A Fusion Approach for Efficient Human Skin Detection

Wei Ren Tan, Chee Seng Chan, *Member, IEEE*, Pratheepan Yogarajah, and Joan Condell

Abstract—A reliable human skin detection method that is adaptable to different human skin colors and illumination conditions is essential for better human skin segmentation. Even though different human skin-color detection solutions have been successfully applied, they are prone to false skin detection and are not able to cope with the variety of human skin colors across different ethnic. Moreover, existing methods require high computational cost. In this paper, we propose a novel human skin detection approach that combines a smoothed 2-D histogram and Gaussian model, for automatic human skin detection in color image(s). In our approach, an eye detector is used to refine the skin model for a specific person. The proposed approach reduces computational costs as no training is required, and it improves the accuracy of skin detection despite wide variation in ethnicity and illumination. To the best of our knowledge, this is the first method to employ fusion strategy for this purpose. Qualitative and quantitative results on three standard public datasets and a comparison with state-of-the-art methods have shown the effectiveness and robustness of the proposed approach.

Index Terms—Color space, dynamic threshold, fusion strategy, skin detection.

I. INTRODUCTION

W ITH the progress of information society today, images have become more and more important. Among them, skin detection plays an important role in a wide range of image processing applications from face tracking, gesture analysis, content-based image retrieval systems to various human–computer interaction domains [1]–[6]. In these applications, the search space for objects of interests, such as hands, can be reduced through the detection of skin regions. One of the simplest and commonly used human skin detection methods is to define a fixed decision boundary for different color space components [7]–[9]. Single or multiple ranges of threshold values for each color space components are defined and the image pixel values

that fall within these predefined range(s) are selected as skin pixels. In this approach, for any given color space, skin color occupies a part of such a space, which might be a compact or large region in the space. Other approaches are multilayer perceptron [10]–[12], Bayesian classifiers [13]–[15], and random forest [16]. These aforementioned solutions that use single features, although, successfully applied to human skin detection; they still suffer from the following. 1) *Low Accuracy:* False skin detection is a common problem when there is a wide variety of skin colors across different ethnicity, complex backgrounds and high illumination in image(s). 2) *Luminance-invariant space:* Some robustness may be achieved via the use of luminance invariant color space [1], [17]; however, such an approach can withstand only changes that skin-color distribution undergo within a narrow set of conditions and also degrades the performance [18]. 3) *Require large training sample:* In order to define threshold value(s) for detecting human skin, most of the state-of-the-art work requires a training stage. One must understand that there are tradeoffs between the size of the training set and classifier performance. For example, Jones and Rehg [15] required 2 billion pixels collected from 18 696 web images to achieve optimal performance. In this paper, we propose a novel approach—fusion framework—that uses product rules on two features; the smoothed 2-D histogram and Gaussian model to perform automatic skin detection. First of all, we employ an online dynamic approach as in [19] to calculate the skin threshold value(s). Therefore, our proposed method does not require any training stage beforehand. Second, a 2-D histogram with smoothed densities and a Gaussian model are used to model the skin and nonskin distributions, respectively. Finally, a fusion strategy framework using the product of two features is employed to perform automatic skin detection. To the best of our knowledge, this is the first attempt that employs a fusion strategy to detect skin in color image(s). The image pixels representation in a suitable color space is the primary step in skin segmentation in color images. A better survey of different color spaces (e.g., RGB, YCbCr, HSV, CIE Lab, CIE Luv, and normalized RGB) for skin-color representation and skin-pixel segmentation methods is given by Kakumanu *et al.* [20]. In our approach, we do not employ the luminance-invariant space. Indeed, we choose the log opponent chromaticity (LO) space [21]. The reasons are twofold: first, color opponency is perceptually relevant as it has been proven that the human visual system uses an opponent color encoding [22], [23]; and second, in this LO color space, the use of logarithms renders illumination change to a simple translation of coordinates. Most of the aforementioned solutions claimed that illumination variation is one of the contributing factors that degrade the performance of skin detection systems. However, our empirical
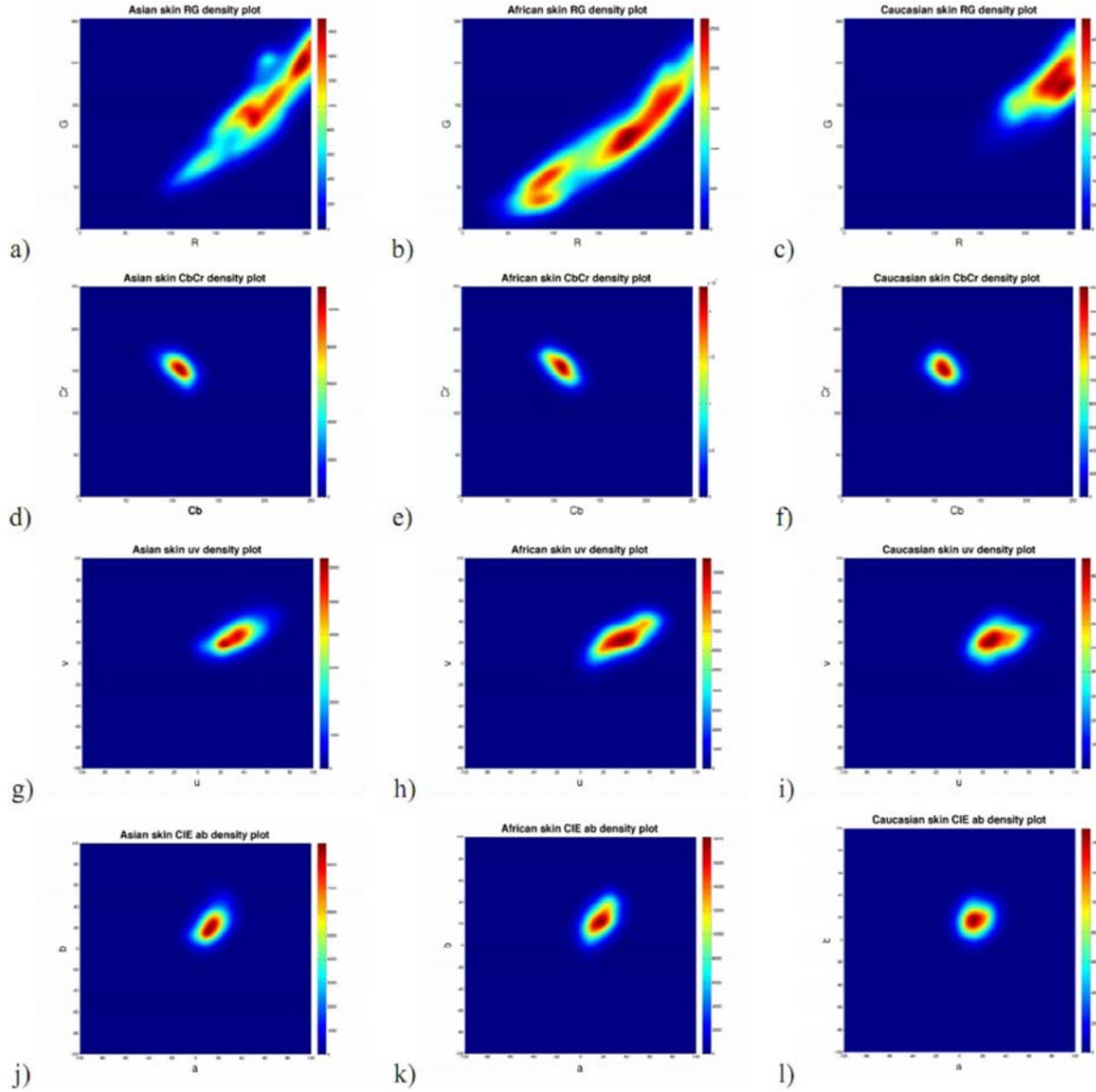
Fig. 1. Density plots of Asian, African, and Caucasian skin in different color spaces, adopted from [25]. Each row represents different color space and columns from left to right represent Asian, African, and Caucasian, respectively.

results and Jayaram *et al.* [18] showed that the absents of luminance component do not affect the system performance. The remainder of this paper is structured as follows. Section II gives a brief description of related work in human skin segmentation. Section III derives our proposed fusion strategy. Section IV presents the experimental results using three different datasets, and Section V concludes this paper with discussions and future work.

II. RELATED WORK

Skin detection is the process of finding skin-color pixels and regions in an image or video. In images and videos, skin color is an indication of the existence of humans in media. In one of the early applications, detecting skin-color regions was used to identify nude pictures on the Internet for content filtering. In another early application, skin detection was used to detect anchors in TV news videos for the sake of video automatic annotation, archival, and retrieval. Interested readers are encouraged to refer to [20] and [24] for a detailed background review. A skin detector typically transforms a given pixel into an appropriate color space and then uses a skin classifier to label the pixel whether it is skin or nonskin. A skin classifier defines a decision boundary of the skin-color class in the color space based on a training database of skin-color pixels. For example, Sobottka and Pitas [7] used fixed range values on the HS color space where the pixel values belong to skin pixels in the range of $R_H = [0, 50]$ and $R_S = [0.23, 0.68]$. Wang and Yuan [9] used threshold values in RG space and HSV space where threshold values are set to be within the range $R_r = [0.36, 0.465]$, $R_g = [0.28, 0.363]$, $R_H = [0, 50]$, $R_S = [0.20, 0.68]$, and $R_V = [0.35, 1.0]$ to differentiate skin and nonskin pixels. In these approaches, high false skin detection is a common problem when there are a wide variety of skin colors across different ethnicity, complex backgrounds, and high illumination. Fig. 1 shows that the skin color of
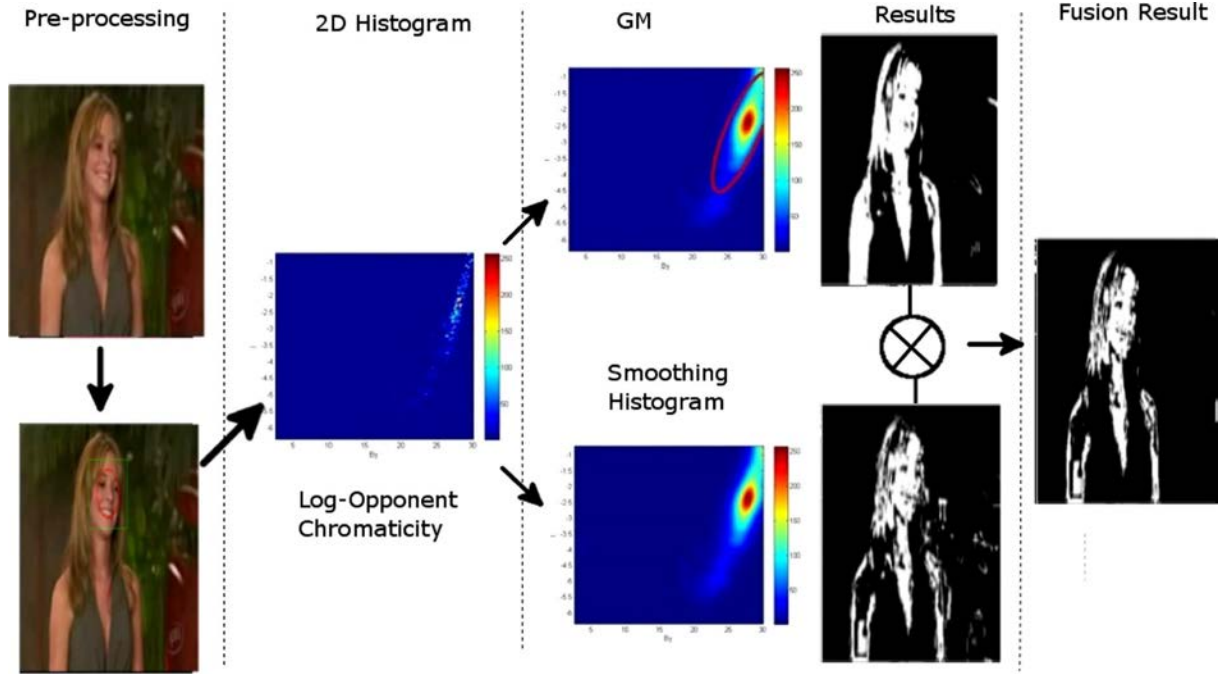
Fig. 2.   Proposed framework: eye detector, 2-D histogram, Gaussian model, and fusion strategy.

people belongings to Asian, African, and Caucasian groups is different from one another and ranges from white, yellow to dark. Some robustness may be achieved via the use of luminance invariant color spaces [1], [17]; however, such an approach can only cope if the change in skin-color distribution is within a narrow set of conditions [18]. Other approaches are multilayer perceptron [10]–[12], Bayesian classifiers [13]–[15], and random forest [16]. In multilayer perceptron-based skin classification, a neural network is trained to learn the complex class conditional distributions of the skin and nonskin pixels. Brown *et al.* [10] proposed a Kohonen network-based skin detector where two Kohonen networks, skin only and skin plus nonskin detectors, were trained from a set of about 500 manually labeled images to obtain an optimal result. Sebe *et al.* [13] used a Bayesian network with training data of 60 000 samples for skin modeling and classification. Friedman *et al.* [14] proposed the use of tree-augmented Naive Bayes classifiers for skin detection. The Bayesian decision rule to minimum cost is a well-established technique in statistical pattern classification. Jones and Rehg [15] used the Bayesian decision rule with a 3-D RGB histogram model built from 2 billion pixels collected from 18 696 web images to perform skin detection. Readers are encourage to read [20] and [24] for a detailed state-of-the-art review. Although these solutions had been very successful, they suffer from a tradeoff between precision and computational complexity. In summary, our proposed method has two advantages in comparison to the state-of-the-art solutions. First of all, our proposed skin detection method employs an online dynamic threshold approach. With this, a training stage can be eliminated. Second, we select a fusion strategy for our skin detector. To the best of our knowledge, this is the first attempt that employs a fusion strategy to detect skin in color image(s).

III. OUR METHOD

Fig. 2 shows the proposed framework for automatic skin detection. First, an approach similar to that of Fusel *et al.* [26] is adopted to obtain the face(s) in a given image. Second, a dynamic method is employed to calculate the skin threshold value(s) on the detected face(s) region. Third, two features—the 2-D histogram with smoothed densities and Gaussian model—are introduced to represent the skin and nonskin distributions, respectively. Finally, a fusion framework that uses the product rule on the two features is employed to obtain better skin detection results. In this paper, the RGB color space is converted to the LO space [21] to mimic visual human perception [22].

A. Preprocessing

In the preprocessing steps, for any given image(s), S_t where t is the number of images, $t \in \{1, 2, \ldots, T\}$ we first locate human eyes as Fusel *et al.* [26]. Then, an elliptical mask model as illustrated in Fig. 3 is used to generate the elliptical face region in the image(s). Here, (x_0, y_0) is the center of the ellipse as well as the eyes symmetry point. Minor and major axes of the ellipse are represented by $1.6D$ and $1.8D$, respectively, where D is the distance between two eyes. For a more detailed description, interested readers are encouraged to read [26]. The detected face regions include smooth (i.e., skin) and nonsmooth (i.e., eyes, eye brown, mouth, etc.) textures. As we are only interested in smooth regions, Sobel edge detection is employed to remove nonsmooth regions. The choice of Sobel edge detection method is due to computational simplicity. Then, the detected edge pixels are further dilated using a dilation operation to get the optimal nonsmooth regions. Finally, we obtain a new image(s), S_t' that only consist(s) of face regions.
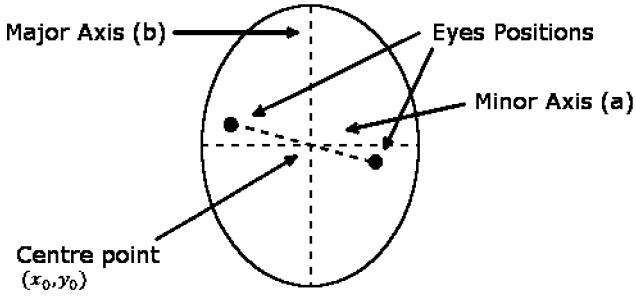
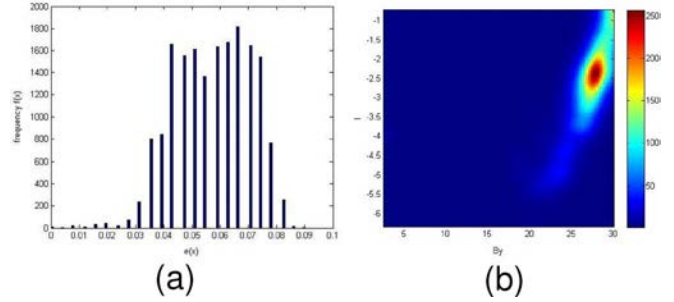Fig. 3. Elliptical mask model generated using eye coordinates.

Fig. 4. (a) and (b) Histograms of 1-D and 2-D: (a) as only one channel with frequency at y axis and (b) has two different channels of the same color space on x axis and y axis.

B. Color Space

It is well established that the distribution of colors in an image is often a useful cue. An image can be represented in a number of different color space models (i.e., RGB, HSV [1], YC_rC_b [27]). These are some color space models available in image processing. Therefore, it is important to choose the appropriate color space for modeling human skin color. In this paper, we propose the use of the LO color space [28]; the reason is twofold: first, color opponency is perceptually relevant as it has been proved that the human visual system uses an opponent color encoding [22], [23]; and second, in this color space, the use of logarithms renders illumination change to a simple translation of coordinates.

1) LO Space: The theory of opponent colors was first studied by Hering [22] in 1892. He observed that certain colors are never perceived together in the human visual system. For instance, we never see yellowish-blue or reddish-green. Based on this theory, the LO is a representation of color information by applying logarithms to the opponency model so that it is simple to model illumination changes. As illumination changes, log component chromaticity distributions undergo a simple translation. These distributions are coded by using means and first k-moments found using principle component analysis [21].

C. Skin Detection

1) Dynamic Threshold With Smoothed 2-D Histogram: Human skin color varies greatly between different ethnicity [25]. Nonetheless, skin appearance in color image(s) can also be affected by illumination, background image, camera characteristic, etc. Therefore, a fixed or prelearned threshold for detecting skin boundaries is not a feasible solution. In our approach, we employ an online dynamic approach as to [19] to calculate the skin threshold value(s) on the face images, S'_t. The assumption is that the face and body of a person always share the same colors. However, instead of using the 1-D histogram, as illustrated in Fig. 4(a); we introduce a 2-D histogram [see Fig. 4(b)] with smoothing densities [29]. In this paper, the feature vector for the smoothed 2-D histogram, Z, is represented by the combination of I and B_y. The smoothed 2-D histogram-based skin segmentation, D_{hist}, at pixel n is given as

$$D_{\text{hist}}(S_t, Z) = \begin{cases} 1 & \text{if } Z(I_n, By_n) > 20 \\ 0 & \text{if } Z(I_n, By_n) \leq 20 \end{cases} \tag{1}$$
$$\forall I_n, By_n \in S_t.$$
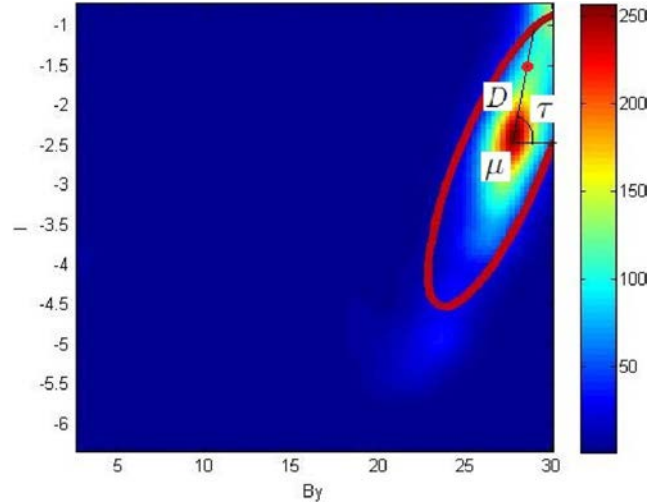
Fig. 5. Gaussian model.

2) Gaussian Model: The Gaussian model is a sophisticated model that is capable of describing complex-shaped distributions and is popular for modeling skin-color distributions. The threshold skin-color distribution in the 2-D histogram is modeled through elliptical Gaussian joint probability distribution functions defined as

$$p(H|\lambda) = \sum_{i=1}^{k} \omega_i g(c|\mu_i, \Sigma_i) \tag{2}$$

where H is the color vector of (I, B_y), $\lambda = \{\omega_i, \mu_i, \Sigma_i\}$, μ_i is the mean vector, and Σ_i is the diagonal covariance matrix, respectively. ω_i refers to the mixing weights, which satisfy the constraint $\sum_{i=1}^{k} \pi_i = 1$. The result of Gaussian model-based skin detection, D_{gmm}, can be obtained by using Fig. 5. μ is the center of the Gaussian model, while τ is the angle between x axis and line D. Let (I_i, By_i) be the coordinate of pixel n and is positioned on the red dot along line D. Distance of (I_n, By_n), d and angle τ are calculated as follows:

$$d = \sqrt{d_x^2 + d_y^2} \tag{3}$$

$$\tau = \tan^{-1}\left(\frac{d_y}{d_x}\right) \tag{4}$$

Fig. 6. Comparison using Pratheepan dataset [19]. Column from left to right represent original images, [32] method, [19] method, and our proposed methods, respectively.
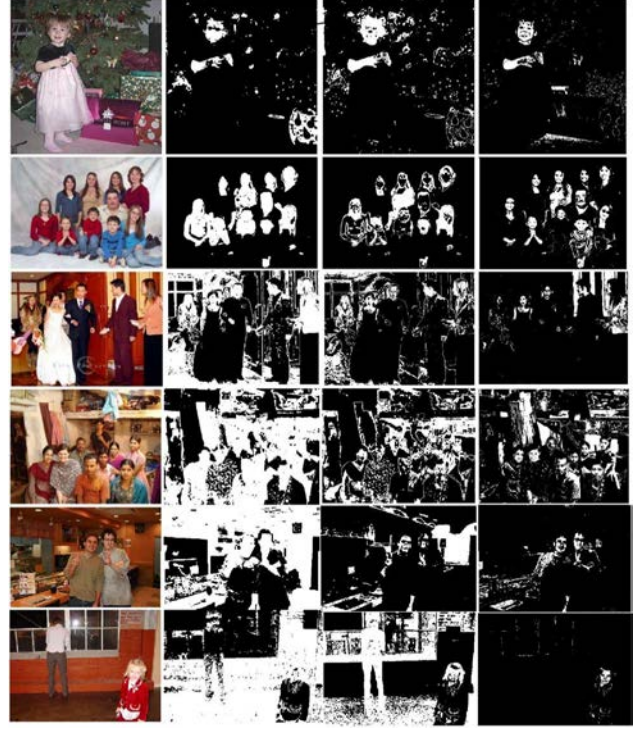

Fig. 7. Comparison using ETHZ dataset [31]. Columns from left to right represent original images, [32] method, [19] method, and our proposed method, respectively.

where d_x and d_y are the distances between (I_n, By_n) and center, μ at x axis and y axis, respectively. μ_x and μ_y are the coordinate of μ at x axis and y axis, respectively. Distance between the boundary and center of the Gaussian model at x axis and y axis, D_x and D_y at given angle, τ are as follows:

$$D_x = \Sigma_x \cos(\tau) \tag{5}$$
$$D_y = \Sigma_y \sin(\tau) \tag{6}$$

where Σ_x and Σ_y are the variance of x axis and y axis for Gaussian model. Distance D is represented as

$$D = \sqrt{D_x^2 + D_y^2}. \tag{7}$$

Therefore, D_{gmm} is given as

$$D_{gmm}(S_t, \mu, \Sigma) = \begin{cases} 1, & \text{if } D > d \\ 0. \end{cases} \tag{8}$$

D. Fusion Strategy

In order to increase the effectiveness and robustness of the skin detection algorithm, a fusion strategy is proposed by integrating the two incoming single features into a combined single representation. Both models will vote for classification of skin and nonskin pixels. This can be done by using product rule to both models. Let $D_{\text{hist}}(S_t, Z)$ and $D_{gmm}(S_t, \mu, \Sigma)$ denote the

matching results produced by the smoothed 2-D histogram Z and Gaussian model $g(c|\mu_i, \Sigma_i)$, respectively. The combined matching results $D(S_t)$ using the fusion rules can be obtained as follows:

$$D(S_t) = \Gamma\{D_{\text{hist}}(S_t, Z), D_{gmm}(S_t, \mu, \Sigma)\} \tag{9}$$

where Γ is the selected fusion rule, which represents the product $\otimes$. In order to make the fusion issue tractable, the individual features are assumed to be independent of each other.

IV. Experiments

In this section, the performance of the proposed approach under different conditions, such as fusion strategy, color spaces, and a comparison with the state-of-the-art methods in terms of qualitative and quantitative performance. We only perform quantitative analysis on the Stottinger dataset [30] as ground truth videos are only available for this dataset.

A. Datasets

Experiments are conducted using three public databases: Pratheepan's dataset [19]; ETHZ PASCAL dataset [31], and the Stottinger dataset [30]. The Pratheepan's dataset [19] consists of a set of images downloaded randomly from Google. These random images are captured with a range of different cameras using different color enhancements and under different illuminations. The ETHZ PASCAL dataset [31] consists of 549 images from PASCAL VOC 2009. The dataset consists mainly of amateur photographs with difficult illuminations and low image quality, and finally, the Stottinger dataset [30] consists

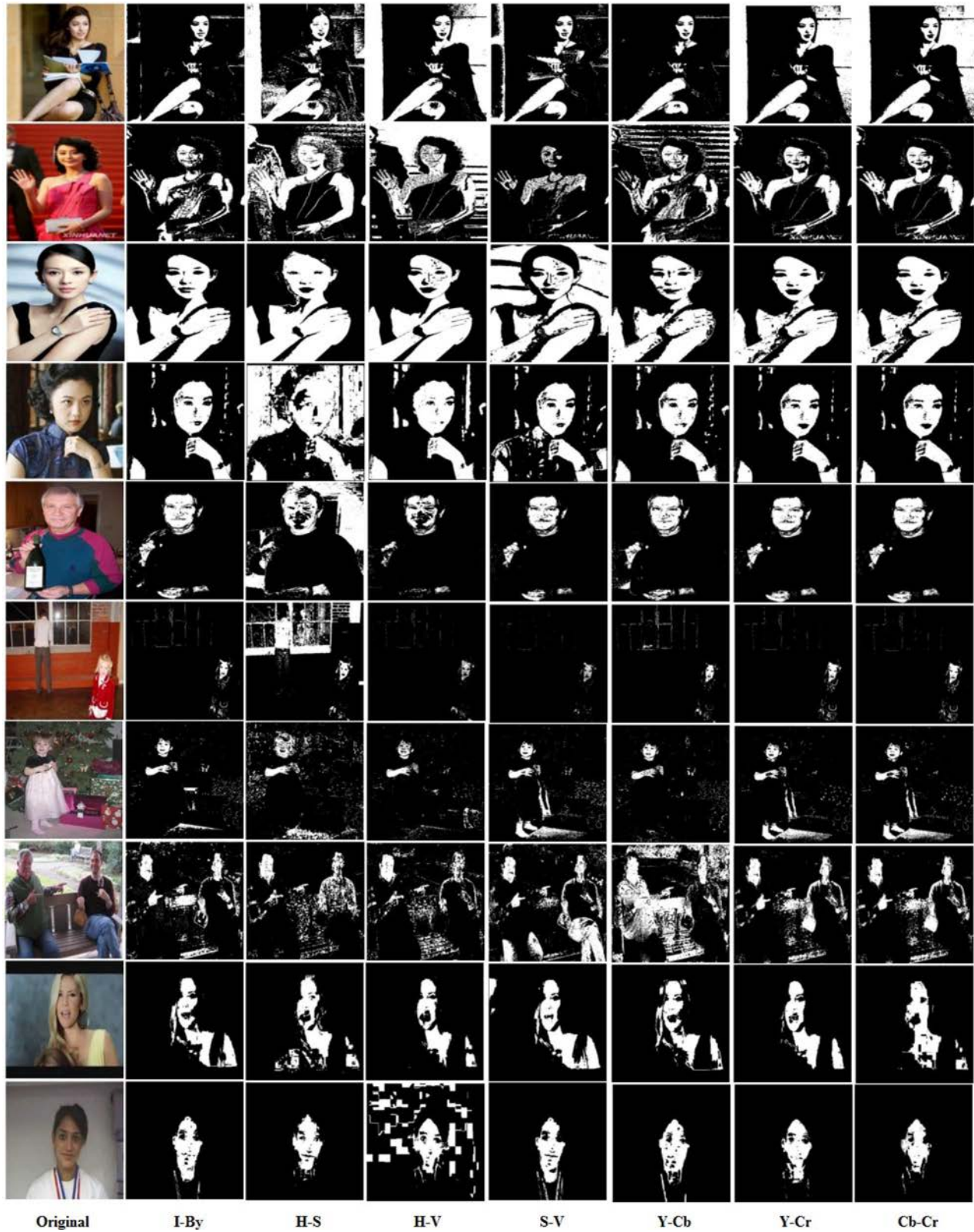Fig. 8. Comparison between results from different color space.

of popular public video clips from web platforms. These are chosen from the community (top-rated) and cover a large variety of different skin colors, illuminations, image quality, and difficulty levels.

B. Results and Analysis

The detection results for each dataset are illustrated in Figs. 6–9, respectively. When there is no face detected on
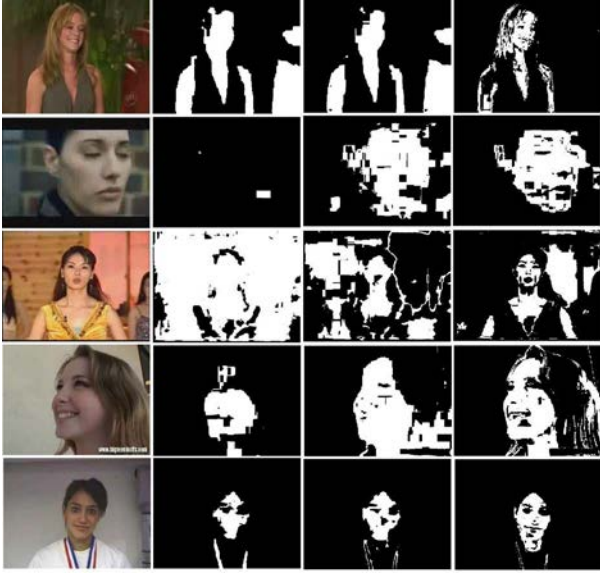
Fig. 9. Stottinger dataset [30]. Columns from left to right represent original images, [32] method, [19] method, and our proposed method, respectively.

TABLE I
COMPARISON BETWEEN DIFFERENT COLOR SPACES IN STOTTINGER DATASETS [30]

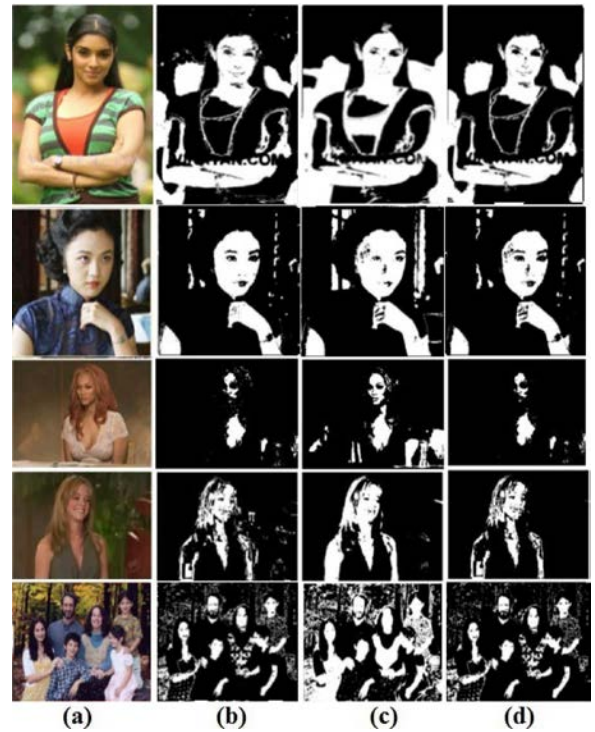| Colour Space	Accuracy	F-score	True Positive Rate	False Negative Rate
IB_y	**0.9039**	**0.6490**	**0.6580**	**0.3420**
HS	0.9057	0.6512	0.6521	0.3479
HV	0.7977	0.4549	0.6251	0.3749
SV	0.8898	0.6285	0.6905	0.3995
YC_b	0.8936	0.6143	0.6277	0.3723
YC_r	0.8985	0.6392	0.6656	0.3344
C_bC_r	0.9151	0.6241	0.5223	0.4777

Fig. 10. (a) Original Image. (b) 2-D histogram's result. (c) Gaussian model's result. (d) Fusion approach's result.

the image, it will return a blank image (black). Therefore, for testing purposes, we assumed that true face(s) are detected in the image. Conclusions are drawn as follows. First, qualitatively, the proposed method has better detection accuracy in comparison to [32] and [19]. For instance, in the image sample (row 4) of Figs. 6 and 7, respectively, both the [19] and [32] methods fail to detect the skin region accurately. One can notice from both images that the particular image sample in Fig. 7 is a fairly complex environment with high illuminations and comprises of humans from different ethnicity. However, our proposed approach is still able to capture almost all skin regions and has the least noise. Second, the proposed approach has better robustness in terms of illumination, background image, camera characteristic, and different ethnicity. The images in the Pratheepan dataset [19] are captured with a range of different cameras using different color enhancements. However, it can be qualitatively noticed that our proposed approach has the least effects on these factors. Nonetheless, in other datasets that are highly complex and illumination, our approach also achieves a better discrimination power than that in [19] and [32]. Finally, the proposed approach does not require any training stage and, hence, is more effective in terms of computational cost as opposed to approaches listed in [16].

C. Comparison Between Different Color Spaces

In this section, we analyze seven different combinations of feature vectors: $IB_y, HS, HV, SV, YC_b, YC_r$, and C_bC_r. The results for each feature vector are presented in Fig. 8 using images from Pratheepan and ETHZ datasets. It can be noticed that IB_y shows better true positive rate than the rest. Further, a better quantitative analysis result is shown in Table I using Stottinger dataset as this dataset only contains ground-truth. It can be seen from Table I that the results from IB_y and HS are comparable. However, we selected IB_y as our color space as it shows higher

true positive rate and lower false negative rate than HS. Also, it has been proven that the human visual system uses an opponent color coding.

D. Fusion Strategy Results

In this section, we show the comparison results of using single feature—smoothed 2-D histograms (s2D) or Gaussian mixture models (GMM) only, and multiple features (the fusion of s2D and GMM). The results are illustrated in Fig. 10 and Table II. Fusion Approach has the highest accuracy and F-score. Moreover, it can also be visualized that the fusion strategy has lower false positive rate compared to the single feature approach. For instance, the smoothed 2-D histogram is able to detect most of the skin regions, but it is highly occluded with noise. The

TABLE II
COMPARISON BETWEEN OUR PROPOSED FUSION AND NONFUSION APPROACH
IN IB_y COLOR SPACE USING STOTTINGER DATASET [30]

Classifier	Accuracy	F-score	True Positive Rate	False Positive Rate
Fusion	**0.9039**	**0.6490**	**0.6580**	**0.0577**
s2D Histogram	0.8930	0.6270	0.6662	0.0716
GMM	0.8595	0.6150	0.8314	0.1361
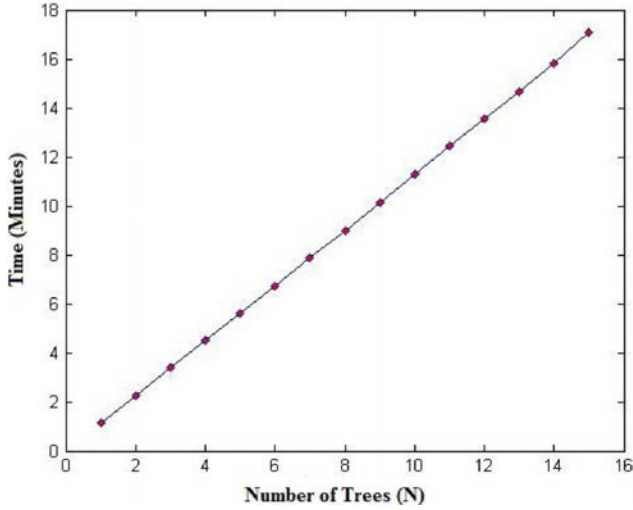
Fig. 11. Plot of time in minutes taken for 15 trees for offline training.

Fig. 12. Comparisons between random forest and our method using Pratheepan dataset [19] and ETHZ dataset [31]. Row from top to bottom represent original image, result from random forest, and result from our method, respectively.

TABLE III
COMPARISON BETWEEN CLASSIFIER PERFORMANCE IN STOTTINGER DATASET
[30] USING IB_y COLOR SPACE

Classifier	Accuracy	F-score	Precision	Recall
Random Forest [16]	0.9305	0.7307	0.7661	0.6984
Our Proposed Method	**0.9039**	**0.6490**	**0.6403**	**0.6580**
Dynamic Threshold [19]	0.8935	0.5922	0.6133	0.5725
Static Threshold [32]	0.8334	0.4745	0.4133	0.5570

Fig. 13. Column from left to right represent original images, skin detection using [19], [32], and our proposed method, respectively.

Gaussian model in the meantime has very high false positive rate (i.e., clothing which is classified as skin region).

E. Quantitative Analysis

Table III shows comparisons between our method and other state-of-the-art methods on the Stottinger dataset [30]. Apart from random forest [16], other methods do not include training. A total of 2985 frames were extracted from seven videos for our experiment. For random forest [16], 1990 image frames are randomly chosen for training and remaining images are used for testing. From those 1990 images, around 3 million pixels were randomly chosen and 15 tress were trained. Each tree extracts 70% of the pixels, respectively, for training. Fig. 11 shows the time in minutes required for 15 trees to be trained in random forest solutions. It took around 17 min for 15 trees to be trained. Fig. 12 shows the comparisons between random forest and our method using the Pratheepan dataset [19]. Here, random forest [16] is trained with Stottinger dataset and tested with Pratheepan dataset [19]. It shows that random forest does not work well. In order to increase the accuracy of random forest on this dataset, huge training samples and/or more trees will be needed. This will cause higher computational power as number trees increases and time consuming during training. Our method is able to maintain quality on skin segmentation. A quantitative analysis is shown in Table III using Stottinger

dataset. The accuracy is a score that uses the sum of true positive and true negative as measurement. While, *F-score* is the product of *Precision* and *Recall*. Table III shows that our proposed method has an acceptable score as compared to random forest [16], even no training stage is required. Nonetheless, it is able to cope with large variable of illumination and complex backgrounds variations.

F. Discussions

In our experiments, we noticed that the final result of our work depends greatly on the outcome of the preprocessing phase (the eye detector [26]). If the algorithm detects a false face region, poor result will be returned. Fig. 13 shows the result of skin segmentation for false face region detected. When a false face region is obtained, false dynamic thresholds will be generated. Therefore, false classifications will be processed, where non-skin regions are classified as skin regions. In our future work,

we will investigate the face detector algorithm to overcome this problem.

V. CONCLUDING REMARKS

In this paper, a fusion framework based on smoothed 2-D histogram and Gaussian model has been proposed to automatic detect human skin in image(s). As exhibited in experiments, the proposed method outperforms state-of-the-art methods in terms of accuracy in different conditions: background model, illumination, and ethnicity. With this, it shows the potential to be applied to a range of applications such as gesture analysis. One drawback of the proposed approach is that its success relies on eye detector algorithms. However, this is the general problem faced by all other researchers who work in this domain. Our future work is focused on building a better preprocessing method, to use field-programmable gate arrays to implement a hardware scheme.

REFERENCES

[1] P. Vadakkepat, P. Lim, L. De Silva, L. Jing, and L. L. Ling, "Multimodal approach to human-face detection and tracking," *IEEE Trans. Ind. Electron.*, vol. 55, no. 3, pp. 1385–1393, Mar. 2008.

[2] C. S. Chan, H. Liu, and D. J. Brown, "Recognition of human motion from qualitative normalised templates," *J. Intell. Robot. Syst.*, vol. 48, no. 1, pp. 79–95, 2007.

[3] N. Kubota and K. Nishida, "Perceptual control based on prediction for natural communication of a partner robot," *IEEE Trans. Ind. Electron.*, vol. 54, no. 2, pp. 866–877, Apr. 2007.

[4] O. Linda and M. Manic, "Fuzzy force-feedback augmentation for manual control of multi-robot system," *IEEE Trans. Ind. Electron.*, vol. 58, no. 8, pp. 3213–3220, Aug. 2010.

[5] C. S. Chan, H. Liu, and D. J. Brown, "Human arm-motion classification using qualitative normalized templates," *Lecture Notes Artif. Intell.*, vol. 4251, no. Part I, pp. 639–646, 2006.

[6] G. Pratl, D. Dietrich, G. P. Hancke, and W. T. Penzhorn, "A new model for autonomous, networked control systems," *IEEE Trans. Ind. Informat.*, vol. 3, no. 1, pp. 21–32, Feb. 2007.

[7] K. Sobottka and I. Pitas, "A novel method for automatic face segmentation, facial feature extraction and tracking," *Signal Process.: Image Commun.*, vol. 12, no. 3, pp. 263–281, 1998.

[8] H. Bae, S. Kim, B. Wang, M. H. Lee, and F. Harashima, "Flame detection for the steam boiler using neural networks and image information in the ulsan steam power generation plant," *IEEE Trans. Ind. Electron.*, vol. 53, no. 1, pp. 338–348, Feb. 2005.

[9] Y. Wang and B. Yuan, "A novel approach for human face detection from color images under complex background," *Pattern Recognit.*, vol. 34, no. 10, pp. 1983–1992, 2001.

[10] D. Brown, I. Craw, and J. Lewthwaite, "A SOM based approach to skin detection with application in real time systems," in *Proc. Brit. Mach. Vis. Conf.*, 2001, pp. 491–500.

[11] M.-J. Seow, D. Valaparla, and V. K. Asari, "Neural network based skin color model for face detection," in *Proc. Appl. Image Pattern Recognit. Workshop*, 2003, pp. 141–145.

[12] S. L. Phung, D. Chai, and A. Bouzerdoum, "A universal and robust human skin colour model using neural network," in *Proc. Int. Joint Conf. Neural Netw.*, 2001, vol. 4, pp. 2844–2849.

[13] N. Sebe, I. Cohen, T. S. Huang, and T. Gevers, "Skin detection: A Bayesian network approach," in *Proc. Int. Conf. Pattern Recognit.*, 2004, pp. 903–906.

[14] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Mach. Learn.*, vol. 29, pp. 131–163, Nov. 1997.

[15] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," *Int. J. Comput. Vision*, vol. 46, no. 1, pp. 81–96, 2002.

[16] R. Khan, A. Hanbury, and J. Stoettinger, "Skin detection: A random forest approach," in *Proc. Int. Conf. Image Process.*, Hong Kong, 2010, pp. 4613–4616.

[17] U. Yang, B. Kim, and K. Sohn, "Illumination invariant skin color segmentation," in *Proc. 4th IEEE Int. Conf. Ind. Electron. Appl.*, May 2009, pp. 636–641.

[18] S. Jayaram, S. Schmugge, M. C. Shin, and L. V. Tsap, "Effect of colorspace transformation, the illuminance compoment, and color modeling on skin detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2004, vol. 2, pp. 813–818.

[19] P. Yogarajah, A. Cheddad, J. Condell, K. Curran, and P. McKevitt, "A dynamic threshold approach for skin segmentation in color images," in *Proc. Int. Conf. Image Process.*, 2010, pp. 2225–2228.

[20] P. Kakumanua, S. Makrogiannisa, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognit.*, vol. 40, no. 3, pp. 1106–1122, 2007.

[21] J. Berens and G. Finlayson, "Log-opponent chromaticity coding of colour space," in *Proc. Int. Conf. Pattern Recognit.*, Barcelona, Spain, 2000, vol. 1, pp. 206–211.

[22] E. Hering, *Outlines of a Theory of the Light Sense*. Cambridge, MA: Havard Univ. Press, 1964.

[23] L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision," *Psychol. Rev.*, vol. 64, pp. 384–404, Nov. 1957.

[24] S. Mitra and T. Acharya, "Gesture recognition: A survey," *IEEE Trans. Syst., Man, Cybern., C: Appl. Rev.*, vol. 37, no. 3, pp. 311–324, May 2007.

[25] A. M. Elgammal, C. Muang, and D. Hu, "Skin detection," in *Encyclopedia of Biometrics*. Germany, Berlin: Springer, 2009, pp. 1218–1224.

[26] I. Fasel, B. Fortenberry, and J. Movellan, "A generative framework for real time object detection and classification," *Comput. Vis. Image Underst.*, vol. 98, pp. 182–210, Apr. 2005.

[27] C. Kumar and A. Bindu, "An efficient skin illumination compensation model for efficient face detection," in *Proc. 32nd IEEE Annu. Conf. Ind. Electron.*, 2006, pp. 3444–3449.

[28] D. A. Forsyth and M. M. Fleck, "Automatic detection of human nudes," *Int. J. Comput. Vis.*, vol. 32, pp. 63–77, Aug. 1999.

[29] P. H. Eilers and J. J. Goeman, "Enhancing scatterplots with smoothed densities," *Bioinformatics*, vol. 20, no. 5, pp. 623–628, 2004.

[30] J. Stottinger, A. Hanbury, C. Liensberger, and R. Khan, "Skin paths for contextual flagging adult video," in *Proc. Int. Symp. Visual Comput.*, 2009, pp. 903–906.

[31] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge 2009 (VOC2009)," 2009.

[32] A. Cheddad, J. Condell, K. Curran, and P. McKevitt, "A skin tone detection algorithm for an adaptive approach to steganography," *J. Signal Process.*, vol. 89, pp. 2465–2478, 2009.

Wei Ren Tan received the B.Sc. (Hons) degree in computer science, majoring in artificial intelligence from the University of Malaya, Kuala Lumpur, Malaysia, in 2010, where he is currently working towards the M.S. degree at the Faculty of Computer Science and Information Technology.

He previously visited the University of Ulster (Magee), Northern Ireland, U.K., as a Visiting Researcher. His research interests include computer vision, image processing, and evolutionary computation.

Chee Seng Chan (S'05–M'09) received the Ph.D. degree from the University of Portsmouth, Portsmouth, U.K., in 2008.

He is currently a Senior Lecturer at the Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia. He previously held appointments in the University of Ulster and the University of Portsmouth. His research interests include fuzzy qualitative reasoning, computer vision, machine learning, cognitive science, in particular, human motion recognition in images and video, biomedical image analysis, compositionality, and graphical models.

Dr. Chan was the recipient of 2010 IET (Malaysia Chapter) Young Engineer Award.

Pratheepan Yogarajah received the B.Sc. (Hons) degree in computer science from the University of Jaffna, Jaffna, Sri Lanka, in 2001, and the M.Phil. degree in computer vision from the Oxford Brookes University, Oxford, U.K., in 2006.

He is currently a Research Associate at the School of Computing and Intelligent Systems, University of Ulster (Magee), Northern Ireland. His research interests include biometrics, computer vision, image processing, steganography and digital watermarking, and machine learning.

Joan Condell received the Ph.D. degree in computing and math from the University of Ulster, Northern Ireland, in 2002, with the research theme in motion tracking in digital images.

She is currently a Lecturer at the University of Ulster (Magee), Northern Ireland. Her achievements include winning and managing European Framework projects and commercialization projects. She has patents filed and has published more than 130 papers in international conferences, books, and journals. Her research interests include steganography, image processing, biometrics and vision for robotics, and multimedia.