

A Hybrid Cascade Approach for Human Skin Segmentation

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 adaBoost 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 preserving a relatively low false alarm rate. 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.

Keywords: Skin Segmentation; Adaboost; Gesture Detection; Pattern Recognition.

2010 Mathematics Subject Classification: 53C25; 83C05; 57N16

1 Introduction

During the course of the past few decades, skin segmentation and detection has been intensively investigated and many improvements have been achieved, however the state-of-the-art of skin segmentation/detection is still in early stages of acceptable performance or suboptimal in performance (Sadek et al., 2013). Furthermore, due to several significant barriers and challenges

that should be navigated to develop practical solutions for efficient and effective skin segmentation, most existing approaches in the literature are still far from finding its way towards applications to real-time recognition systems (Mofaddel and Sadek, 2010).

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 (Yang et al., 1998), hand gesture recognition (Zhu et al., 2000), adult image content filtering (Zhu et al., 2004), and videophone (Chai and Ngan, 1999).

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 (Yang et al., 1998; Zhu et al., 2000, 2004; Phung et al., 2005; Jones and Rehg, 2002) and generic color modeling (Yang et al., 1998; Chai and Ngan, 1999). 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 (Ibraheem et al., 2013). 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, skin detection/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 rest of the paper unfolds as follows. Section 2 reviews few related studies. In section 3, the proposed skin segmentation method is described in detail. The experimental results are given and the corresponding analysis is discussed in section 4. Finally, in section 5, we summarize the paper and suggest some possible future works.

2 Related Work

Human skin detection/classification that aims at finding skin-color pixels or regions in a given image is quite rigorous part of state of the art applications in various disciplines, such as facial recognition, gesture analysis, etc. In (Brand and Maso, 2000), the authors make sue of skin detection approaches for biometric identification applications. Moreover, in (Chun, 2013), skin detection is used in cosmetic surgery application for detecting facial wrinkles and fixes to affected areas of face. In this work, the

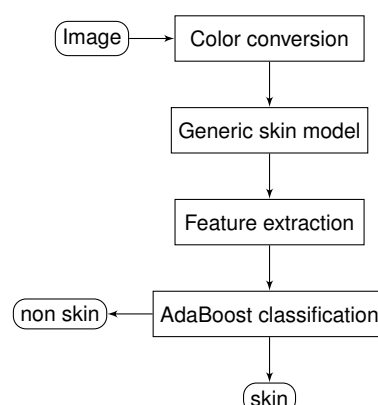


Figure 1: Overall workflow of the proposed approach

authors introduce an algorithm for wrinkles curves and pattern matching for the purpose of facial recognition. Furthermore, in (Batool et al., 2013), 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 (Prema and Manimegalai, 2012; Jones and Rehg, 2002). 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. Among these three approaches, 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 (Prema and Manimegalai, 2012). 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, to select the most appropriate color space seems to be a very daunting task. To the best of our knowledge, Littman and Ritter (E. Littmann, 1997) have, for the first time, compared color spaces. In their work, a neural approach based on linear maps for skin color with normal distributions using three different colors: RGB, YIQ and YUV on a relatively small dataset of hand images is proposed. In all, the authors have concluded that the neural approach is produces better classification and its performance is quasi-independent of the color space representation.

In (Albiol et al., 2001), 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, YCbCr, and HSV), and their performances are found to be in close agreement. Moreover, the authors have concluded that the detection performance in the 3D YCbCr color space is better than that of the 2D CbCr color space, since the transformation form the 3D space to the 2D space is not invertible.

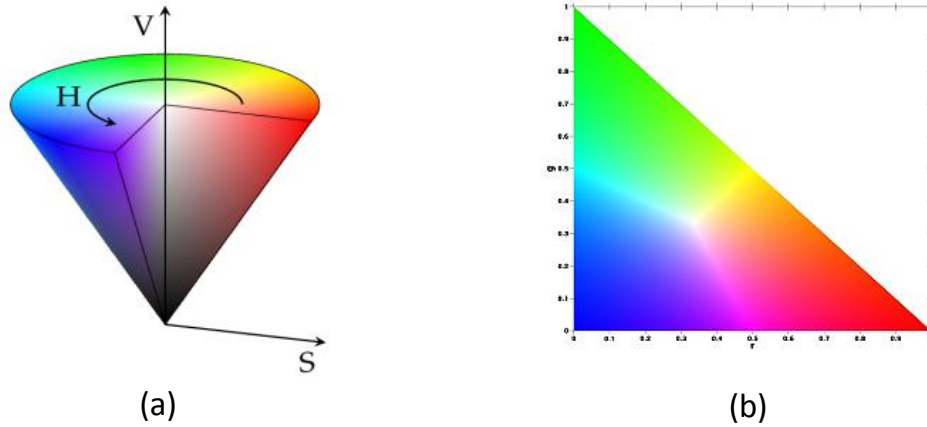


Figure 2: HSV and normalized RG color models.

3 Proposed Approach

In this section, we provide an overview of the proposed approach, then each component and their interactions are described in more details. Fig.1 shows a flow diagram of the proposed approach. This approach is essentially based on segmentation of the normalized RGB 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. We use the following normalized color representation space (Nevatia, 1977), given as:

$$\begin{aligned} y &= c_1 r + c_2 g + c_3 b \\ \tau_1 &= \frac{r}{r+g+b} \\ \tau_2 &= \frac{g}{r+g+b} \end{aligned} \quad (3.1)$$

where r, g , and b stands for RGB channels with values from 0 to 255, and $c_1 + c_2 + c_3 = 0$ 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 normalized RGB 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, 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

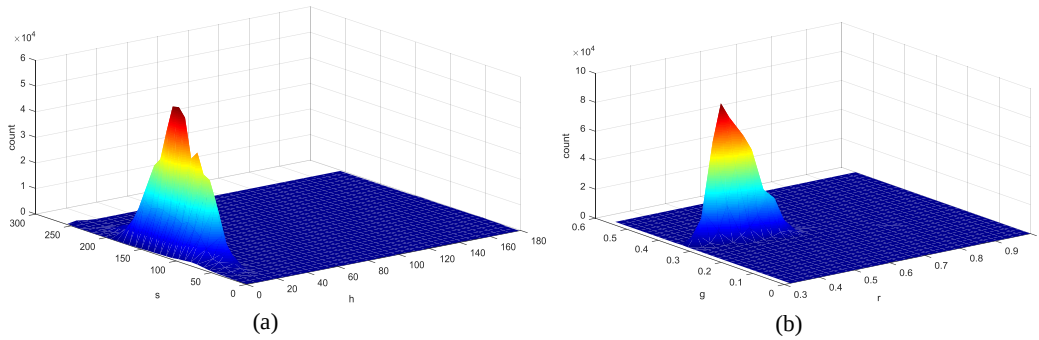


Figure 3: 3D histogram distributions of skin color: (a) HS space, (b) normalized RG space.

applied:

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

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 the current 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 extra channel to improve overall color segmentation quality.

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 (Kovac et al., 2003), we find it more appropriate to define a fixed skin color range in two color spaces (i.e., normalized RGB and HSV) to represent the skin pixels, as follows,

1. In normalized RGB 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 normalized RGB 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.

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 (NN), Support Vector Machine (SVM), k-Nearest Neighbor (kNN), etc. that can be employed for skin segmentation. To perform this task, we opted for an AdaBoost algorithm (Wilso and Fernandez, 2006) that provides high classification rate and can be applied in real-time.

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. It is a meta-algorithm because it addresses sub-procedures. This algorithm has a capability to upgrade so called "weak learners" to a strong learner. The terms 'weak learner' and 'strong learner' are derived from the PAC. (Probably Approximately Correct) 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 adjust 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 (Kim et al., 2012).

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%). Roughly speaking, provided a so-called Weak Learning Assumption holds, boosting can then 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 classifier, 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:      $\alpha_t \leftarrow \frac{1}{2} \ln \left( \frac{1 - \epsilon(h_t, \tilde{W}_t)}{\epsilon(h_t, \tilde{W}_t)} \right)$ 
7:      $\tilde{W}_{t+1}[0 : n - 1] \leftarrow \frac{1}{Z} \tilde{W}_t[0 : n - 1] e^{-\alpha_t \lambda}$ 
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, we find (any) classifier $h(\mathbf{x})$ for which the weighted classification error:

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

is better than chance. The weights are updated at Step 7 (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 mention that after each boosting iteration, assuming we can find a component classifier whose weighted error is better than chance, the combined classifier is then guaranteed to have a lower exponential loss over the training examples (see 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 (TP_r) and false alarm rate (FP_r). In the context of this segmentation task, TP_r corresponds to the number of correctly classified skin pixels with respect to the total number of skin pixels in ground truth data, while FP_r 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 the detection rate (recall) TP_r and false alarm rate FP_r are defined as follows:

$$TP_r = \frac{TP}{TP+FN}, \quad FP_r = \frac{FP}{TN+FP} \quad (3.3)$$

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; that 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 TP_r and the lower the FP_r , the more accurate the method.

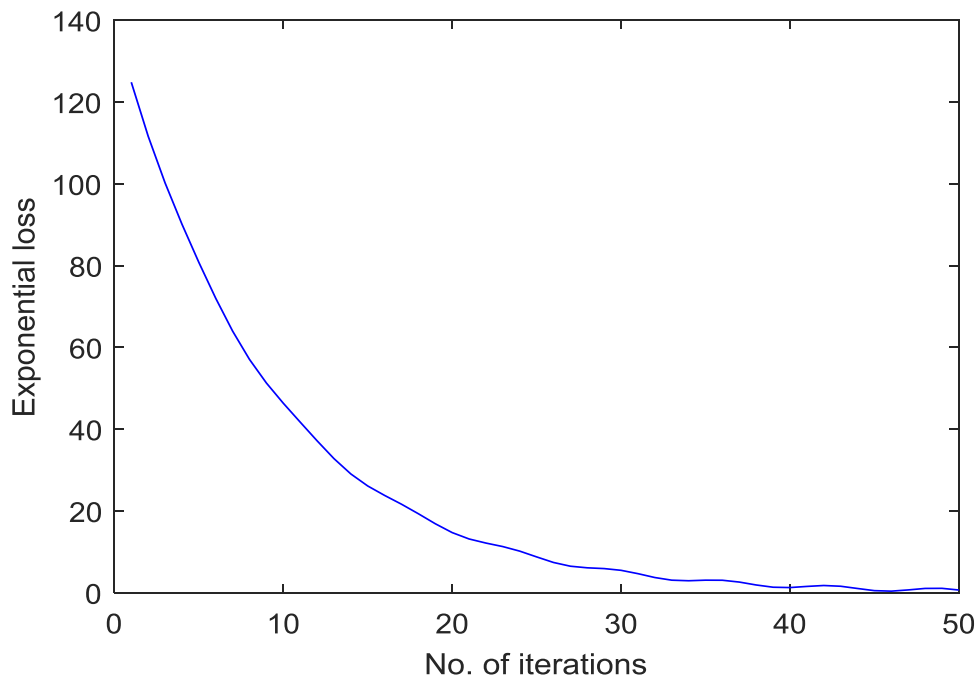


Figure 4: A lower exponential loss over the training samples.

4 Experiments and Results

In this section, the performance of the proposed approach is evaluated under different conditions, such as scale change, occlusion, orientation change, color spaces, and 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 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 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 efficiency of our approach. To this end, some illustrative images are randomly taken from Stottinger(Stottinger et al., 2009) and Pratheepan datasets (Yogarajah et al., 2010) 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



Figure 5: Samples of training skin data.

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.

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 the 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. 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 the correct detection rate (TP_r) and false alarm rate (FN_r) have been employed to measure the performance of the presented method (Phung et al., 2005). Table 1 reports the overall segmentation performance produced by the method in terms of detection rate and false alarm measures. From the results contained in the above table, it is remarkable that in HSV space not

Segmentation	TP_r	FP_r
normRG+HSV	0.91	0.11
normRG+HS	0.88	0.15

Table 1: Overall performance of the proposed method.

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 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

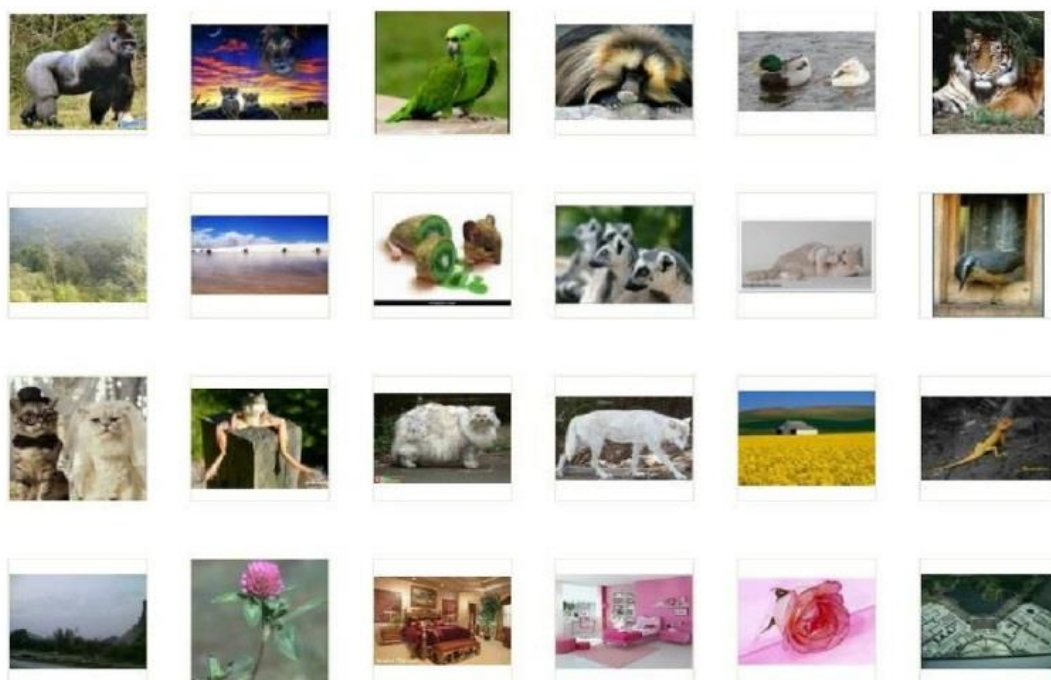


Figure 6: Samples of training non-skin data.

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 (P. B. Khadse, 2015; Phung et al., 2005; Khan et al., 2010). 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.

5 Conclusions and Future Work

In this paper, a multistage skin segmentation method built as a cascade of a nonparametric generic model and an adaBoost classifier is proposed. With 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. In order 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 adapt the proposed approach for other



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

semantic classes having an intrinsic color (e.g., sky, vegetation, etc.) in order to get an overall impression about its generalization ability.

Acknowledgment

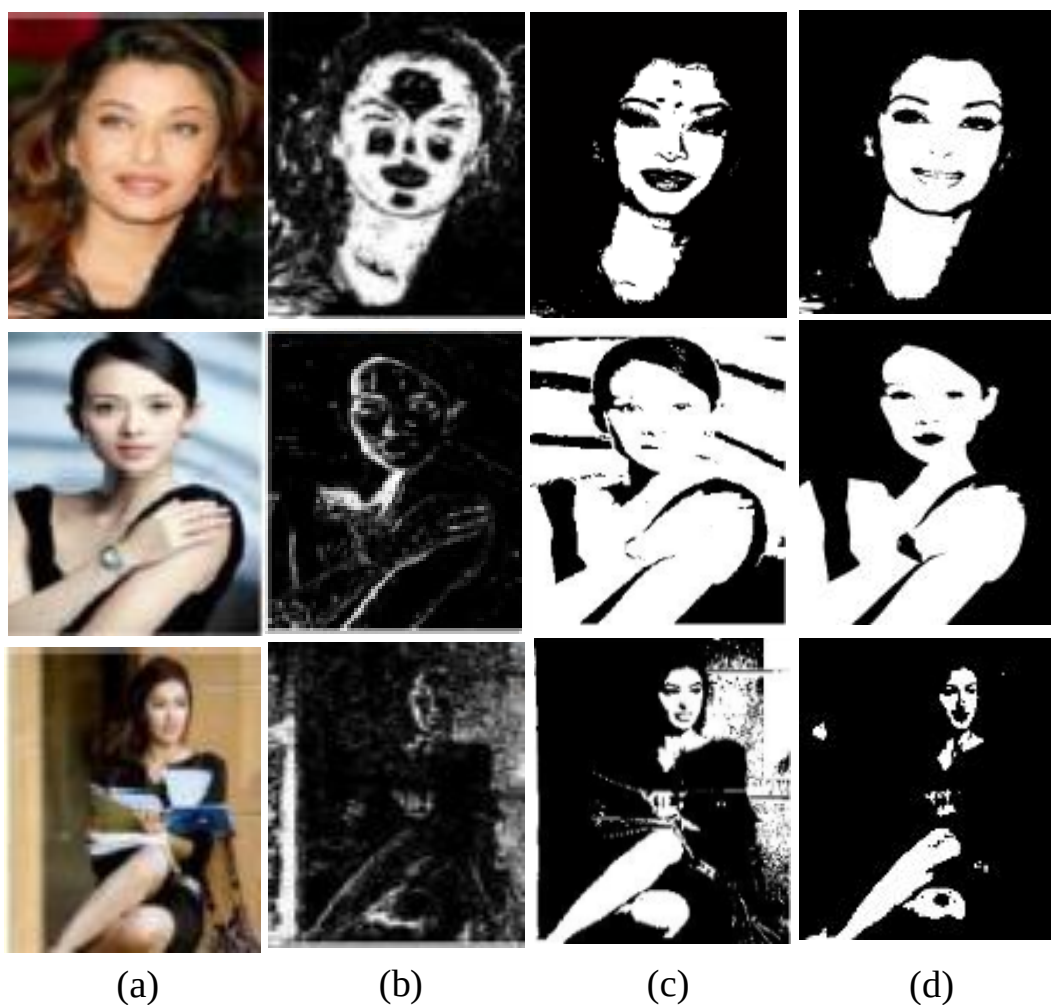


Figure 8: Comparison with other methods:(a) Original images, (b) Random forest result (Khan et al., 2010), (c) Dynamic-threshold method result (P. B. Khadse, 2015) and (d) Our method result.

A brief acknowledgement section may be given after the conclusion section just before the references. The acknowledgments of people who provided assistance in manuscript preparation, funding for research, etc. should be listed in this section.

References

- Albiol, A., Torres, L., and Delp, E. (2001). Optimum color spaces for skin detection. In *ICIP*, volume 1, pages 122–124. IEEE.
- Batool, N., Taheri, S., and Chellappa, R. (2013). Assessment of facial wrinkles as a soft biometrics. In *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*.
- Brand, J. and Maso, J. (2000). A comparative assessment of three approaches to pixel level human skin-detection. In *Proc. of the International Conference on Pattern Recognition*, volume 1, pages 1056–1059.
- Chai, D. and Ngan, K. N. (1999). Face segmentation using skin-color map in videophone applications. *IEEE Trans. Circuits Syst. Video Technol.*, 9:551–564.
- Chun, W. (2013). Real-time hand interaction for augmented reality on mobile. In *Proceedings of the international conference on Intelligent user interfaces*, pages 307–314.
- E. Littmann, H. R. (1997). Adaptive color segmentation: A comparison of neural and statistical methods. *IEEE Trans. Neural Networks*, 8(1):175–185.
- Ibraheem, N., Khan, R., and Hasan, M. (2013). Comparative study of skin color based segmentation techniques. *Int. J. Applied Inform. Syst.*, 5:24–38.
- Jones, M. J. and Rehg, J. M. (2002). Statistical color models with application to skin detection. *Intl Journal of Computer Vision*, 46:81–96.
- Khan, R., Hanbury, A., and Stoeftinger, J. (2010). Skin detection: A randomforest approach. In *Proc. Int. Conf. Image Process.*, pages 4613–4616.
- Kim, S., Sim, J., and Yang, S. (2012). Vision-based cleaning area control for cleaning robots. *IEEE Trans. Consum. Electron*, 58(2):685–690.
- Kovac, J., Peer, P., , and Solina, F. (2003). Human skin color clustering for face detection. In *Proc. of EUROCON*, pages 144–148.
- Mofaddel, M. A. and Sadek, S. (2010). 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)*, pages 366–370, Luxor, Egypt.
- Nevatia, R. (1977). A color edge detector and its use in scene segmentation. *IEEE Transactions on Systems, Man and Cybernetics*, 7(11):820–826.
- P. B. Khadse, S. G. (2015). Human skin detection: Using fusion approach with iby color space. *International Journal of Advance Electrical and Electronics Engineering*, 4(3).
- Phung, S. L., Bouzerdoun, A., and Chai, D. (2005). Skin segmentation using color pixel classification: Analysis and comparison. *IEEE Trans. Pattern Anal. Machine Intell.*, 21:148–154.
- Prema, C. and Manimegalai, D. (2012). Survey on skin tone detection using color spaces. *International Journal of Applied Information Systems*, 2(2):18–26.
- Sadek, S., Mofaddel, M. A., and Michaelis, B. (2013). Multicolor skin modeling with application to skin detection. *Journal of Computations & Modelling*, 3(1):153–167.

- Stottinger, J., Hanbury, A., Liensberger, C., and Khan, R. (2009). Skin paths for contextual flagging adult video. In *Proc. Int. Symp. Visual Comput*, pages 903–906.
- Wilso, P. and Fernandez, J. (2006). Facial feature detection using haar classifiers. *Journal of Computing Sciences in Colleges*, 21(4).
- Yang, J., Lu, W., and Waibel, A. (1998). Skin-color modeling and adaptation. In *Proc. of Asia Conf. Computer Vision*, pages 687–694.
- Yogarajah, P., Cheddad, A., Condell, J., Curran, K., and McKeivitt, P. (2010). A dynamic threshold approach for skin segmentation in color images. In *IEEE ICIP10*, pages 2225–2228.
- Zhu, Q., Wu, C. T., Cheng, K. T., and Wu, Y. L. (2004). An adaptive skin model and its application to objectionable image filtering. In *Proc. of ACM Multimedia*, pages 56–63.
- Zhu, X., Yang, J., and Waibel, A. (2000). Segmenting hands of arbitrary color. In *Proc. of IEEE Conf. Automatic Face and Gesture Recognition*, pages 687–694.