British Journal of Mathematics & Computer Science
17(6): 1-14, 2016, Article no.BJMCS.26412
ISSN: 2231-0851

SCIENCEDOMAIN *international*
www.sciencedomain.org

A Hybrid Cascade Approach for Human Skin Segmentation

Samy Bakheet[1,2*] and Ayoub Al-Hamadi[2]

[1] *Department of Mathematics & Computer Science, Faculty of Science, Sohag University, P.O.Box 82524 Sohag, Egypt.*

[2] *Institute for Information Technology and Communications, Otto-von-Guericke-University Magdeburg, P.O.Box 4120, 39016 Magdeburg, Germany.*

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2016/26412

Editor(s):

(1) Victor Carvalho, Polytechnic Institute of Cvado and Ave, Portuguese Catholic University and Lusiada University, Portugal.

(2) Paul Bracken, Department of Mathematics, The University of Texas-Pan American, Edinburg, TX 78539, USA.

Reviewers:

(1) Chuanxu Wang, Qingdao University of Science and Technology, China.

(2) Anonymous, Tanta University, Egypt.

(3) Darabant Sergiu Adrian, Babes Bolyai University, Romania.

(4) Luis Enrique Colmenares Guillen, Benemerita Universidad Autonoma de Puebla, Mexico.

(5) Nirvair Neeru, Punjabi University Patiala, Punjab, India.

Complete Peer review History: http://sciencedomain.org/review-history/15572

Received: 16th April 2016

Accepted: 15th July 2016

Published: 29th July 2016

Original Research Article

Abstract

Human skin segmentation is fundamental to a wide range of computer vision applications ranging from face recognition, facial expression recognition and gesture analysis to various human computer interaction domains. In this paper, we propose a multistage skin segmentation method built as a cascade of a nonparametric generic model and an adaBoost classifier. Several entities are used to train the adaBoost classifier. Feature vectors fed into the ada Boost contain color information from two different color spaces. Extensive experiments are conducted on two datasets in order to evaluate the performance of the approach. The various results obtained show that the proposed method is a promising approach and it successfully achieves high quality segmentation, while concurrently retaining reasonably low false alarm rates. The comparison of the proposed method with related state-of-the-art competitors reveals the superiority and effectiveness of the proposed method, while maintaining real-time performance.

**Corresponding author: E-mail: samy.bakheet@gmail.com*

Keywords: Skin segmentation; adaBoost; gesture detection; pattern recognition.

2010 Mathematics Subject Classification: 53C25, 83C05, 57N16.

1 Introduction

Over the course of the last three decades or so, skin segmentation has been intensively investigated and many improvements have been achieved. However, most of the current state-of-the-art approaches for skin segmentation/detection are still in early stages of acceptable performance or suboptimal in performance [1]. Furthermore, due to several significant barriers and challenges that should be navigated to develop practical solutions for efficient and effective skin segmentation, many of these approaches often fail to find their way into applications in real-time recognition systems [2]. In the contexts of pattern recognition and computer vision, human skin detection is a very popular and useful technique for detecting and tracking human-body parts (e.g., face, hand, etc.) and it receives extraordinary attention, as a primary preprocessing step, in a wide range of computer vision applications such as, face detection, tracking and recognition [3, 4], hand gesture recognition [5, 6], adult image content filtering [7, 8], and videophone [9, 10].

The crucial assumption on which the pivotal idea of most existing approaches for skin detection relies is that human skin color is quite different from other objects' color in background in most situations, and its distribution tends to constitute a distinct cluster in some specific color spaces. Broadly speaking, almost all research studies in human skin detection seem to be propagated in two different directions: statistical color modeling [3, 7, 11, 12] and generic color modeling [3, 9, 13]. For the first approach, i.e., the generic color model, the fundamental idea behind this approach is to define a fixed color range to isolate skin pixels from non-skin pixels. Generally, the fixed color range is accomplished empirically utilizing a few collected training samples. Moreover, such an approach can be implemented very efficiently and thus is more amenable for integration in real-time applications. However, adopting generic color model by itself cannot apparently provide sufficient guarantees to cope with serious problems such as, human skin variations and illumination changes [14, 15, 16]. Seeking to reduce these inherent limitations, the so-called statistical color model has emerged with the aim of developing a more robust model for skin color. In this approach, the skin segmentation problem is formulated as a traditional classification task, and both positive (skin pixels) and negative (non-skin pixels) instances are fed to a machine learning (ML) classifier. Briefly, within the statistical model, the processes proceed as follows. First, a histogram-based technique or Gaussian model is applied on large amount of training data to estimate skin color distribution. Then, a learning classifier (e.g., Bayesian) is used to distinguish the skin pixels from non-skin pixels. To ensure that the entire segmentation model is quite robust with respect to changes in lighting conditions and a wide range of human skin tones, it is recommended that the classifiers are adaptively designed.

It is noteworthy to mention that the statistical color model exhibits the following two drawbacks. Firstly, the overall classification accuracy of the model relies heavily on the accuracy of the supervised learning module that is, in turn, much dependant on the sensitivity and the quantity of the training data (i.e., color pathes) that should be collected in such a way as to be properly representative of a wide range of human skin tones. The other drawback arises from an inherent property of the model itself; as the basic principle of operation relies on assumption that the skin color distribution is a single or mixture Gaussian. Unfortunately, such an assumption often does not hold in practice,which leads to a decay in the segmentation accuracy. The remaining of the paper unfolds as follows. Section 2 reviews few related studies. In section 3, the proposed skin segmentation method is described in detail. Exhaustive experiments and analysis of results are presented in section 4. Finally, in section 5, we summarize the paper and suggest some possible future works.

2 Related Work

Human skin segmentation that aims at finding skin-color pixels or regions in a given image is recognized as a key indispensable component in a large number of applications in diverse disciplines such as facial recognition, gesture analysis, human tracking, adult image filtering, etc. In [17], the authors make use of skin detection approaches for biometric identification applications. In [18], skin detection is used in a cosmetic surgery application for detecting facial wrinkles and fixes to affected areas of face. In this work, the authors introduce an algorithm for wrinkles curves and pattern matching for the purpose of facial recognition. Furthermore, in [19], a technique for skin detection is proposed to detect human body parts and identify gestures. In literature, skin detection methods can be further categorized as region-based methods and pixel-based methods. For region based methods, they are normally established on the basis of region based approach. Therefore, the spatial arrangement of neighboring pixels is often considered to improve performance of skin pixel detection. Moreover, in this approach, it is necessary to make use of other extra information such as texture [20, 21]. On the other hand, pixel based methods follow an alternative approach where each pixel is solely categorized as a skin pixel or non-skin pixel apart from its neighboring pixels. In this context, skin detection is typically treated as a standard binary classification task where the input is a color vector and the output is a pallet of skin and non-skin pixel. Consequently, skin detection task belongs to the category of pattern recognition.

In the existing literature, there are three major approaches for pattern recognition, including statistical, neural, and symbolical methods. Among these three approaches, the statistical approach is the most versatile, as it has the potential to properly model skin color, and it is adopted more frequently in skin detection [20]. In many cases, the pixel based methods have proved to have the potential to provide the accuracy and computational efficiency, and consistency required for skin detection task. In skin detection, one of the most significant challenges is how to choose the most appropriate color space for skin detection. In other words, the selection of the most appropriate color space seems to be a very daunting task. To the best of our knowledge, Littman and Ritter [22] have, for the first time, explored the use of different color spaces and their impacts on the performance of skin detection task. In their work, a neural approach based on linear maps for skin color with normal distributions is proposed, which is implemented using three different color spaces (i.e. RGB, YIQ and YUV) on a relatively small dataset of hand images. In summary, the authors concluded that the neural approach produces most accurate segmentation results and its performance is quite robust and, most notably, quasi-independent of the color representation.

In [23], Albiol et al. proved that given an invertible transformation between color spaces, an optimum skin detection with superior performance can be achieved in each color space. In their work, Bayes classifier is trained on a database of 200 images in three color spaces (i.e., RGB, YCrCb, and HSV), and their performances are found to be in close agreement. Moreover, the authors have concluded that the detection performance in the 3D YCrCb color space is better than that of the 2D CrCb color space, since the transformation form the 3D space to the 2D space is not invertible.

3 Proposed Approach

In this section, we first provide a basic outline of the proposed approach for skin segmentation and then discuss each component and their interactions in more detail. In Fig. 1, a flow diagram of the proposed approach is shown. This approach is essentially based on segmentation of the normalized RGB (normRGB) and HSV color spaces which follow a human intuition of color classification. In comparison to different color spaces features, we believe that we, in this work, choose proper regular color spaces for skin segmentation. The following normalized color representation space [24] is used.

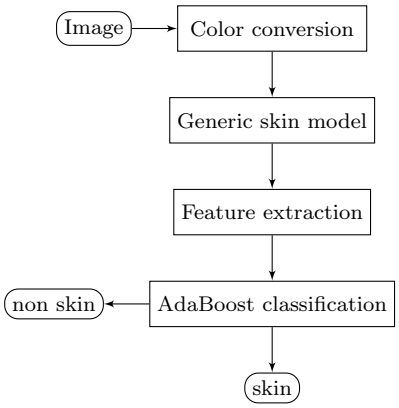


Fig. 1. Overall workflow of the proposed approach

$$\begin{aligned} y &= c_1 r + c_2 g + c_3 b \\ \tau_1 &= \tfrac{r}{r+g+b} \\ \tau_2 &= \tfrac{g}{r+g+b} \end{aligned} \tag{3.1}$$

where r, g, and b stands for RGB channels with values from 0 to 255, and $c_1 + c_2 + c_3 = 1$ and c_1, c_2, and c_3 are constants that can united to yield the image illumination. We can readily conceive that as τ_1 and τ_2 are determined only by the percentage of the RGB components, they can describe the color information of the image pixels. In addition to that, they are independent of the image brightness which is regularly cited as one of the primary advantages of the normRGB space. Moreover, the normalization process not only turns out to be more robust to the change of the illumination, but also can diminish the sensitivity of the distribution to the color variability.

The other color space that the presented method adopts is HSV color space. The main reason behind our choice of utilizing the HSV color representation is based on the fact that in this space the color information is elegantly separated from the image intensity. In the HSV color space, color is split into three perceptual components: Hue, Saturation, and Value. Geometrically, these components form a cone, as shown in Fig. 2. In order to convert RGB color space to HSV color space, we proceed as follows. In case of an 8-bit image, let $r, g, b \in [0, 255]$ be color values of image pixels in RGB space, first r, g, and b are converted to the floating-point format and scaled to fit to $[0, 1]$. Then, the following transformation is applied.

$$\begin{aligned} v &= \max(r, g, b) \\ s &= \begin{cases} \frac{\alpha}{v}, & \text{if } v \neq 0 \\ 0, & \text{otherwise} \end{cases} \\ h &= \begin{cases} 60(g-b)/\alpha, & \text{if } v = r \\ 120 + 60(b-r)/\alpha, & \text{if } v = g \\ 240 + 60(r-g)/\alpha, & \text{if } v = b \end{cases} \end{aligned} \tag{3.2}$$

where $\alpha = \max(r, g, b) - \min(r, g, b)$ and h, s, and v are the HSV values of image pixels. If $h < 0$, then $h \leftarrow h + 360$. Therefore, the HSV values satisfy the inequalities: $0 \leq v, s \leq 1$ and $1 \leq h \leq 360$. For visualization purposes, the values are converted to destination data type (i.e., 8-bit images), as follows: $h \leftarrow h/2, s \leftarrow 255s, v \leftarrow 255v$ to fit to $[0, 255]$. Unlike several other methods, in our method we opt to exploit all the three channels (i.e., hue, saturation, and value) of the HSV color space for segmenting skin color. The main objective behind doing this is to take advantage of the color information of the extra channel to improve overall color segmentation quality.

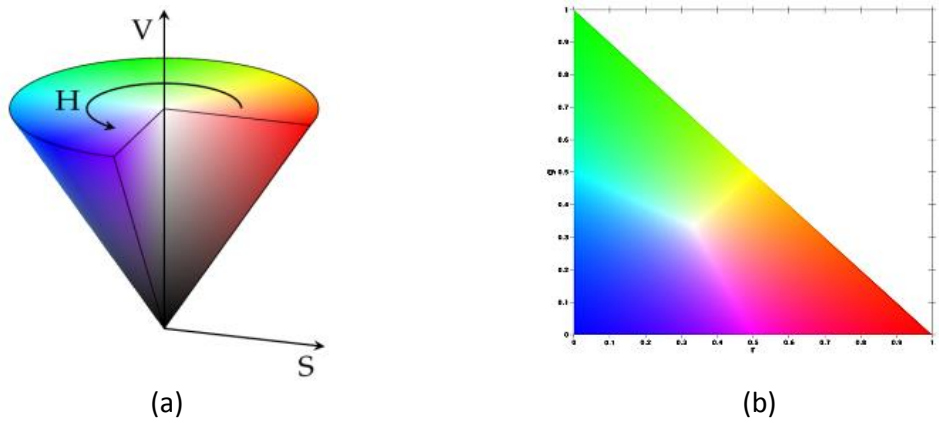


Fig. 2. Color representation: (a) HSV and (b) normalized RG color models [25]

3.1 The generic skin model

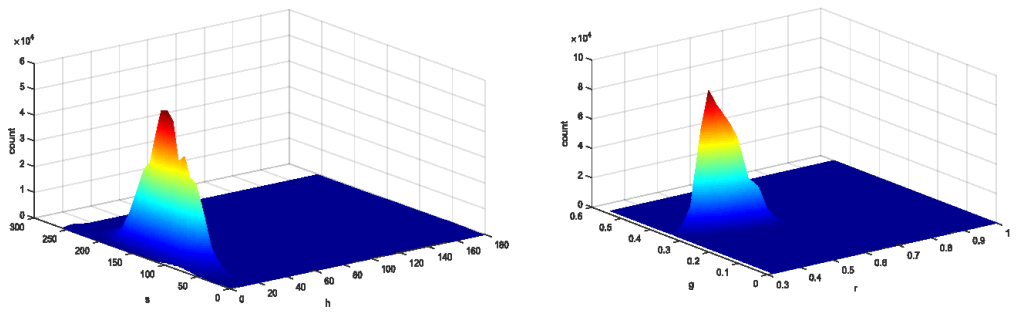
For more efficient and robust skin segmentation, a multistage skin detection method built as a cascade of a nonparametric generic model and an adaBoost classifier is introduced. In the context of our approach, the primary objective of using a generic skin model in the training stage is solely confined to the process of collecting a set of training instances to train the adaBoost classifier. To construct the generic skin model, unlike [26], we find it more appropriate to define a fixed skin color range in two color spaces (i.e., normRGB and HSV) to represent the skin pixels as follows,

1. In normRGB space:
   $r \in [0.36, 0.465] \wedge g \in [0.28, 0.363]$

2. In HSV space:
   $[0, 58, 88] \leq (h, s, v) \leq [25, 173, 229]$

where the symbol $\wedge$ denotes a bitwise AND operator. The two 3D histogram distributions of skin colors that are created from the above generic model in the HS and normalized RG color spaces are shown in Fig. 3. As can be seen in the figures, a best fit to the histograms distributions can be achieved with Gaussian approximations. It is worth mentioning here that in the normRGB space the third component does not hold any significant information and is normally dropped to obtain a reduction in dimensionality, as each of the red, green and blue pixel values are divided by the sum of the RGB pixel values, so that the sum normalizes to one (i.e., $r + g + b = 1$). Moreover, as stated earlier, the color representation in this space is invariant with respect to illumination direction and intensity, as well as to the viewing direction and surface orientation.

As for our particular choice of the HSV (Hue, Saturation, and Value) color space in this color detection task, the reason is mainly attributed to the relative advantages of the HSV over the common RGB color space, as HSV space appears more intuitive about colors in terms of brightness and spectral names rather than the mixture coefficients of R, G and B. Additionally, in HSV only H value can determine various color information, whereas in RGB combination of R, G and B is used to determine particular color.



(a) (b)

Fig. 3. 3D histogram distributions of skin color: (a) HS space, (b) normRG space

3.2 AdaBoost learning for skin segmentation

In this work, we approach the task of human skin segmentation by casting it as a typical binary classification problem where the skin and non skin spectra comprise positive and negative classes. Since this task is formulated as a supervised classification problem, there are a variety of supervised classification models available in machine learning such as, Naive Bayes, Artificial Neural Network (ANN), Support Vector Machines (SVMs), k-Nearest Neighbor (kNN), etc. that can be employed for skin segmentation. To accomplish the current classification task, we opt for an adaBoost algorithm [27] that achieves much higher classification rate and it meets real-time requirements.

As a machine learning algorithm, the boosting algorithm is deeply rooted in the success of several computer vision tasks such as, object detection and pattern classification, image segmentation, and data mining. AdaBoost (short for adaptive boosting) has been developed as an improvement for the original boosting algorithm. AdaBoost is categorized as an adaptive machine-learning meta-algorithm as it combines various machine learning algorithms to perform classification. This algorithm has a capability to upgrade so called "weak learners" to a strong learner. The terms 'weak learner' and 'strong learner' are originally derived from the Probably Approximately Correct (PAC) model, and they have technical definitions. In general, the classifier is said to be a strong learner if it is well correlated with true labels, while a weak learner is slightly correlated with the labels (it can label examples better than random guessing).

The major difference between the two algorithms is that whilst in the boosting algorithm it is necessary to know in advance the lower error rate, adaBoost is able to adaptively adjusts the lower error rate according to the results of the weak learning feedback. Moreover, the adaBoost algorithm does not have to know the details of weak learning to achieve the same efficiency as Boosting [28, 29]. As stated earlier, adaBoost algorithm is a direct outgrowth or an epimorphic extension of the regular boosting algorithm, which can improve stage by stage overall accuracy, regardless of what the type of the data is. It adaptively selects the 'best' features in each step and combine a series of weak classifiers into a strong classifier. In adaBoost learning algorithm, an equal weight is initially assigned to each training sample.

In this work, the classification module relies on the weak learning assumption that the weak learners can consistently find weak classifiers slightly different from random (rules of thumb that classify the data correctly at better than 50%). Strictly speaking, assume the so-called Weak Learning Assumption holds, then boosting can be used to generate a single weighted classifier that correctly classifies training data at a very conservative rate of 1% false negatives. AdaBoost can optimally weight training instances by Focusing on difficult data points that have been misclassified

most by the previous weak classifier. Finally, based on an optimally weighted majority vote of weak classifiers, adaBoost combines the weak classifiers into a comprehensive prediction. More formally, adaBoost learning algorithm can be concisely formulated as follows. Given the training data $\mathcal{D} = \{(x_i, y_i)\}_{i=0}^{n-1}$, $x_i \in \mathbb{R}^d, y_i \in \{-1, 1\}$, adaBoost can then use an ensemble of weak classifiers (hypotheses) to build a more expressive one (i.e., the final hypothesis $H(x)$), by going through the steps of the following algorithm:

1: **procedure** ADABOOST;
2: $\mathcal{D} \leftarrow \{(\mathbf{x}_i, y_i)\}_{i=0}^{n-1}$
3: $\tilde{W}_0[0 : n-1] \leftarrow \frac{1}{n}$
4: **for** $t = 0 \rightarrow \tau - 1$ **do**
5: $h_t \leftarrow \arg\max_{h \in \mathcal{H}} |0.5 - \epsilon(h, \tilde{W}_t)|$
6: $\alpha_t \leftarrow \frac{1}{2} \ln(\frac{1-\epsilon(h_t, \tilde{W}_t)}{\epsilon(h_t, \tilde{W}_t)})$
7: $\tilde{W}_{t+1}[0 : n-1] \leftarrow \frac{1}{Z} \tilde{W}_t[0 : n-1] \, \mathrm{e}^{-\alpha_t \lambda}$
8: $\lambda \leftarrow y[0 : n-1] h_t(\mathbf{x}[0 : n-1])$
9: **end for**
10: **return** $\{\alpha_t, h_t\}$.
11: **end procedure**

In the above algorithm, Z is the normalization factor of $\tilde{W}_{t+1}$, τ is the maximum number of weak-classifiers in the ensemble. At each iteration step, the goal is to find any classifier $h(\mathbf{x})$ for which the weighted classification error, defined as follows,

$$\epsilon(h, \tilde{W}) = 0.5 - \frac{1}{2} \sum_{i=0}^{n-1} \tilde{W}[i] y_i h(\mathbf{x}_i)$$

is better than chance. The weights are updated at Step 7 (where Z is chosen so that the new weights sum to one). The final classifier is then given by:

$$H(\mathbf{x}) = \text{sign}\left(\sum_{t=0}^{\tau-1} \alpha_t h_t(\mathbf{x})\right)$$

At this point, we ought to point out that after each boosting iteration, assuming we can find a component classifier whose weighted error is better than chance, then the combined classifier is guaranteed to have a lower exponential loss over the training examples, as depicted in Fig. 4. Before moving forward into the next section of 'Experiments and Results', it is worthwhile to show how the proposed method is evaluated. The evaluation protocol we adopt for the proposed method consists of assessing the obtained results in terms of the detection rate (so called recall) and false alarm rate. In other words, the segmentation performance of our approach is characterized in terms of detection rate (or recall), τ and false alarm rate, ε. In the context of this segmentation task, τ corresponds to the number of correctly classified skin pixels with respect to the total number of skin pixels in ground truth data, while ε is the number of non skin pixels incorrectly classified as skin pixels with respect to the total number of non skin pixels. More formally, the values of τ and ε are defined as follows,

$$\tau = \frac{\text{TP}}{\text{TP+FN}}, \qquad \varepsilon = \frac{\text{FP}}{\text{TN + FP}} \tag{3.3}$$

where TP measures the number of skin pixels correctly identified, while FN the number of skin pixels identified as non skin pixels. On the other hand, FP is the number of false alarms; which is non skin pixels that are incorrectly classified as skin pixels, whereas TN indicates the number of the non-skin pixels correctly classified as non-skin pixels. Note that the evaluation of the true positive rate TP and false positive rate FP is more common. Furthermore, from the above formulas in Eq. (3.3), one can generally confirm that the higher the τ and the lower the ε, the more accurate the method.

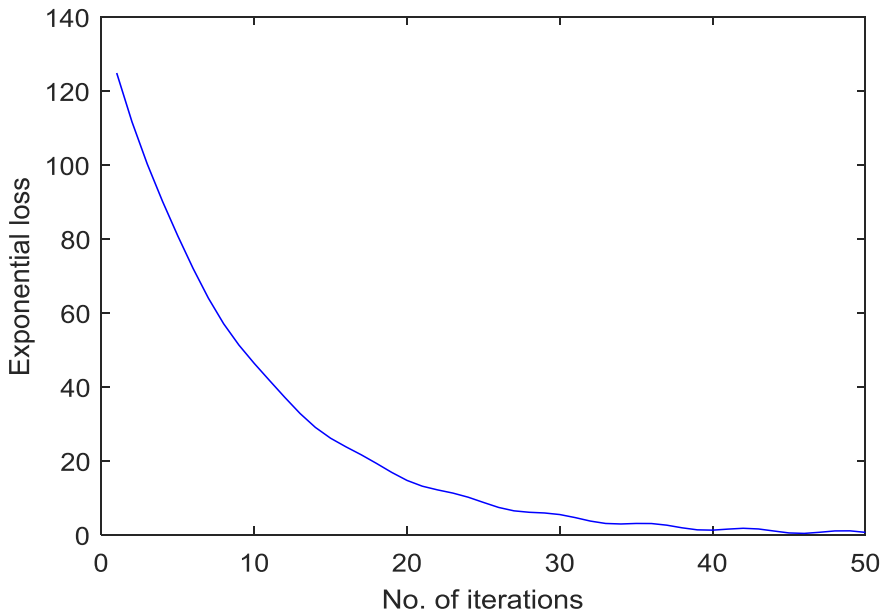


Fig. 4. A lower exponential loss over the training samples

4 Experiments and Discussion of Results

In this section, the performance of the proposed approach is evaluated under different conditions such as scale change, occlusion, orientation change, and color spaces, along with a comparison with the state-of-the-art methods. To verify the performance of the proposed method on an independent validation dataset, extensive experiments have been carried out on several datasets from various sources and domains. The first is our own constructed gesture dataset which is composed of a maximum of 300 images randomly downloaded from the internet by using keyword-based Google image search. Most of these images are taken from the political domain resources, where the gestures are normally performed by several political chief executives (e.g., presidents and prime ministers, etc.) of different countries in the world. A sample comprised out of 30 image is randomly selected from the dataset and manually segmented to build the ground truth.

In order to allow the proposed method to be comparable to other contemporary relevant methods, we evaluate our experiments with other datasets to determine the effectiveness and adequacy of the presented approach. To this end, some illustrative images are randomly taken from Stottinger[30] and Pratheepan datasets [31] that contain a large collection of images captured with different-angle cameras under illumination conditions. In our experiments, in order to build a training dataset, around 400,000 pixels for skin and 550,000 pixels for non-skin pixels have been extracted from images randomly downloaded from the web by using Google image search. Since the approach is developed with the intention to be applied in gesture recognition applications, most of skin samples are extracted from human hands. Some of the samples for human skin and non-skin are shown in Fig. 5 and Fig. 6, respectively. For human skins, a variety of skins are selected from different ethnicities with varying illumination conditions. It is perhaps important to mention that the training and test data are completely independent sets, and also the color information are obtained from more than one color space instead of using just the information from one color space.

Fig. 5. Samples of training skin data

We have conducted a special experiment intended to investigate the effect of the third channel in HSV color space (i.e., the Value) on the segmentation performance and some samples of the obtained results are given in Fig. 7. In this figure, the first column shows the original gesture images, while the second and third columns show the segmentation results in HSV and HS color spaces, respectively. Moreover, from the figure, one can convincingly conclude that the additional information provided by the Value (V) channel of HSV color space contributes potentially in achieving even finer segmentation results. As stated earlier, two metrics, i.e., correct detection rate (τ) and false alarm rate (ε), have been employed to measure the performance of the proposed method [11]. Table 1 reports the overall segmentation performance produced by the method on the test images in terms of detection rate and false alarm measures.

From the results contained in the table, it is quite remarkable that in HSV space not only the correct detection rate steadily increases, but also the false alarm shows a specific downtrend. To show the feasibility of the proposed method in practical applications, the visual comparison of the proposed method with other related state of the art methods is given in Fig. 8 to demonstrate that the method provides a further improvement over existing methods. Another interesting remark motivated by Fig. 8 is that the proposed method delivers even better performance in face detection applications, however it has developed with the primary aim to be used in hand gesture recognition applications. In general, the achieved results can be rated as rather satisfactory and in a good agreement with those obtained with other relative state of the art methods including those reported in [32, 11, 33]. It should be added as a final note that, in this work, most experiments have been conducted on still images, however the proposed framework can handle real-world image sequences quite efficiently, without requiring any specific change to the implementation of the algorithm. All algorithms in the experiments were implemented in Microsoft Visual Studio 2010 with OpenCV vision library version 2.5 for the graphical processing functions. All tests and evaluations were performed on a PC with Intel(R) Core(TM) 2 - 2.83 GHz processor, 4GB RAM, running Windows 7 Ultimate 64-bit operating system.

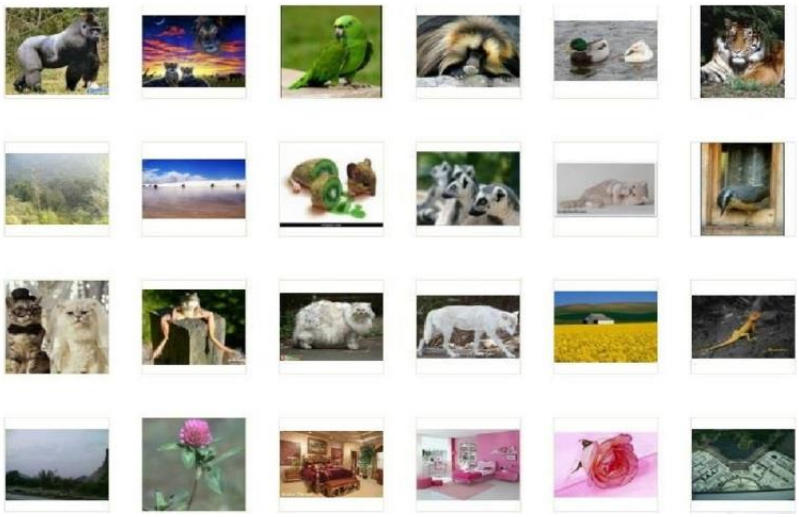


Fig. 6. Samples of training non-skin data

(a) (b) (c)

Fig. 7. Skin segmentation results: (a) Original images, (b) Segmentation without V information, and (c) Segmentation with V information

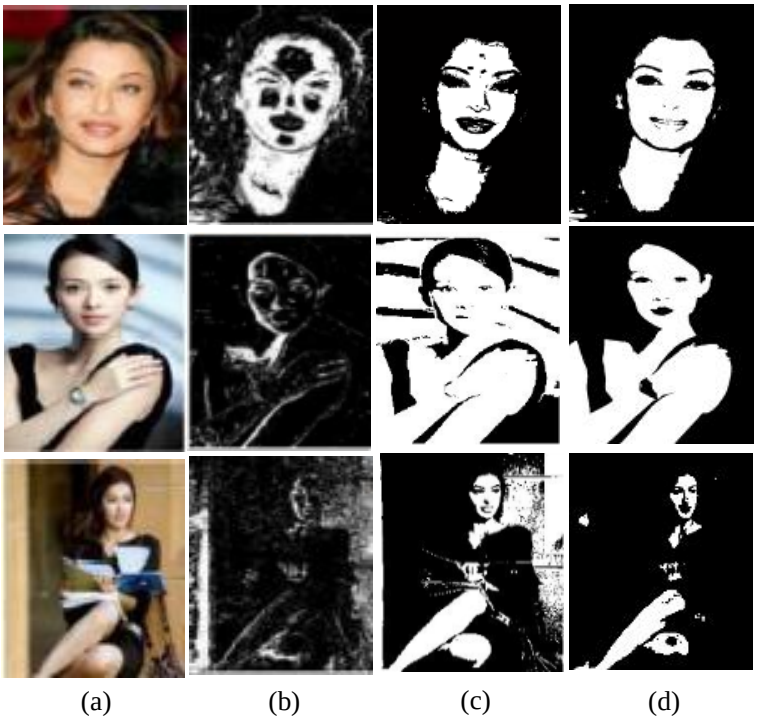


Fig. 8. Comparison with other methods:(a) Original images, (b) Random forest result [33], (c) Dynamic-threshold method result [32] and (d) Our method result

Table 1. Overall performance of the proposed method on the test images

Color features	Recall (τ)	False alarm (ε)
normRG + HSV	0.91	0.11
normRG + HS	0.88	0.15

5 Conclusions and Future Work

This paper has presented a multistage skin segmentation method built as a cascade of a nonparametric generic model and an adaBoost classifier. In this approach, several entities are used to train the adaBoost classifier, and the feature vectors fed into the adaBoost contain color information from two different color spaces. To evaluate the performance of the approach, extensive experiments on two datasets have been conducted. The achieved results are in good agreement and compared very favorably with those of alternative methods previously reported in the literature, revealing that the proposed framework is accurate, robust, and fast. As for future work, we plan to investigate the use of texture features in addition to color features to verify if better segmentation accuracy can be achieved. Furthermore, we also plan to apply the proposed approach to other semantic classes having an intrinsic color (e.g., sky, vegetation, etc.) in order to gain an overall impression about its generalization ability.

Acknowledgement

This research was supported by the Transregional Collaborative Research Center SFB/TRR 62 Companion Technology for Cognitive Technical Systems funded by the German Research Foundation (DFG). The financial support from the DFG is gratefully acknowledged. Moreover, anonymous reviewers are sincerely thanked for their insightful reviews and their valuable comments on earlier versions of the manuscript that helped to improve the quality of the manuscript.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Mofaddel MA, Sadek S. Adult image content filtering: A statistical method based on multi-color skin modeling. In IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'10), Luxor, Egypt. 2010;366-370.

[2] Sadek S, Mofaddel MA. Multicolor skin modeling with application to skin detection. Journal of Computations & Modelling. 2013;3(1):153-167.

[3] Yang J, Lu W, Waibel A. Skin-color modeling and adaptation. In: Proc. of Asia Conf. Computer Vision. 1998;687-694.

[4] Sadek S, Al-Hamadi A, Michaelis B, Sayed U. Human action recognition: A novel scheme using fuzzy log-polar histogram and temporal self-similarity. EURASIP Journal on Advances in Signal Processing; 2011.

[5] Zhu X, Yang J, Waibel A. Segmenting hands of arbitrary color. In Proc. of IEEE Conf. Automatic Face and Gesture Recognition. 2000;687-694.

[6] Sadek S, Abdel-Khalek S. Generalized α-entropy based medical image segmentation. Journal of Software Engineering and Applications. 2014;7(1):62-67.

[7] Zhu Q, Wu CT, Cheng KT, Wu YL. An adaptive skin model and its application to objectionable image filtering. In Proc. of ACM Multimedia. 2004;56-63.

[8] Sadek S, Mofaddel MA, Mahdy YB.Vision-Based Adult Image Content Identification. LAP LAMBERT Academic Publishing, Saarbrücken, Germany; 2013. ISBN 978-3-659-38723-4

[9] Chai D, Ngan KN. Face segmentation using skin-color map in videophone applications. IEEE Trans. Circuits Syst. Video Technol. 1999;9:551-564.

[10] Sadek S, Mahdy YB, Mofaddel MA. A statistical approach to adult image content detection using multi-color skin modeling. International Journal of Advances in Science and Technology. 2010;1(4).

[11] Phung SL, Bouzerdoum A, Chai D. Skin segmentation using color pixel classification: Analysis and comparison. IEEE Trans. Pattern Anal. Machine Intell. 2005;21:148-154.

[12] Sadek S, Al-Hamadi A, Michaelis B. Toward real-world activity recognition: An SVM based system using fuzzy directional features. WSEAS Transactions on Information Science and Applications. 2013;10(4):116-127.

[13] Sadek S, Al-Hamadi A, Michaelis B, Sayed U. Towards robust human action retrieval in video. In Proceedings of the British Machine Vision Conference (BMVC'10), Aberystwyth, UK; 2010.

[14] Sadek S, Al-Hamadi A, Elmezain M, Michaelis B, Sayed U. Human activity recognition using temporal shape moments. In IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'10), Luxor, Egypt. 2010;79-84 .

[15] Ibraheem NA, Khan RZ, Hasan MM. Comparative study of skin color based segmentation techniques. Int. J. Applied Inform. Syst. 2013;5:24-38.

[16] Sadek S, Al-Hamadi A. Entropic image segmentation: A fuzzy approach based on tsallis entropy. International Journal of Computer Vision and Signal Processing. 2015;5(1):1-7.

[17] Brand J, Maso J. A comparative assessment of three approaches to pixel level human skin-detection. In Proc. of the International Conference on Pattern Recognition. 2000;1:1056–1059.

[18] Chun W. Real-time hand interaction for augmented reality on mobile. In Proceedings of the international conference on Intelligent user interfaces. 2013;307-314.

[19] Batool N, Taheri S, Chellappa R. Assessment of facial wrinkles as a soft biometrics. In 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG); 2013.

[20] Prema C. Manimegalai D. Survey on skin tone detection using color spaces. International Journal of Applied Information Systems. 2012;2(2):18–26.

[21] Jones MJ, Rehg JM. Statistical color models with application to skin detection. Intl Journal of Computer Vision. 2002;46:81-96.

[22] Ritter H, Littmann E. Adaptive color segmentation: A comparison of neural and statistical methods. IEEE Trans. Neural Networks. 1997;8(1):175-185.

[23] Albiol A, Torres L, Delp EJ. Optimum color spaces for skin detection. In ICIP, IEEE. 2001;1:122-124.

[24] Nevatia R. A color edge detector and its use in scene segmentation. IEEE Transactions on Systems, Man and Cybernetics. 1977;7(11):820-826.

[25] Sadek S, Al-Hamadi A, Michaelis B, and Sayed U. A robust neural system for objectionable image recognition. In Proceedings of Second International Conference on Machine Vision (ICMV'09) Dubai, UAE. 2009;32-36.

[26] Kovac J, Peer P, Solina F. Human skin color clustering for face detection. In Proc. of EUROCON. 2003;144-148.

[27] Wilso P, Fernandez J. Facial feature detection using haar classifiers.Journal of Computing Sciences in Colleges. 2006;21(4):127-133.

[28] Kim S, Sim JY, Yang S. Vision-based cleaning area control for cleaning robots. IEEE Trans. Consum. Electron. 2012;58(2):685-690.

[29] Sadek S, Al-Hamadi A, Michaelis B, Sayed U. Human action recognition via affine moment invariants. In 21st International Conference on Pattern Recognition (ICPR'12), Tsukuba Science City, Japan. 2012;218-221.

[30] Stottinger J, Hanbury A, Liensberger C, Khan R. Skin paths forcontextual flagging adult video. In Proc. Int. Symp. Visual Comput. 2009;903-906.

[31] Yogarajah P, Cheddad A, Condell J, Curran K, McKevitt P. A dynamic threshold approach for skin segmentation in color images. In IEEE ICIP10. 2010;2225-2228.

[32] Gaikwad S, Khadse PB. Human skin detection: Using fusion approach with iby color space. International Journal of Advance Electrical and Electronics Engineering. 2015;4(3).

[33] Khan R, Hanbury A, Stoettinger J. Skin detection: A randomforest approach. In Proc. Int. Conf. Image Process. 2010;4613-4616.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/15572