

# Dynamic approach for real-time skin detection

Sara Bilal · Rini Akmeliawati ·  
Momoh Jimoh E. Salami · Amir A. Shafie

Received: 23 April 2012 / Accepted: 24 November 2012  
© Springer-Verlag Berlin Heidelberg 2012

**Abstract** Human face and hand detection, recognition and tracking are important research areas for many computer interaction applications. Face and hand are considered as human skin blobs, which fall in a compact region of colour spaces. Limitations arise from the fact that human skin has common properties and can be defined in various colour spaces after applying colour normalization. The model therefore, has to accept a wide range of colours, making it more susceptible to noise. We have addressed this problem and propose that the skin colour could be defined separately for every person. This is expected to reduce the errors. To detect human skin colour pixels and to decrease the number of false alarms, a prior face or hand detection model has been developed using Haar-like and AdaBoost technique. To decrease the cost of computational time, a fast search algorithm for skin detection is proposed. The level of performance reached in terms of detection accuracy and processing time allows this approach to be an adequate choice for real-time skin blob tracking.

**Keywords** Skin detection · Colour space · Haar-like · Face detection · Hand detection

## 1 Introduction

Skin detection plays an important role in a wide range of image processing applications specifically in sign language recognition (SLR) and other human to computer interaction (HCI) domains [16]. There are various approaches for blobs detection and tracking based on the skin colour information in the visual spectrum. However, developing an accurate classifier that can detect all skin types under different illuminations, shadows and clustered backgrounds is still an unsolved problem. This paper presents a real-time dynamic skin detection approach for blobs detection and tracking. The proposed method reliability is defined by its robustness with different environmental lighting conditions.

The paper is organized as follows. Section 2 presents the research background. Section 3 describes the proposed method for blob detection. The experimental results are then presented in Sect. 4. In Sect. 5, we compare our approach with two existing adaptive and statistical techniques. Finally, Sect. 6 states the concluding part of the paper.

## 2 Research background

Diverse approaches for skin detection have been developed which could be classified into parametric [1, 4, 9, 12], statistical, such as [15] and adaptive methods [7, 11, 17] as described in [16, 20, 23]. The parametric and statistical methods are quite effective for skin detection, but they are highly sensitive to changes in lighting conditions and they also produce high false-positive errors in case of objects with colour similar to human skin exist in the environment. Most of these approaches assume that human skin colours of different individuals cluster in a small region in colour

---

International Conference of Mechatronics and Automation (ICMA).

---

S. Bilal · R. Akmeliawati (✉) · M. J. E. Salami · A. A. Shafie  
Department of Mechatronics Engineering, International Islamic  
University Malaysia (IIUM), JI Gombak,  
53100 Kuala Lumpur, Malaysia  
e-mail: rakmelia@iium.edu.my

S. Bilal  
e-mail: smosb@hotmail.com

space provided that the images are taken under illumination controlled environments [27].

The parametric skin models are based on a mathematical model to describe the human skin colour distribution. In the existing parametric skin models, the mathematical model is usually a Gaussian model [23]. The parametric models require a higher computing specification because they model a wide range of skin pixels' colour. Aspects of parametric models are defined in RGB, HSV,  $Y_CbCr$  colour spaces; other approaches [18] have combined the RGB and  $Y_CbCr$  colour spaces for increasing the skin colour detection accuracy. The statistical methods [14, 25, 26, 30] have shown advantages over the parametric models. As in the parametric models, the statistical models have used RGB,  $IC_rCb$ , HSV, HIS, HS and IUV colour spaces. In those methods, the skin pixels are segmented manually and used as a skin colour probability density function and find the threshold value to segment the skin area.

The survey paper [16] has shown that the adaptive skin detection methods, which dynamically adapt themselves to the changes of colour skin features due to the environmental noise or the change of light or object similar to the skin colour, demonstrate a better performance than other statistically based approaches. Many of the adaptive skin detectors are mainly employed to achieve high detection accuracy. Neural network-based methods support vector machines and other kernel methods have been proposed [16, 20, 23], where in the vast majority of cases, lighting conditions are constant within an image.

However, the problem with these approaches is that when we lose track of the skin, the adaptive model might falsely adapt to non-skin pixels regions. Furthermore, the work in Ref. [17] proposed an adaptive skin model, which has been created based on a detected face in luminance channel and geometric properties followed by a statistical skin modeling approach. Hence, to build flexible systems that work in a variety of lighting conditions which are robust and efficient for skin detection algorithms, are still an open research area. Sara et al. [2] have suggested a hybrid method by combining skin colour segmentation with Haar-like and Adaboost algorithm [24] for human hand posture detection.

The work presented in [2] contains preliminary results for skin detection and blob tracking, which used only  $Y_CbCr$  and open palm as the initial clue. In this paper, face or open hand palm has been used for initiating the developed method. The skin colour choice between RGB, HSV and  $Y_CbCr$  colour spaces has been experimented to show the best skin colour choice that fit this approach. In addition, extra analysis on various skin colours of different people using either face or hand as a first clue to define the skin areas is presented. This proposed method has been benchmarked using 40 videos; 35 videos collected from over the Internet and 5 videos have been obtained under

different environments. The approach has been developed for real-time implementation with 30 fps, which requires few milliseconds (ms) to find either the hand or the face of a visible person in the video frames. In addition, it requires a visible person standing about 1 m far from the camera. A comparison between our skin detection method and other skin colour detection methods [6, 13] is presented in this paper. The next section explains the proposed technique.

### 3 Proposed method

The idea of the proposed method comes from the assumption that illumination conditions are not changeable within an image frame and the skin colour does not vary much for a single person. Therefore, if skin colour could be found for everyone in the image frame; these will reduce the false positives of skin regions. This can be achieved if it is known that a particular region of the image contains skin pixels. Based on that region, a colour space can be defined or conformed dynamically and applied to the whole image. Figure 1 shows the framework of the proposed method in which a face or/and hand is detected in an image.

Following that a region that contains only skin pixels with high probability is highlighted. This will reduce false positives within the consecutive frames. The proposed method consists of the following phases: face or hand has been detected in the first frames of the video. Then, a seed point was extracted from a center of the detected blob and number of pixels was selected to find skin pixels.

#### 3.1 Face or hand detection with Haar-like features and AdaBoost algorithm

Initializing the proposed method with a face or specific hand shape will provide an adequate search region for skin pixels. Neural network-based methods, support vector machines and other kernel methods have been proposed for face or hand posture detection. However, most of these algorithms use raw pixel values as features which make their approaches sensitive to noise and changes in illumination. Other methods using region-based approaches, such as Haar-like features, which are similar to Haar basis functions has been proposed by [24]. The features encode differences in average intensities between two rectangular regions, and they can extract textures independent from absolute intensities. The proposed method uses Haar-like features and AdaBoost algorithm for face or hand detection.

##### 3.1.1 Haar-like feature family

The simple Haar-like features described in Fig. 2 (so-called because they are computed similarly to the coefficients in

Fig. 1 Overall system view

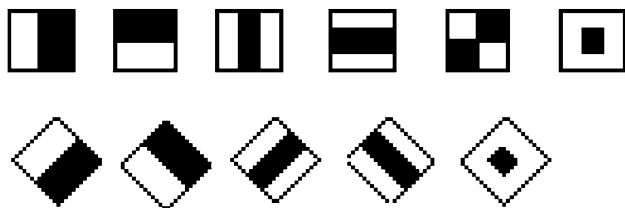
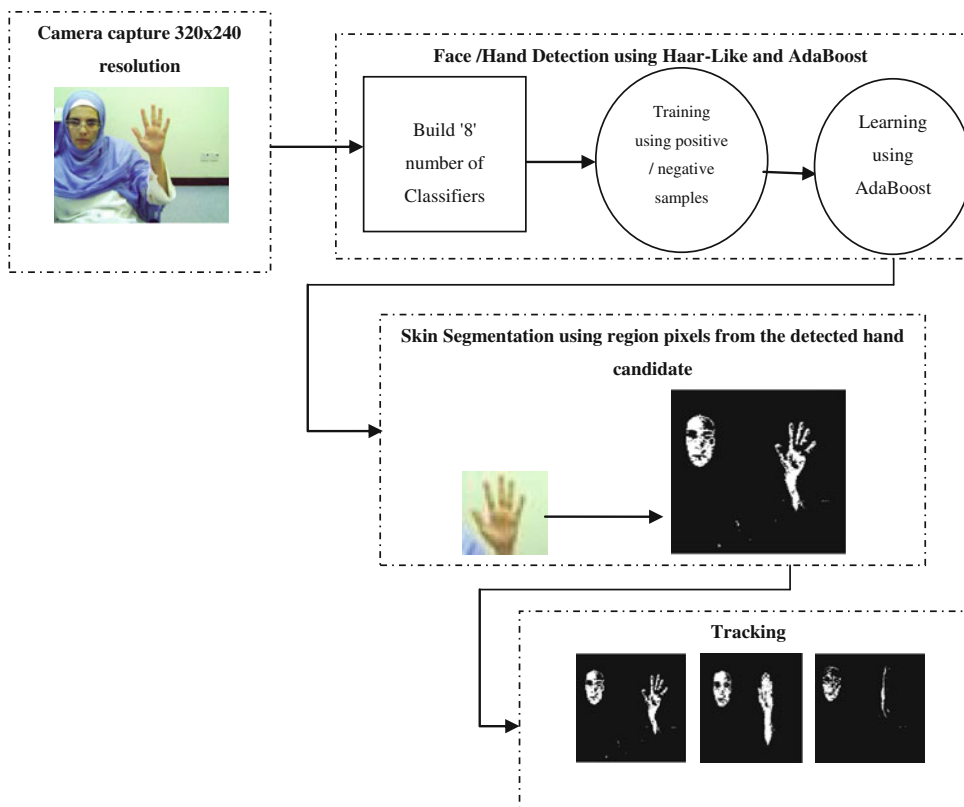


Fig. 2 A set of Haar-like features

the Haar wavelet transform) are used by [19, 24]. There are two motivations for the employment of the Haar-like features rather than raw pixel values. The first motivation is that the Haar-like features can encode ad hoc domain knowledge, which is difficult to describe using a finite quantity of training data [5, 10]. When compared with raw pixels, the Haar-like features can efficiently reduce (increase) the in-class (out-of class) variability and thus making classification easier [10]. Thus, Haar-like feature-based system can operate faster than a pixel-based system; especially, when using the integral image concept. In addition to the above advantages, the Haar-like features are also relatively robust to noise and lighting changes because they compute the grey level difference between the white and black rectangles.

3.1.1.1 Integral image Using the integral image, any rectangular sum can be computed in four array references

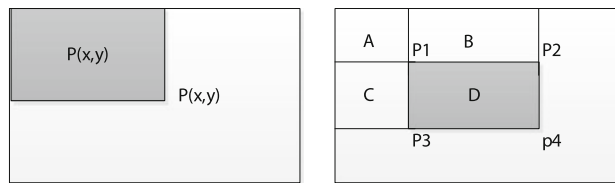


Fig. 3 The integral image

which can reduce the image processing time [12]. The integral image at location  $x, y$  contains the sum of the pixels above and to the left of  $x, y$ . Using the following pair of recurrences:

$$s(x, y) = s(x, y - 1) + i(x, y) \tag{1}$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \tag{2}$$

where  $ii(x, y)$  is the integral image and  $i(x, y)$  is the original image,  $s(x, -1) = 0$ , and  $ii(-1, y) = 0$ .

The integral image can be computed in one pass over the original image. The sum of the pixels within rectangle **D** (see Fig. 3) can be computed with four array references. The value of the integral image at location **P1** is the sum of the pixels in rectangle **A**. The value at location **P2** is  $A + B$ , at location **P3** is  $A + C$  and at location **P4** is  $A + B + C + D$ . The sum within **D** is computed as  $P4 + P1 - (P2 + P3)$ .

The above technique is used in the proposed method combined with AdaBoost algorithm for learning to detect the face or hand region which has reduced the image processing time.

### 3.1.2 AdaBoost learning algorithm

AdaBoost learning algorithm is a method aimed at improving the accuracy based on a series of weak classifiers stage-by-stage [10]. Initially, it maintains a uniform distribution of weights over each training samples. In the first iteration, the algorithm trains a weak classifier using one Haar-like feature that achieves the best recognition performance for the training samples. In the second iteration, the training samples, which were misclassified by the first weak classifier, receive higher weights so that the new one is selected. The iteration goes on and the final result is a cascade of linear combinations of the elected weak classifiers, i.e. a strong classifier, which achieves the

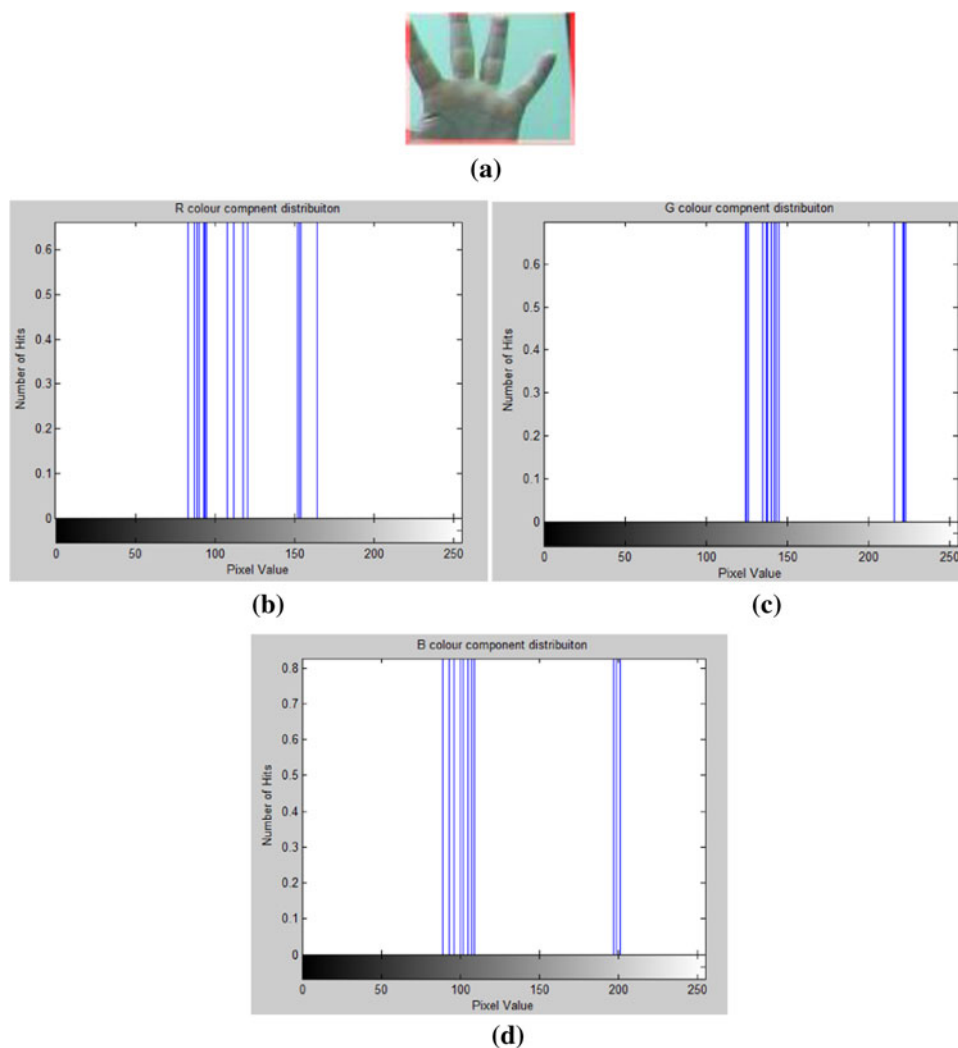
required accuracy. In the first stage of the training process, the threshold of the weak classifier is adjusted so that 100 % of the target objects can be detected, while keeping the false-negative rate close to zero. Detailed implementation of Haar-like and adaptive AdaBoost can be found in [24]. After the face or hand region has been detected in the image, skin colour regions was found.

### 3.2 Colour space choice

To segment human skin regions from non-skin regions based on the colour, different approaches have found that human skin colours from different ethnicities fall in a compact region in colour spaces [28, 29]. In related work, [3, 19, 22] have shown that the best results for skin colour detection can be achieved using RGB, HSV, and  $YCbCr$  colour spaces. Their assumption was due to the fact that using Hue factor with one dimension colour space for skin segmentation will be the fastest methods for implementing

**Fig. 4** RGB colour distribution within the skin detected area.

**a** Skin search area, **b** R component, **c** G component and **d** B component



a skin detector. Obviously, a large number of dimensions in colour space require a larger amount of training data and memory space. In addition, larger number of dimensions requires more thresholds to specify the object region [7]. In our approach, the method is initiated by detecting the face or hand region; the dimension of the colour space are then obtained using the range of the skin pixels within the detected area. This technique would reduce the memory space and processing time.

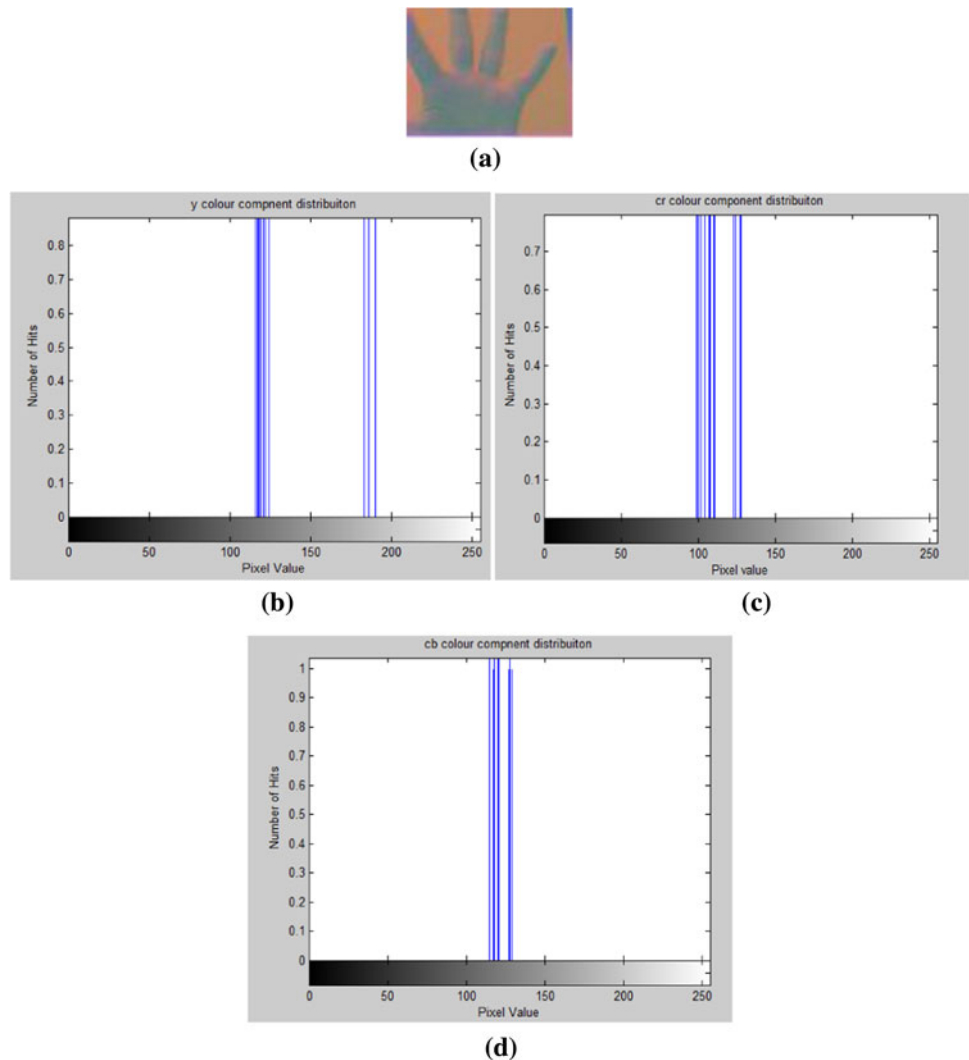
The detected skin area is analysed in the image sequence using RGB and  $YC_bC_r$  to choose the appropriate colour space for the proposed method. A small  $10 \times 10$  skin's region from the detected skin area is extracted and the histograms for each colour component from colour spaces RGB and  $YC_bC_r$  (see Figs. 4, 5) are analysed. Figure 4 shows that the RGB colour space is widely distributed. Figure 5 shows that the components in  $YC_bC_r$  colour space are narrowly distributed which exposes the  $YC_bC_r$  colour space as our best colour space choice for this proposed method.

### 3.3 Skin segmentation

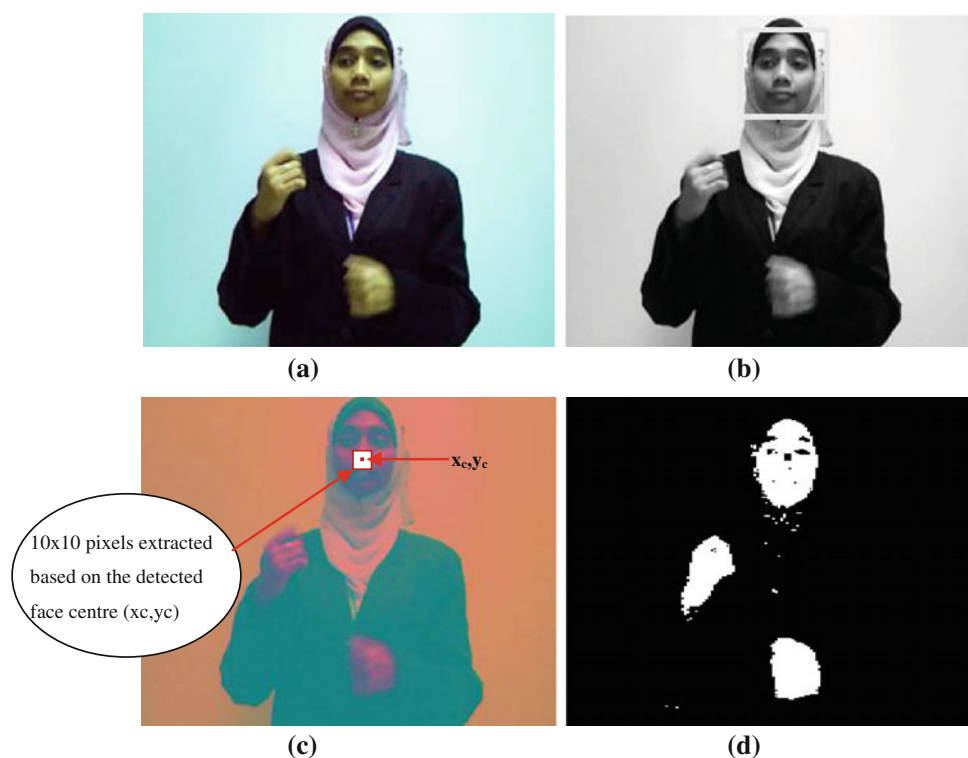
To perform the skin detection from an image, the image needs to be converted from RGB to  $YC_bC_r$  colour space after the face or hand has been detected. Then, to find skin pixels that fall within the same colour space range of the detected hand or face, a generation of  $10 \times 10$  pixels box from the centre of the face has been extracted as shown in Fig. 6c.

In the  $YC_bC_r$  colour space format, the image is stored in three components: the Y component which contains the luminance information of the image, while the  $C_b$  and  $C_r$  components hold the colour difference information of the image. Then, a range of skin pixels has been specified based on the detected face, the distribution of skin pixels values is highlighted using  $C_b$  and  $C_r$  components and  $(C_b - C_r$  or  $C_r - C_b)$  as an additional threshold while luminance Y component is discarded. The segmented image is shown in Fig. 6d.

**Fig. 5**  $YC_bC_r$  colour distribution within the skin detected area. **a** Search skin area, **b** Y component, **c**  $C_b$  component and **d**  $C_r$  component



**Fig. 6** **a** Original RGB image, **b** detected face on grey scale image, **c** YCbCr image, **d** the segmented image based on the  $C_b$ ,  $C_r$  and  $C_r-C_b$  threshold



### 3.3.1 Skin area choice

As stated in Sect. 3.1, face or hand has been chosen as an initial skin area, where the search for skin region pixels can start. However, one can find that face region contains pixels that are not skin colour such as eyes and lips area. Therefore, we found that open palm is more than appropriate for searching skin pixels than face region for non-dark skin coloured people. For dark skin coloured people, we found that their front palm area differs in skin colour from other skin areas (see Fig. 7a). In such case, it is more practical first to define a box of  $10 \times 10$  pixels around the centre of both hand and face of the person. Then, the  $C_r$  and  $C_b$  of the detected pixels from the hand palm are merged with the ones from the detected face to form a combined histogram as shown in Fig. 7b, c. The result in Fig. 7c has proven that in general, it is useful for applications, such as sign language recognition (SLR) to use hand as an initial blob for skin detection.

Meanwhile, for applications that require face detection, such as facial recognition, we have to initialize the developed method with a detected face to capture more features in real-time applications as shown in Fig. 7e. The threshold ratio of  $C_b$  and  $C_r$  is merged to obtain more skin pixels to give a better skin detection result. This is shown in Fig. 7f.

### 3.4 Tracking the blobs

There are various approaches for tracking hand and face that have demonstrated better results for specific

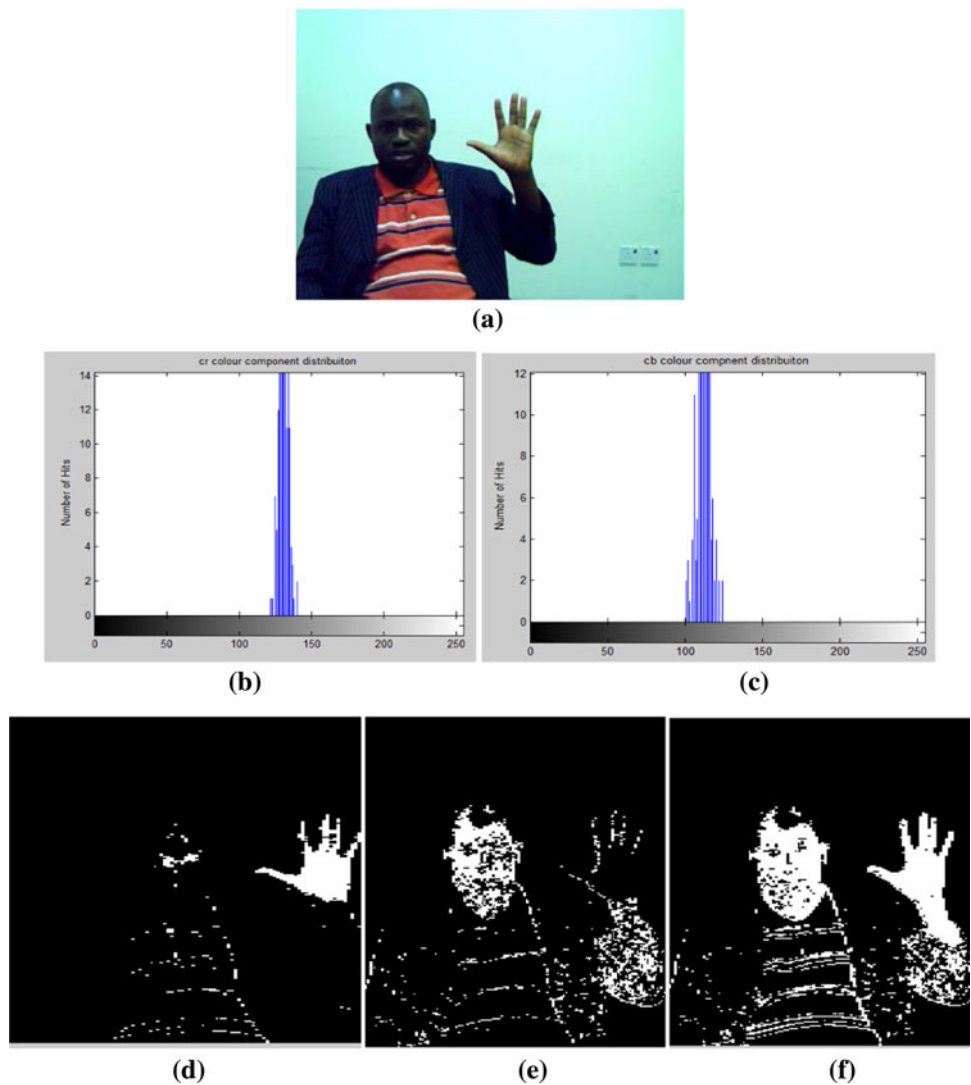
applications. For example, frame subtraction and optical flow motion tracking are sufficient and reliable for sensing the motion of moving objects. Continuously adaptive mean shift (CAMSHIFT) [3] is designed for face and coloured object tracking. However, CAMSHIFT alone has not been used to track more than one object in the image frame.

In this work, we have considered three skin blobs; a simple center tracking technique has been used to track the detected blobs in the image. The center tracking uses the center pixels  $x_c$  and  $y_c$  of the detected skin blob to track the object.

## 4 Experimental results

We have experimented that appearance-based methods such as Haar-like are view specific algorithms; i.e. they are limited to certain shapes with high accuracy of detection, so these methods are applied in our proposed method as an initial stage. As mentioned in the previous sections, a skin blob has been detected in the initial frames and therefore pixels outside the detected blobs are regarded as non-skin pixels. All experiments were run on  $320 \times 240$  pixel resolution images or videos in AVI format on a Dell Desktop workstation T3500 (Quad Core Intel Xeon W3520, 6 GB DDR3 SDRAM 1066 MHZ).

**Fig. 7** **a** Open hand posture with part of arm. **b, c**  $C_r$  and  $C_b$  pixel distribution of the detected hand, respectively. **d** Results using the palm area, **e** results using the face area. **f** Results using the merged  $C_r$  and  $C_b$  from the detected hand and face



**Table 1** System performance using Haar-like and AdaBoost for face detection

Face	Classification			Average detection time (ms)	Accuracy
	Images	Hits	Missed		
Auto illumination	50	49	1	84.25	98 %

#### 4.1 Face detection using Haar-like and AdaBoost algorithm

The proposed method has made use of available face detection algorithm in OpenCV library. The available algorithm for face detection has been trained using 3,000 positive images from FERET database [8] and 3,000 negative image sets which were taken from different websites over the Internet and resized to  $20 \times 20$ . AdaBoost is used to linearly combine many weak classifiers to detect the face. Our proposed method has tested the

available face classifier using 50 images collected in an auto-illuminated environment. The accuracy of the face detector is shown in Table 1.

#### 4.2 Hand detection using Haar-like and AdaBoost algorithm

In OpenCV library, only face detectors and upper body detectors are available, but no hand detector available as of June, 2011. Therefore, similarly to the face detection approach in Sect. 4.1, a set of eight weak classifiers using a

simple Haar features has been built for hand detection. Each weak classifier is a simple detector based on Haar-like and AdaBoost algorithms have been used to detect the open hand palm. The training process was done using 600 samples of an open palm as shown in Fig. 8 collected using a camera with  $320 \times 240$  resolutions. The open hand has been detected as shown in Fig. 9 and the proposed method accuracy is shown in Table 2 [2].

#### 4.3 Skin segmentation using detected blob

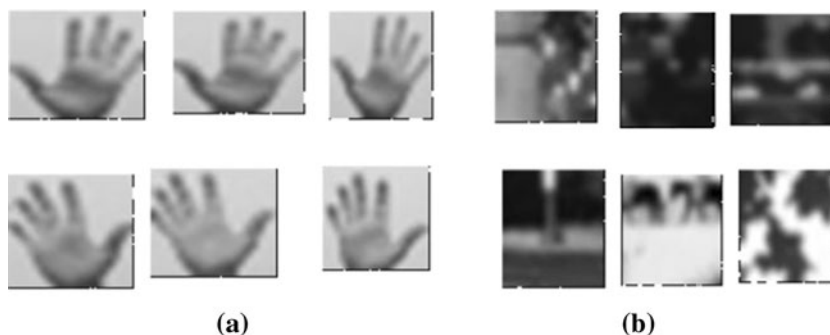
The proposed approach focuses on reducing the search area where possible skin pixels appear. This has been achieved by detecting a face or hand as stated in Sects. 4.1 and 4.2. Following that, a range of  $10 \times 10$  of pixels from the center of origin is used as skin pixels range for detecting the full blob colour areas. Then, a model of the desired  $YC_bC_r$  is created using the blob region colour range. For real-time applications, the open hand palm is used as an initial stage for detecting skin region. Meanwhile, for images and videos that downloaded from the Internet to test the algorithm, we use the detected face region as stated

in Sect. 4.1 as an initial stage to search for skin pixels. It can be concluded that, if the skin search area within the image is narrowed down then the  $YC_bC_r$  colour space performs better than other colour spaces (Fig. 9b). The experimental results using the three colour spaces HSV, RGB and  $YC_bC_r$  are shown in Fig. 10.

#### 4.4 Evaluation of the proposed method using face, hand and face combined with hand

The developed method for skin blobs detection described in the previous sections has proven that either face or hand, or both face and hand can be detected as the initial clue for skin pixels within the image frame. In Sect. 3.3.1, it has been described how the developed skin detection method can give different true-positive results when either face or hand has been used to initiate the proposed method for dark skin person. In addition, proposed method performance using combined pixels' range from face and hand has been described in Fig. 7. The proposed method performance differs when using either face or hand because of colour variations of the dark person hand palm and face.

**Fig. 8** **a** Positive samples, **b** negative samples



**Fig. 9** **a** Coloured image, **b** hand posture detection



**Table 2** System performance using Haar-like and AdaBoost for open hand posture detection [2]

Open posture	Classification			Average detection time (ms)	Accuracy
	Images	Hits	Missed		
Auto illumination	50	45	5	83.23	90 %



**Fig. 10** Using the proposed method with three colour space. **a** Coloured image, **b** detected hand used for initialization, **c**  $YCbCr$ , **d** RGB and **e** HSV



In addition, two images have been chosen randomly to show the proposed method effectiveness when using face, hand and, face and hand combined together for other skin colour range.

In Fig. 11b, c, it is obvious that using hand or face, respectively, as an initial clue to detect skin blobs, do not make much difference in the segmented skin areas. Meanwhile, by combining the two images of Fig. 11b, c, the result in Fig. 11d appears better than Fig. 11b, c, where more true-positive pixels appear in the white segmented region.

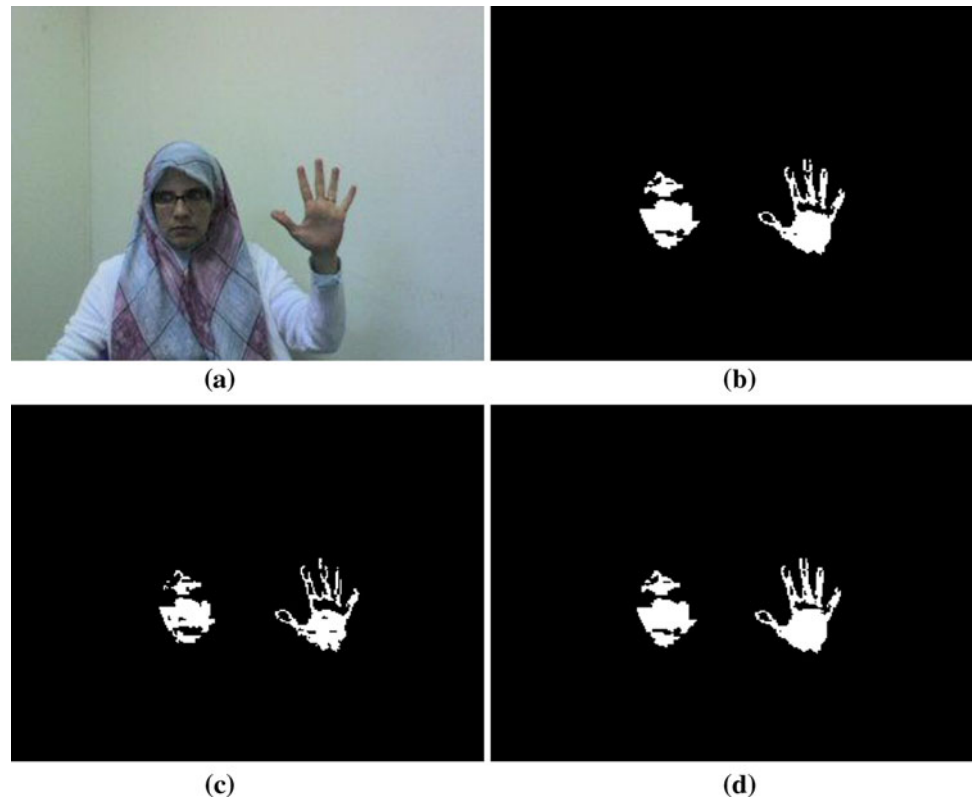
On the other hand, the algorithm has been tested on an image as shown in Fig. 12a, which has variations in illumination and colour range within the same region, such as

face. Figure 12b shows that when hand has been chosen for the proposed method initiation, more true-positive pixels appear in the segmented hand area. Meanwhile, in Fig. 12c, when face has been used as an initial clue for skin pixels, more true-positive skin pixels appear in the face region. Finally, when the skin pixels' range has been extracted from both hand and face region, it is obvious that better skin segmentation results appear for both hand and face regions as shown in Fig. 12d.

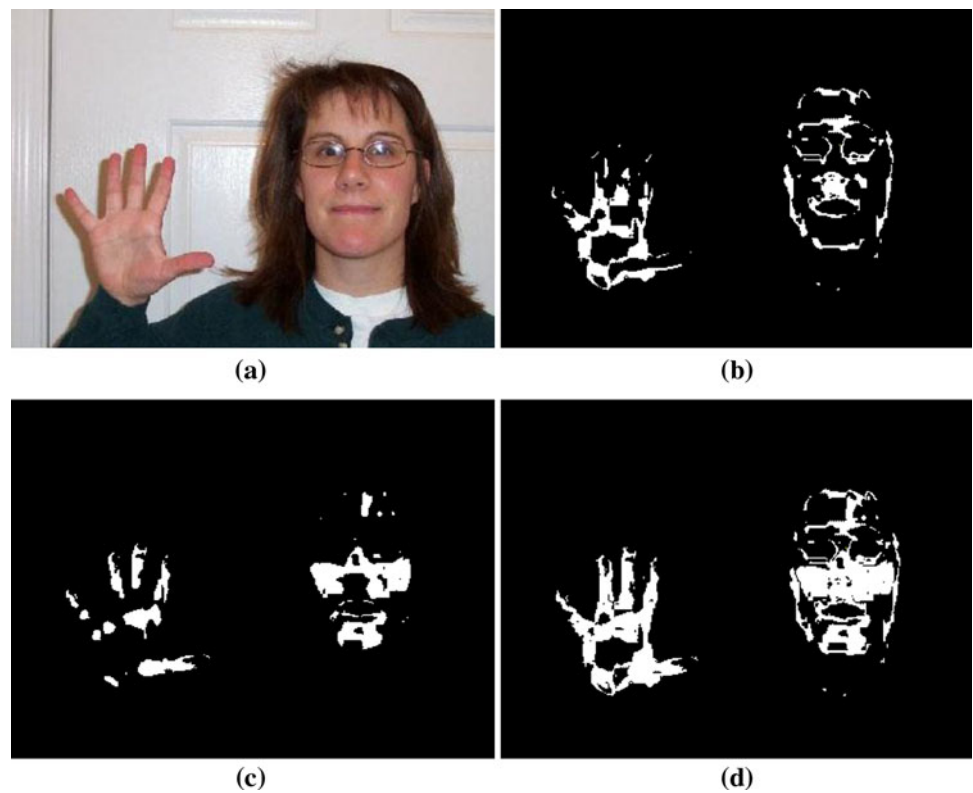
From the aforementioned results, we summarize that:

- (1) Based on the initial region either face or hand, the segmented image will have more true-positive skin pixels in that region.

**Fig. 11** a Original image, b using hand pixel range, c using face pixel range and d combined face and hand pixel range



**Fig. 12** a Original image, b using hand pixel range, c using face pixel range and d combined face and hand pixel range



- (2) Combining face and hand extracted skin pixels, followed by applying the range of the extracted pixels on the entire image, the segmentation results appear the same if we add the segmented image based on pixels extracted from hand with the segmented image based on pixels extracted from face as in Fig. 11d and 12d.

#### 4.5 Blob tracking

After the skin blobs have been detected, the center pixels are used to track the blobs over time. In case of variable lighting conditions within same video frames, the proposed method will update the values for  $C_r$ ,  $C_b$  and  $C_r - C_b$  as an additional threshold value from the moving blobs. Figure 13 shows the tracking of skin blobs over time.

#### 4.6 The overall proposed method stages

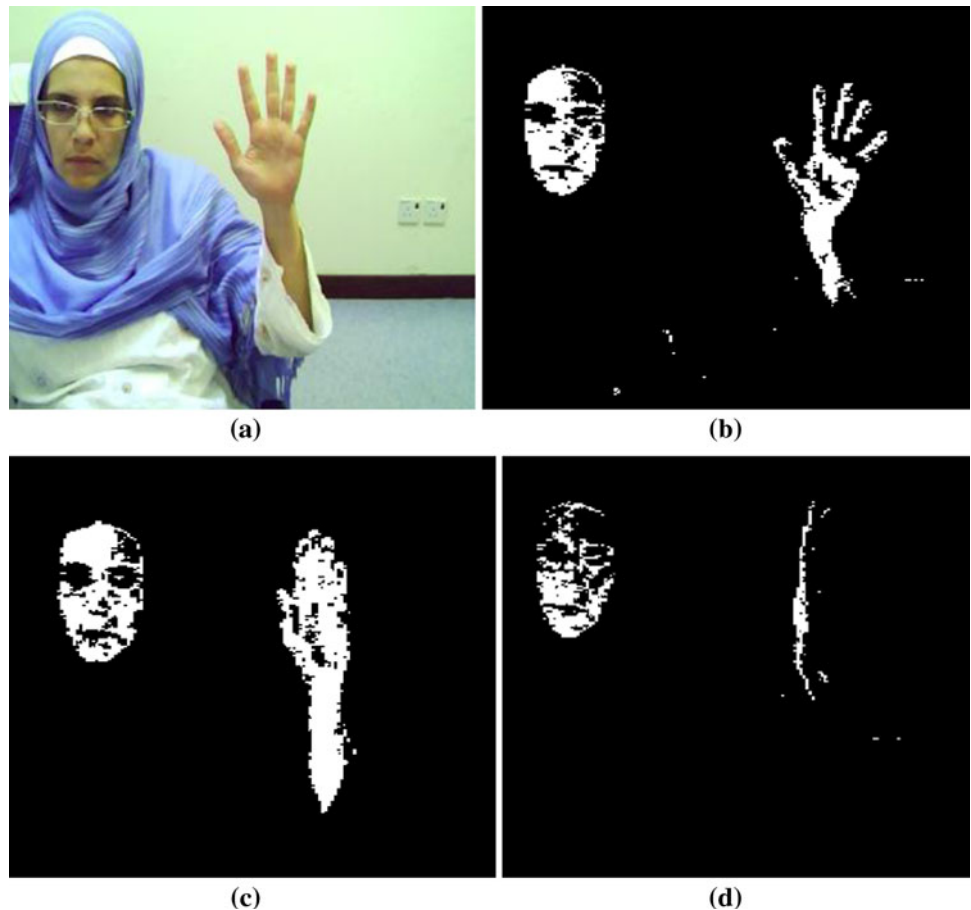
Haar-like approach with AdaBoost algorithm has given better detection rate for face rather than hand (see Tables 1, 2). From the experiments in Sect. 3.2.1, it is obvious that

the application plays an important role in choosing a prior face or hand to initiate the proposed method with an average delay time of 83–85 ms. Moreover, the overall time required for the proposed method to complete the skin segmentation algorithm is about 90 ms, with the maximum delay in the face or hand detection. During the process stage of skin detection, there are some false pixels which are not a skin, but have been observed as skin pixels. Figure 14c shows the false pixels, however, this area is considerably small as compared to the original skin blobs during the tracking process.

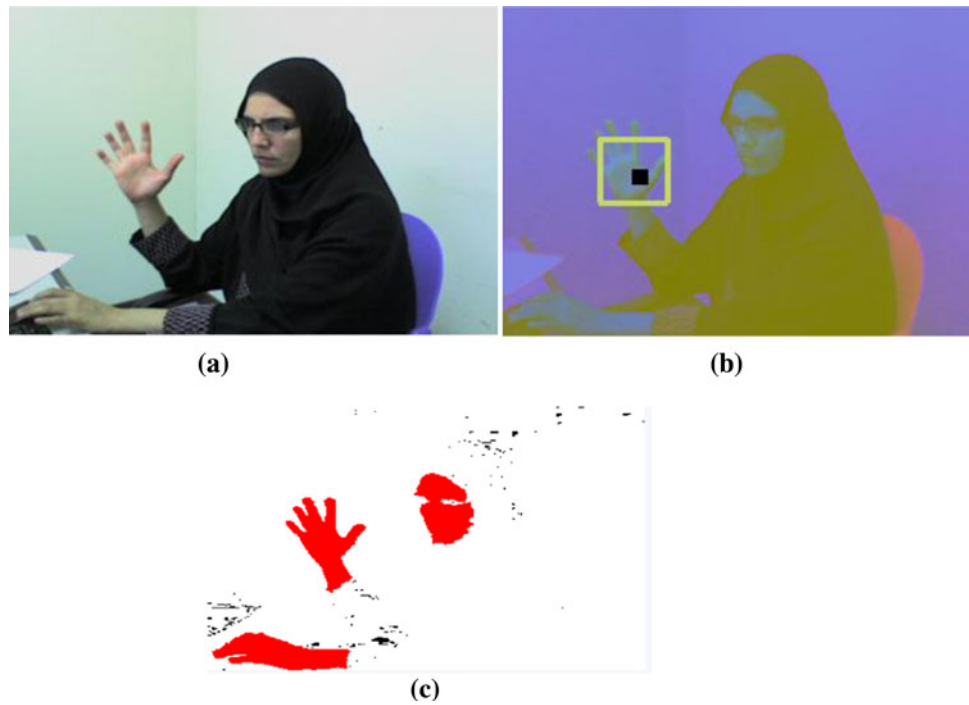
The complete proposed method can be summarized as follows (see Fig. 14):

1. The face or hand has been detected using Haar-like feature and AdaBoost algorithm introduced by [19, 24].
2. Skin colour pixels were obtained from the center of the detected blob; a seed pixel  $(x_c, y_c)$  has been used to find the skin blob colour. Pixels that have values different from the majority have been neglected.
3. The skin segmentation has been achieved using the thresholds values of  $C_b$ ,  $C_r$  and  $C_b - C_r$  as an additional threshold for filtering the image. Blob extraction has

**Fig. 13** The changes in output over time. **a** Frame 1, **b** Frame 90, **c** Frame 120 and **d** Frame 220



**Fig. 14** **a** Original image, **b** hand detect using Haar-like and **c** skin segmentation and tracking



**Fig. 15** Comparison using: **a** tested image, **b** GMM skin detection model [13], **c** [6] approach and **d** proposed method using  $YCbCr$  colour space



- been accomplished with a region of pixels from the center seed pixel.
- The tracking of the detected skin colour areas using the center of the detected skin blobs has been achieved.

## 5 Comparison with other approaches

A robust skin classifier should be able to detect different skin types under a large variety of illumination conditions and different backgrounds. Many of the existing skin detection techniques consider only a few skin types and a few possible illumination conditions [1]. Not many reported skin detection approaches [17] used face as an initial stage for detecting skin pixels without modeling skin pixels. The proposed method uses face as well as hand to detect and model skin pixels. Skin detection performance was tested for 40 videos with variable frame length which contain human upper body. All presented results were obtained for the same set of three chosen frames from the videos so they can be compared with each other. Skin detection effectiveness was measured by

- False-positive rate ( $FP$ ), which is a percentage of non-skin pixels classified as skin pixels,

- The F-measure is defined as a harmonic mean of Precision ( $P$ ) and Recall ( $R$ ), where an F-measure score reaches its best value at 1 and worst score at 0 [21]. Equations (3–5) shows the F-measure calculations:

$$F - measure = 2 \times \frac{precision \times recall}{precision + recall} \quad (3)$$

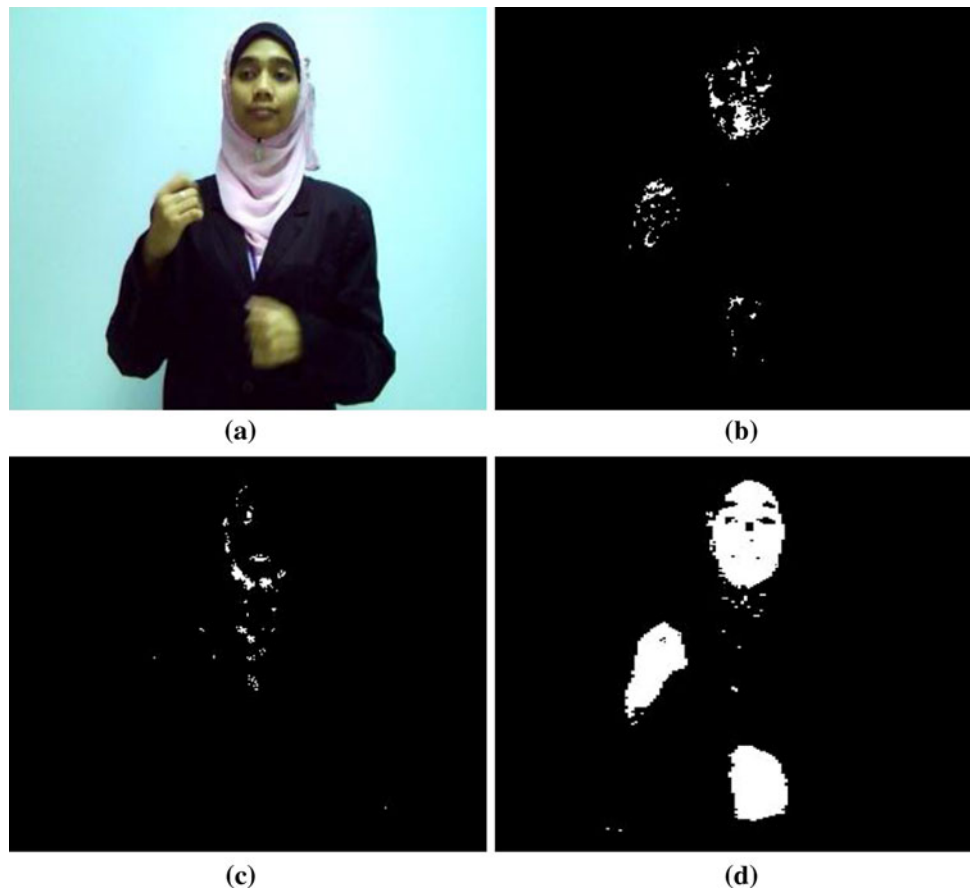
$$P = \frac{TP}{(TP + FP)} \quad (4)$$

$$R = \frac{TP}{(TP + FN)} \quad (5)$$

where  $TP$  is the true positive, which is a percentage of skin pixels classified as skin pixels.  $FN$  is the false negative, which is a percentage of skin pixels classified as non-skin pixels.

The proposed approach is compared with the statistical skin model in [6] and Gaussian mixture model (GMM) in [13]; the results are shown in Table 3. The performance of the statistical models is higher when compared with the parametric methods as stated in [17]. The proposed method is developed to be robust and adequate for real-time applications. Therefore, when we have compared our approach with the recently reported approaches [6, 13], which model various ranges of human skin colour under different illumination conditions, relatively high  $FP$  values

**Fig. 16** Comparison using: **a** tested image, **b** GMM skin detection model [13], **c** Conaire et al. approach [6] and **d** proposed method using  $YC_bC_r$  colour space



**Table 3** False positives and false negatives

Skin detection method	FP (%)	F-measure
Statistical [6]	1.2928	0.5219
Adaptive GMM [13]	2.5816	0.6109
Proposed approach	0.9808	0.4970

from (Table 3) show that their approaches were so noisy in real-time applications.

The statistical model and the adaptive GMM gave sufficiently good F-measure, but higher *FP* (see Table 3) as compared to the proposed method. From the experiments, we observed that using the detected hand or face as an initial stage rather than modeling skin colours types to identify skin area can reduce the *FP* and increase the proposed method accuracy. Figures 15 and 16 show the results of skin detection using the GMM approach, the statistical approach and the proposed method. In Fig. 15, it is clear that the GMM method and the proposed method have similar skin detection areas; meanwhile the statistical approach has a high FN. The results in Fig. 16 show that the proposed method has the highest skin detection areas comparable to the existing methods. This is true because the proposed method requirements, such as a person is required to stand approximately 1.5–3 m in front of the camera and be visible in the scene, were imposed.

## 6 Conclusion

A robust method has been developed for human skin colour detection. The method, which requires a person with visible face and hand on the scene with range of 1.5–3 m, was developed for real-time implementation with 30 fps. An average delay time of 83–85 ms was measured during the initiation stage of face or hand detection, using Haar-like and Adaboost algorithm. Haar-like and Adaboost approaches have given sufficiently good blob detection results in different illumination conditions and in the presence of background cluster as the developed method was well-trained offline. The advantage of the proposed method lies in its initial stage that it does not require to build a skin samples model or search for skin pixels over a number of frames, except for the pixels within the detected blob region. The user does not need to be concerned with illumination, neither is there any need for a uniform background.

The original contribution of this work is the proposal of a novel technique for combining appearance-based method with the colour space  $YC_bC_r$  techniques for achieving a real-time skin blobs detector. Another important contribution is that the proposed method enables to deal with larger

set of individuals' skin colours without concerning about the skin colour range. The proposed method can also narrow down the search area for skin without the need for skin databases. In addition, it is computationally fast and independent of background, illumination or skin colour. As a future direction, the proposed dynamic skin blob detection algorithm can be combined with other motion detection algorithms for better blob tracking if occlusion appears between face and hand.

**Acknowledgments** Funded by Research Matching Grant Scheme RMGS 09-03, IIUM.

## References

- Albiol, A., Torres, L., Delp, E.: J. Optimum color spaces for skin detection. In: The Proceeding of the International Conference on Image Processing, vol. 1, pp. 122–124 (2001)
- Bilal, S., Akmeliawati, R., Salami, M., Shafie, A., Bouhabba, E.: A hybrid method using haar-like and skin-color algorithm for hand posture detection, recognition and tracking. In: The International Conference on Mechatronics and Automation (ICMA), Xi'an, China, pp. 934–939 (2010)
- Bradski, G.R.: Computer vision face tracking for use in a perceptual user interface. Intel. Technol. J. **2**(3), 1–15 (1998)
- Brand, J., Mason, J.: A comparative assessment of three approaches to pixel level human skin-detection. In: The Proceeding of the International Conference on Pattern Recognition, vol. 1, pp. 1056–1059 (2000)
- Chen, Q., Georganas, N.D., Petriu, E.M.: Real-time vision-based hand gesture recognition using haar-like features. In: The Instrumentation and Measurement Technology Conference—IMTC, pp. 1–6 (2007)
- Conaire, C.Ó., O'Connor, N.E., Smeaton, A.F.: Detector adaptation by maximising agreement between independent data sources. In: The IEEE International Workshop on Object Tracking and Classification Beyond the Visible Spectrum, pp. 1–6 (2007)
- Dadgostar, F., Sarrafzadeh, A.: An Adaptive real-time detector based on Hue thresholding: a Comparison on two motion tracking Methods. Pattern. Recogn. Lett. **27**, 1342–1352 (2006)
- Feret Face Database. (2012). [http://www.itl.nist.gov/iad/humanid/feret/feret\\_master.html](http://www.itl.nist.gov/iad/humanid/feret/feret_master.html)
- Filipe, T., Tiago, C., Hamid, S.: Improved automatic skin detection in color images. Dig. Image. Comp. Tech. App. **3**, 419–427 (2003)
- Freund, Y., Schapire, R.E.: A short introduction to boosting. J. Japanese. Soc. Artificial. Intel. **14**(5), 771–780 (1999)
- Gomez, G., Moralez, E.: Automatic feature construction and a simple rule induction algorithm for skin detection. In: The Proceeding of the ICML Workshop on Machine Learning in Comp. Vision, pp. 31–38 (2002)
- Hsu, R.-L., Abdel-Mottaleb, M., Jain, A.K.: Face detection in color images. IEEE. Trans. Pattern. Anal. Mach. Intel. **24**(5), 696–706 (2002)
- Hsuan, M., Yang, Ahuja, N.: Gaussian Mixture Model for human skin color and its application in image and video databases. In: The Proceeding of SPIE: Conference on Storage and Retrieval for Image and Video Databases, vol. 3656, pp. 458–466 (1999)
- Imagawa, K., Lu, S., Igi, S.: Color-based hands tracking system for sign language recognition. In: The Proceeding of Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 462–467 (1998)

15. Jones, M.J., Rehg, J.M.: Statistical color models with application to skin detection. *Int. J. Comput. Vis.* **46**(1), 81–96 (2002)
16. Kakumanu, P., Makrogiannis, S., Bourbakis, N.: A survey of skin-color modeling and detection methods. *Pattern. Recogn.* **40**(3), 1106–1122 (2007)
17. Kawulok, M.: Dynamic skin detection in color images for sign language recognition. In: *The 3rd International Conference on Image and Signal Processing (ICISP)*, LNCS vol. 5099, pp. 112–119 (2008)
18. Kukharev, G., Nowosielski, A.: Fast and efficient Algorithm for Face Detection in Colour Images. *Mach. Graphics. Vis.* **13**(4), 377–399 (2004)
19. Lienhart, R., Maydt, J.: An extended set of Haar-like features for rapid object detection. In: *The Proceeding of IEEE International Conference on Image Processing ICIP*, vol. 1, pp. 900–903 (2002)
20. Lu, J.W., Gu, Q., Plataniotis, K.N., Wang, J.: A comparative study of skin-color models. In: *The International Conference of Image Analysis and recognition (ICIAR)*, pp. 729–736 (2005)
21. Sasaki, Y. (2007). The truth of the F-measure. In *S. o. C. Science* (Ed.): University of Manchester
22. Sigal, L., Sclaroff, S., Athitsos, V.: Skin color-based video segmentation under time-varying illumination. *IEEE. Trans. Pattern. Anal. Mach. Intel.* **26**(7), 863–877 (2004)
23. Vezhnevets, V., Sazonov, V., Andreeva, A.: A survey on pixel-based skin color detection techniques In: *The Proceeding of the GraphiCon 2003, Moscow-Russia*, pp. 85–92 (2003)
24. Viola, P., Jones, M.J.: Robust real-time face detection. *Int. J. Comput. Vis.* **57**(2), 137–154 (2004)
25. Wong, K.W., Lam, K.M., Siu, W.C.: An efficient algorithm for human face detection and facial feature extraction under different conditions. *Pattern. Recogn.* **34**, 1993–2004 (2001)
26. Wong, K.W., Lam, K.M., Siu, W.C.: A robust scheme for live detection of human face in color images. *Signal. Process. Image. Commun.* **18**(2), 103–114 (2003)
27. Yang, J., Lu, W., Waibel, A.: Skin-color modeling and adaptation. In: *The Asian Conference on Computer Vision ACCV'97*, pp. 687–694 (1997)
28. Yang, J., Waibel, A.: A Real-Time Face Tracker. In: *The Proceeding of WACV*, pp. 142–147 (1996)
29. Yoo, T., Oh, S.: A Fast Algorithm for Tracking Human Faces Based on Chromatic Histograms. *Pattern. Recognit. Letters.* **20**, 967–978 (1999)
30. Zhu, Q., Cheng, K.-T., Wu, C.-T., Wu, Y.-L.: Adaptive learning of an accurate skin-color model. In: *The Proceeding of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 37–42 (2004)

## Author Biographies

**Sara Bilal** Did her B. Tech in Electronic Engineering from Sudan University of Science and Technology 2002. Sara holds an M. Sc degree in Electrical and Electronic Engineering from University Technology Malaysia in 2005. She is currently pursuing PhD at International Islamic University Malaysia, IIUM. Her research

interests are in the areas of Image Processing, Computer Vision and Pattern Recognition.

**Rini Akmeliawati** (M'04–SM'07) was born in Jakarta, Indonesia. She obtained the Bachelor of Engineering (Honours) in Electrical Engineering from Royal Melbourne Institute of Technology (RMIT) University, Australia in 1997, and her PhD in Electrical and Electronics Engineering from the University of Melbourne, Australia in 2002. She is currently an Associate Professor of the Department of Mechatronics Engineering at the International Islamic University Malaysia. Previously, she was a lecturer at Royal Melbourne Institute of Technology (RMIT) University, Australia (2001–2004) and Monash University (2004–2008). Her main research interests include nonlinear control systems theory and applications, intelligent systems and image processing. She has published more than 80 international journal and conference papers. Rini Akmeliawati is a Senior Member of IEEE. She was the Chair of IEEE Instrumentation and Measurement Society—Malaysia Chapter 2007–2009, the treasurer and the vice-chair of the same society in 2010 and 2011, respectively, and the secretary of IEEE Control Systems Society—Malaysia Chapter in 2010. She is also currently the Chair of Intelligent Mechatronics System Research Unit at the International Islamic University Malaysia.

**Momoh-Jimoh Eyiomika Salami** has received my B.S.E.E degree from the University of Ife, Ile-Ife, Nigeria in June 1977, and a PhD degree in Electrical Engineering from the University of Calgary, Calgary, Canada in 1985. I have been actively involved in teaching and research activities at the university level in Nigeria (at the University of Ilorin), Saudi Arabia (at King Saud University) and Malaysia since the completion of my graduate studies. I joined International Islamic University Malaysia (IIUM) in September 1996 and I am currently a Professor in the Department of Mechatronics Engineering as well as the Deputy Dean of Postgraduate and Research, Faculty of Engineering, IIUM. I have authored/co-authored more than 100 publications in both local and international journals and conference proceedings as well as being one of the contributors in a recently published book entitled 'The Mechatronics Handbook' edited by Prof. Bishop. My research interests include digital signal and image processing, intelligent control system design and instrumentation. I am a senior member of IEEE.

**Amir Akramin Shafie** (PhD) received his B.Eng. (Hons) in Mechanical Engineering from the University of Dundee and Master of Science in Mechatronics from University of Abertay Dundee, Scotland. He has been conferred a doctorate in the field of Artificial Intelligence by University of Dundee, Scotland in 2000. He is currently attached to International Islamic University Malaysia as the Dean of Kulliyah of Engineering and Associate Professor after serving in industry as Industrial Researcher. Dr Amir has been principal researcher for various research grants over the last 5 years mainly in the area of autonomous mechatronic system which aims to develop intelligent robotic system for industrial and other applications. He currently has various published articles in books, refereed journals and various international conferences, some of which have been highly cited.