHOR | Ahmad Fakhri Ab. Nasir | ✉ afakhri@ump.edu.my

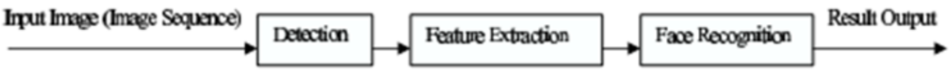

Figure 1. Basic block diagram for face recognition [1].

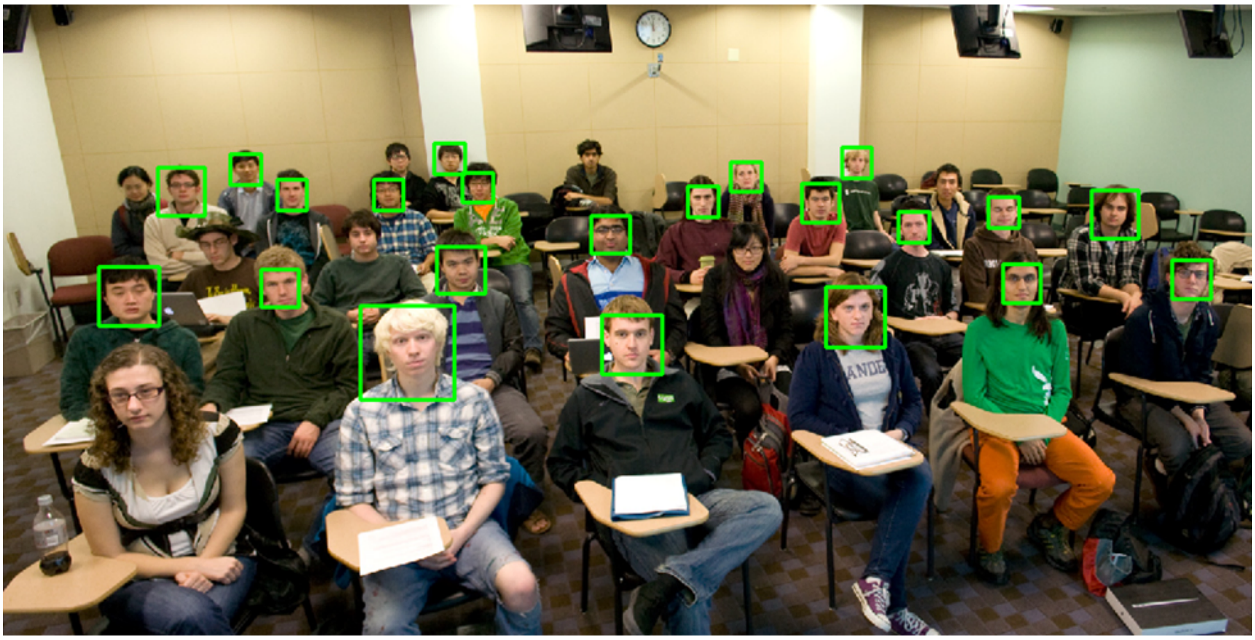

Figure 2. Sample of face detection in a class picture used in Reference [6].

features are edge, line, centre–surround and special diagonal line.

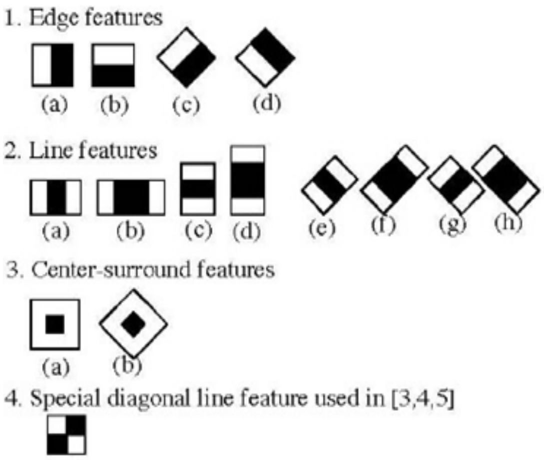

Figure 3. Different types of Haar feature [10].

Haar feature consists of black and white regions. The black and white regions are replaced by +1 and −1 values, respectively. Each feature results in a single value calculated by subtracting the sum of pixels under the white region minus the sum of pixels under the black region of each feature. To determine whether the feature is a human or nonhuman face, the single value obtained from feature calculation indicates human face similarity. Feature matching is performed on each spot in the image that results in a combination of 160,000 types of features. For example, Figure 4 exhibits that the line feature matches the shape of the human nose, thereby resulting in a high calculation value.

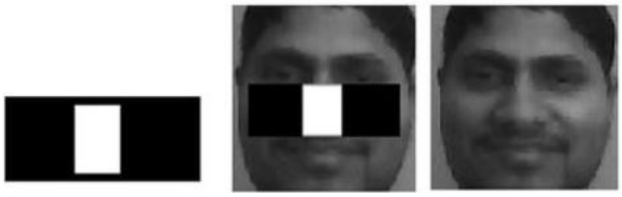

Figure 4. Example of line feature matching with the human nose.

AdaBoost removes and simplifies irrelevant human facial features; only few relevant features are useful among the 160,000 features generated by Haar feature [8]. Relevant features match with the human face. Thus, AdaBoost reduces the features from 160,000 to 7,000 relevant features. AdaBoost also determines whether a feature is a strong or weak classifier, wherein a weak classifier can be considered a relevant feature. The output of weak classifiers is a binary value (1 or 0), wherein an output is 1 when the classifier can identify features when applied to the image. Otherwise, the output is 0.

Lastly, cascading shortens the time for human or nonhuman feature identification for each feature. Cascading is composed of many stages, and each stage consists of a strong classifier. Each stage is grouped with a certain number of features. If the feature does


(a)


(b)

Figure 5. (a) Original and (b) YCbCr colour filtered images.

not pass Stage 1, then the feature is considered a nonhuman face, thus eliminating the time-consuming process of going through all the stages. After passing Stage 1, the feature may be considered a human face, which will go through subsequent stages until successfully identifying such a feature as a human face.

YCbCr Colour Model

YCbCr is a colour space composed of Y, Cb and Cr as the luma, blue-difference and red-difference chroma components, correspondingly. A luma component determines the brightness of an image. According to [13-14], blue and red space ranges are suitable for distinguishing human skin, such as Cb = [77,127] and Cr = [133,173]. The YCbCr colour space distinguishes colour well under different brightness levels. In this research, the skin colour range, which is not in Cb and Cr, is removed before using the Viola–Jones algorithm (Figure 5).

HSV Colour Model

The HSV colour model is a colour space that consists of the colour elements H (hue) and S (saturation), wherein both values vary from 0 to 1.0. HSV is a linear distribution of the colour spectrum that can distinguish very specific colours of the same kind. For this research, the colour ranges that represent skin colour are set to H = [0.26,1.00], S = [0.00,0.15] and V = [0.91,1.00] .

Results and Discussions

In this research, MATLAB software was used to code the algorithms. Three algorithms, namely, [Approach (i) – standard model using Viola–Jones], [Approach (ii) – YCbCr + Viola–Jones] and [Approach (iii) – HSV + Viola–Jones], were coded. The initial step is to obtain images with different surrounding conditions and human ethnicities from the Internet. A total of 70 images with different characteristics were obtained from the Internet to test the efficiency of the abovementioned algorithms. Table 1 lists these images which were separated into groups with each group having its own properties.

Table 1. Image categories.

Type	Image descriptions	Image
i	Chinese ethnicity	1–10
ii	Malay ethnicity	11–20
iii	Indian ethnicity	21–30
iv	African ethnicity	31–40
v	American ethnicity	41–45
vi	Dim surrounding light condition	46–55
vii	Bright surrounding light condition	56–60
viii	Different human face orientation & distances	61–70

Figure 6 displays the sample for each type of image. The images were tested for face detection and calculated on the basis of the confusion matrix shown in Table 2.

Table 2. Confusion matrix table (contingency table).

		Predicted	
		Face	Nonface
Actual	Face	TN	FP
	Nonface	FN	TP

The performance measures for the confusion matrix are presented as follows [11-12]:

$$\text{Accuracy} = \frac{(TN + TP)}{(TN + FP + FN + TP)} \times 100 \tag{1}$$

Figure 7 presents that the experimental results are divided into different types of images and approaches. This figure illustrates that [Approach (ii)] is optimally used for face detection on all types of images, except Indian and African ethnicities and dim surrounding light condition. [Approach (iii)] seems the poorest algorithm for all the images because it obtains the lowest face detection accuracy results. On average,

Figure 6. Image sample for types (a) i, (b) ii, (c) iii, (d) iv, (e) v, (f) vi, (g) vii and (h) viii.

[Approach (ii)] is better than the two other approaches with an accuracy of 88.08%, followed by [Approach (i)] 84.17% and [Approach (iii)] 78.23%.

As mentioned previously, the YCbCr and HSV have limitations for Indian and African ethnicities and dim surrounding light condition considering the dark skin complexion and slightly out of the colour range preset from normal fair human skin complexion. For example, the YCbCr and HSV tend to remove most of the human skin region for people of African ethnicity, thus making the human face difficult for detection. Figure 8 depicts an African or dark-skinned human image that is unsuitable for the YCbCr and HSV colour segmentation, which will remarkably affect the result of the Viola–Jones algorithm detection process. The figure demonstrates that many faces are removed after the YCbCr and HSV colour segmentation, thereby indicating that detection cannot be performed through the Viola–Jones algorithm.

Conclusions

Given that most industries currently use applications that are controlled by intelligent systems, face detection technology is recommended for monitoring and surveillance systems and access control, such as smart door entrance identification and attendance system, to predict human behaviour. In this research, we identified that YCbCr + Viola–Jones algorithm obtains a highly accurate result and low false detection in most of the surrounding conditions and different ethnicities of people. Although using the Viola–Jones algorithm alone obtains a highly accurate detection, the surrounding conditions have different types of shapes and orientations, thereby causing several features to appear similar to human faces and generate false detection outputs. The HSV has minimal benefits for face detection because this type of colour segmentation is sensitive to identifying colours, thus causing it to remove more colour regions near the human face. The removal of certain parts of the human face will prevent the Viola–Jones algorithm from detecting human features, thus leading to undetected human faces.

Acknowledgements

The authors acknowledge the internal grants from Universiti Malaysia Pahang (RDU1703159 and RDU180383). The authors are also grateful to Mr. Loh Khai Shyang, an undergraduate student of the Faculty of Manufacturing Engineering, Universiti Malaysia Pahang, for the results.

References

[1] Y. Lu, J. Zhou, and S. Yu, "A survey of face detection, extraction and recognition," *Comput. Informatics*, vol. 22, no. 2, pp. 163–195, 2003.

[2] A. S. Tolba, A. H. El-Baz, and A. A. El-Harby, "Face recognition: A literature review," *Int. J. Signal Process.*, vol. 2, no. 2, pp. 88–103, 2006.

[3] X. Zhu and Deva Ramanan, "Face Detection, Pose Estimation, and Landmark Localization in the Wild," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2879–2886.

[4] R. Sznitman and J. Bruno, "Active Testing for Face Detection and Localization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32(10), pp. 1914–1920, 2010.

[5] S. Samy, A.-H. Ayoub, M. Bernd, and S. Usama, "Face Detection and Localization in Color Images: An Efficient Neural Approach," *J. Softw. Eng. Appl.*, vol. 4, pp. 682–687, 2011.

Table 3. Experimental data of different types of images.

Image Type / Approaches	[Approach (i)] (%)	[Approach (ii)] (%)	[Approach (iii)] (%)
Chinese	89.09	90.57	83.02
Malay	87.50	96.15	84.91
Indian	86.21	84.75	74.14
African	77.14	76.12	64.29
American	81.25	83.87	80.65
Dim surrounding light condition	83.33	80.00	77.78
Bright surrounding light condition	80.00	100.00	80.00
Different human face orientations and distances	88.82	93.20	81.08
Average	84.17	88.08	78.23

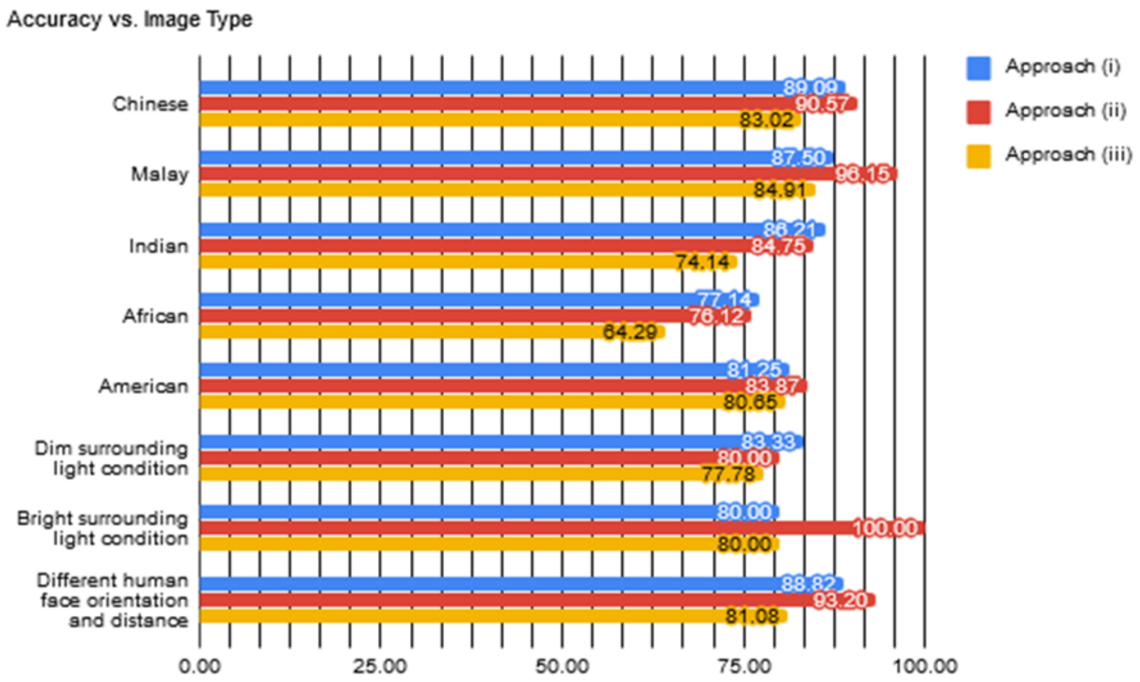

Figure 7. Experimental results of different types of images.

(a) (b)

Figure 8. Face detection using the (a) YCbCr and (b) HSV colour segmentation.

[6] "Building Live WebCam Face Detector using Node.js and OpenCV," 2018. [Online]. Available: https://codeforgeek.com/2017/07/building-live-webcam-face-detector-node-js-opencv/.

[7] D. Rudolfo Rizki, S. Delima, P. Hendra, S. Hendrik, and G. Frisman, "An application of viola jones method for face recognition for absence process efficiency," in *IOP Conf. Series: Journal of Physics: Conf. Series*, 2013.

[8] W. Yi-Qing, "An Analysis of the Viola-Jones Face Detection Algorithm," *Image Process. Line*, vol. 4, pp. 128–148, 2014.

[9] V. Paul and J. Michael, "Rapid Object Detection using a Boosted Cascade of Simple Features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, pp. 511–518.

[10] A. P. Mena, M. M. Bachiller, and E. Díaz-Lópe, "Comparative Study of the Features Used by Algorithms Based on Viola and Jones Face Detection Algorithm," *Bioinspired Comput. Artif. Syst. Lect. Notes Comput. Sci.*, vol. 9108, 2015.

[11] Z. Taha, R. M. Musa, A. P. P. Abdul Majeed, M. R. Abdullah, A. F. Ab. Nasir, and M. H. . Hassan, "Classification of High Performance Archers by Means of Bio-physiological Performance Variables via k-Nearest Neighbour Classification Model," in *Intelligent Manufacturing & Mechatronics. Lecture Notes in Mechanical Engineering*, 2018.

[12] A. N. A Fakhri, A. R. M Nordin, M. Nashriyah, M. A Rasid, and A. S. Abdul Ghani, "Ficusdeltoidea (Jack) Moraceae Varietal Identification Using Statistical Recognition Approach," *World Appl. Sci. J.*, vol. 35, pp. 82–88, 2017.