LEARNING SKIN DISTRIBUTION USING A SPARSE MAP

Rajkiran Gottumukkal and Vijayan K. Asari

Department of Electrical and Computer Engineering,
Old Dominion University, Norfolk, VA 23529, USA.
E-mail: {rgott002, vasari}@odu.edu

ABSTRACT

We present a new skin modeling technique based on SNoW (Sparse Network of Winnows) for accurate and robust skin region detection. A Skin Distribution Map (SDM) representing the sparse network is trained with skin pixels to learn their distribution in a color space. We then train the SDM with non-skin pixels to unlearn the distribution of the non-skin pixels, which overlap with the skin pixels in the color space. This skin model can be used for skin detection on any color space. We found the accuracy of skin detection using SDM to be slightly better than that using the Skin Probability Map (SPM) method. The main advantage of using SDM method over SPM method is that the complexity, memory requirements and time for skin detection are reduced significantly.

1. INTRODUCTION

Skin detection can be defined as the process of identifying the pixels of a given color image which corresponds to human skin. Humans use color as a fundamental clue for detecting objects from complex scenes. In recent years many researchers have used human skin color in applications such as gesture recognition [10], video indexing [9], face detection [16], face tracking [7], region of interest segmentation [8] etc. Skin detection is a difficult problem since skin colors vary based on the ambient light and the cameras, which produce different colors, even for the same person, under the same illumination conditions. Skin colors also vary from person to person.

Probabilistic methods have been widely used for pixel based skin detection. One of the most widely used probabilistic methods is the Skin Probability Map (SPM) [1, 2], which has been found to be the best one in terms of accuracy and running time [2]. Other well-known statistical skin color models are single Gaussian model [11] and the mixture of Gaussians model [12]. The single Gaussian skin model has the advantage of being simple and fast, it however does not adequately represent the variance of the skin distributions under different conditions. To overcome this drawback, the mixture of the Gaussians model has been suggested. It is however hard to be trained and it is slow. Some of the earlier skin detection techniques classify pixels based on pre-defined ranges in the color space [14]. More recently machine-learning algorithms have been used to perform pixel classification based on predefined ranges in color space [13]. The method we use for skin modeling is also a machine-learning algorithm (SNoW). It was first used in image processing for face detection with great success [15]. We consider the skin pixels and non-skin pixels as two classes and perform the training using SNoW.

The choice of color space has a significant impact on the skin detection result. Researchers have tried to find the most effective color space for skin detection by performing detailed analysis on different color spaces [5, 6]. But there is no single color space, which is most effective for all skin detection methods [3]. It has been argued theoretically that for every color space, there exists an optimum skin detection scheme such that the performance of all these skin detection schemes is the same in that color space [4].

2. SKIN COLOR MODEL USING SNOW

SNoW is a learning architecture, which can be trained to learn the feature (color) space of skin pixels by providing it with positive and negative examples. The process of training the network is described in the remainder of this section. We created a set of skin pixels from images in the AR face database [17] and images taken from a digital camera. The images have illuminations varying from fluorescent light to direct sunlight. From the outdoor scenes we used skin regions from both direct sunshine and shadows. We didn't use any images from the internet during training since there is no control on the image properties. Training was done on the *Cb-Cr* color space since the transformation from *RGB* is simple and it has

0-7803-8554-3/04/$20.00 ©2004 IEEE.

explicit separation of luminance and chrominance components. We say more about the choice of color space in the next section.

A two-dimensional skin distribution map (SDM) of dimension 256 × 256 was created and all its elements were initialized to 0. The network is then trained in the *Cb-Cr* color space with the image created from the skin pixels. During training with the skin pixels, the elements of the SDM are selected for updating based on the index computed as *(Cb(x,y), Cr(x,y))*. The winnow update rule is used to update the matrix and the update procedure is as shown:

$$\begin{aligned} &\text{if} \quad W(Cb(x,y), Cr(x,y)) == 0 \qquad \forall x,y \\ &\qquad W(Cb(x,y), Cr(x,y)) = \eta \end{aligned} \tag{1}$$

The elements or weights of the SDM, which are active, are initialized to η as shown in expression (1); η is 0.1 in our implementation. Next the weights are promoted for all the pixels in the skin image as shown in expression (2).

$$\begin{aligned} &\text{if} \quad W(Cb(x,y), Cr(x,y)) < \eta \times \alpha^n \qquad \forall x,y \\ &\qquad W(Cb(x,y), Cr(x,y)) = \alpha \times W(Cb(x,y), Cr(x,y)) \end{aligned} \tag{2}$$

In the above expressions W is the SDM; Cb, Cr represent the color components of the skin image and x, y vary from 1 to X, Y respectively where X is the number of columns in the skin image and Y is the number of rows in the skin image. The term α in expression (2) is used for promoting the weights, and $\alpha = 2$ in our case. The term n=200 in our implementation, which signifies that we stop promoting the weights if more than 200 pixels from the skin image occur at the same point in the *Cb-Cr* color space. The result of an initial skin detection performed on a test image from a digital camera, using the SDM updated only with training skin pixels is shown in figure 1. Skin detection was performed by projecting the pixels into the SDM based on their *Cb* and *Cr* values. If the weight for a pixel in the SDM is greater than 1000 (threshold selected after experimental analysis) it is classified as a skin pixel. In figure 1, pixels classified as skin are represented as white and non-skin pixels are represented as black. As can be seen in figure 1, many non-skin pixels are falsely classified as skin pixels. Non-skin pixels falsely classified as skin pixels are used as negative examples to update the SDM. More examples of non-skin pixels were obtained by performing skin detection on test images and the misclassified non-skin pixels were incorporated into the non-skin image. The SDM was then updated with the non-skin image as shown in expression (3), where t=1000 and β=0.5.

(a) (b)

Fig. 1. (a) Original test image and (b) Initial skin detection.

$$\begin{aligned} &\text{if} \quad W(Cb(x,y), Cr(x,y)) > t \qquad \forall x,y \\ &\qquad W(Cb(x,y), Cr(x,y)) = \beta \times W(Cb(x,y), Cr(x,y)) \end{aligned} \tag{3}$$

Expression (3) performs the demotion of weights for pixels from the non-skin examples, which overlap with the skin examples in the *Cb-Cr* color space. The advantage of using the winnow update rule for updating the weights is that we can reduce the false positives without adversely affecting the true positives. This process is continued as long as a large decrease in false positives is observed without decreasing the true positives. The SDM in *Cb-Cr* color space after updating it with non-skin pixels and the detected skin regions obtained on a test image are shown in figure 2.

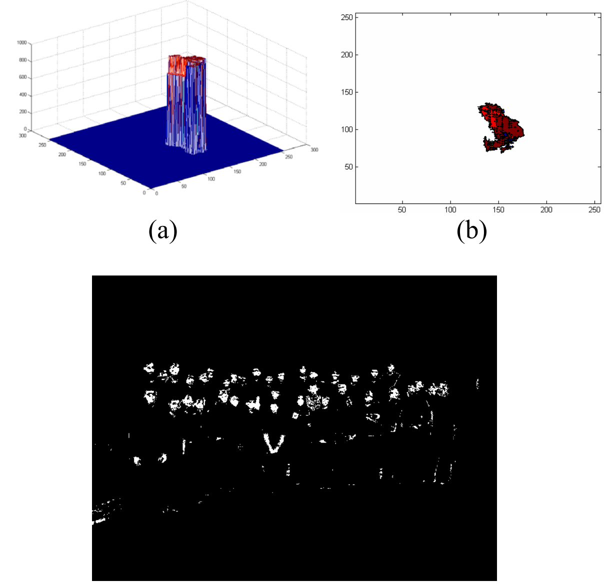
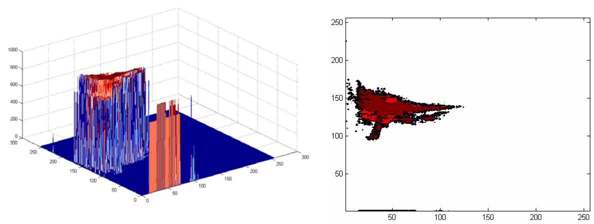


Fig. 2. (a) Side view of the SDM in *Cb-Cr* color space, (b) Top view of the SDM in *Cb-Cr* color space and (c) Detected skin pixels using the SDM.

Please note from figure 2 that most of the non-skin pixels which were previously misclassified are now classified correctly. From expression (2) it can be observed that the maximum possible value an element in the SDM can take is 0.1×2^{200}, therefore it is not possible to display the SDM since some of the weights are very high. Hence to display the SDM as in figure 2, all the weights greater that 1000 were scaled down to 1000.

3. SELECTING COLOR SPACE

We are interested in selecting a color space, which would give the best skin detection results for our technique. From the literature we found that one of the best skin detection was obtained using the *T-S* color space [5]. Hence we trained our skin detection algorithm on the *T-S* color space expecting to obtain better skin detection accuracy than that obtained using *Cb-Cr* color space. The SDM is trained in a similar manner as described in the previous section, the only difference being that we train it on the *T-S* histogram for skin and non-skin examples. The SDM obtained in *T-S* color space is shown in figure 3.



(a) (b)

Fig. 3. (a) Side view of the SDM in *T-S* color space, (b) Top view of the SDM in *T-S* color space.

From [4] we may hypothesize that for any skin detection scheme there exists a color space for which the given skin detection scheme performs at its best. This being the case, we may obtain an optimum color space for our skin detection scheme using the technique presented in [18]. However, finding the optimum color space for skin detection using SDM is one of the future research goals. In this paper we restrict ourselves to comparing SDM and SPM methods on *Cb-Cr* and *T-S* color spaces.

4. COMPARATIVE EVALUATION

We evaluated the performance of SDM and SPM based skin detection methods using the same training data and tested on the same set of images collected from different sources. We obtained the true positives (*TP*) and false positives (*FP*) of both the methods for 10 test images. The ROC curves for SDM and SPM methods in *Cb-Cr* and *T-S* color spaces are shown in figure 4 and figure 5 respectively. We counted the number of pixels correctly classified as skin and number of pixels falsely classified as skin at different thresholds for the test images to obtain the *TP* and *FP*. The *TP* rate (*TP* rate = (*TP / total number of skin pixels*) × *100%*) and *FP* rate (*FP* rate = (*FP / total number of non- skin pixels*) × *100%*) at different thresholds for all the test images were averaged to obtain the ROC curves. The results shown are for ~0.4 million skin pixels and ~3 million non-skin pixels from the test images.

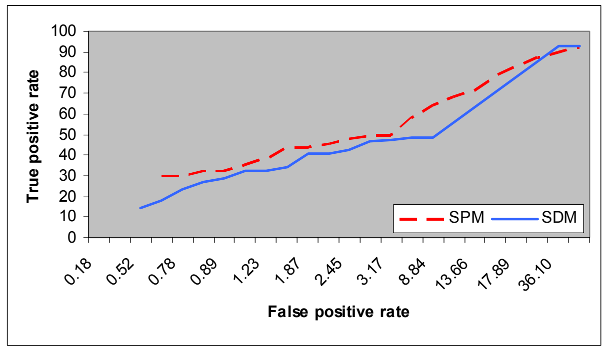


Fig. 4. ROC curves comparing the accuracy of SDM and SPM methods for skin detection in *Cb-Cr* color space.

The underlying principle of SDM and SPM based skin detection methods are similar. Instead of using two matrices to represent skin and non-skin probabilities, SDM uses a single matrix and updates its elements using winnow update rule for skin and non-skin examples. The weight search complexity of SDM technique is roughly $O(d)$, where as for SPM it is $O(2d)$; where d is the dimension of the map. Also, in the SPM method $X \times Y$ divisions are performed to obtain the ratio of $P(Skin|Cb\text{-}Cr)$ and $P(Non\text{-}skin|Cb\text{-}Cr)$; where X and Y are the dimensions of the image on which skin detection is performed. In the case of SDM no computations other than a comparison with a threshold is required during skin detection. The memory required for SDM is half (N^2) of that used by SPM ($2N^2$) since both the skin and non-skin training data is incorporated into a single matrix. Table 1 summarizes these comparisons; here N^2 stands for number of elements in the SPM/SDM and the time for skin detection was computed for an image of size 1024 × 768 pixels using MATLAB program running on a 1.5GHz computer with 768MB of RAM.

Algorithm	No. Div.	Memory	Time (sec)	Weight search complexity
SPM	$X \times Y$	$2N^2$	7.0	$O(2d)$
SDM	0	N^2	5.9	$O(d)$

Table 1. Performance comparison of SPM and SDM methods for skin detection.

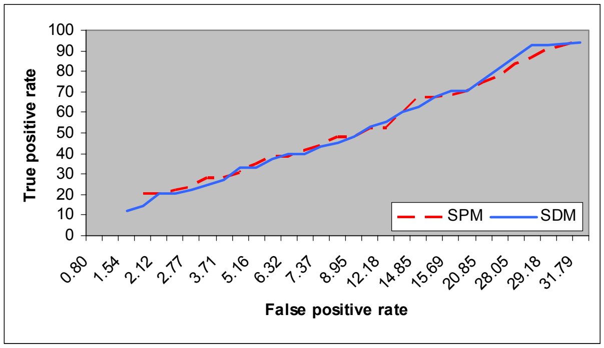


Fig. 5. ROC curves comparing the accuracy of SDM and SPM methods for skin detection in *T-S* colour space.

5. CONCLUSION

In this paper we presented an approach to learn the skin color distribution based on the SNoW learning scheme. The learning scheme is used to update the skin distribution map (SDM) using the skin and non-skin training data. It was found that both the methods have comparable skin detection accuracy on the two color spaces. Also a slight improvement in accuracy was observed in SDM method over the SPM method for true positive rates greater than 90%. The main contributions are that we have reduced the weight search complexity, computational complexity, memory requirements and time required for skin detection using SDM based skin detection while maintaining the accuracy comparable to SPM method. One of the important future research directions would be to identify the color space on which SDM method provides the best skin detection results.

6. REFERENCES

[1] M.J. Jones, and J.M. Regh, "Statistical color models with applications to skin detection," in *Proc. of CVPR*, vol. I, pp. 274-280, 1999.

[2] J. Brand, and J.S. Mason, "A comparative assessment of three approaches to pixel-level human skin-detection," in *Proc. of ICPR*, vol. I, pp. 1056-1059, 2000.

[3] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques," in *Proc. Graphicon-2003*, pp. 85-92, 2003.

[4] A. Albiol, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection," in *Proc. of the ICIP*, vol. 1, pp. 122-124, 2001.

[5] J.C. Terrillon, and S. Akamasu, "Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images," in *Proc. of the 12th Conference on Vision Interface (VI '99)*, vol.2, pp. 180-187, 1999.

[6] G. Gomez, "On selecting color components for skin detection," in *Proc. of the ICPR*, vol. 2, pp. 961-964, 2002.

[7] Y. Raja, S.J. McKenna, and S. Gong, "Tracking and Segmenting People in Varying Lighting Conditions Using Colour," in *Proc. Automatic Face and Gesture Recognition*, pp. 228-233, 1998.

[8] M. Fleck, D.A. Forsyth, and C. Bregler, "Finding naked people," in *Proc .of ECCV*, Berlin, Germany: Springer-Verlag, pp. 592-602, 1996.

[9] F. Hsu, S. Lee, and B. Lin, "Video Data Indexing by 2D C-Trees," *Journal of Visual Languages and Computing*, vol. 9, no. 4, Elsevier Science, pp. 375-397, 1998.

[10] Y. Zhu, G. Xu, and D. Kriegman, "A Real-Time Approach to the Spotting, Representation, and Recognition of Hand Gestures for Human-Computer Interaction," *Computer Vision and Image Understanding*, vol. 85, no. 3, pp. 189-208, 2002.

[11] J. Yang, W. Lu, and A. Waibel, "Skin Color Modeling and Adaptation," in *Proc .of ACCV*, vol. II, pp. 687-694, 1998.

[12] M.-H. Yang, and N. Ahuja. "Gaussian mixture model for human skin color and its application in image and video databases," in *Proc. of the SPIE: Storage and Retrieval for Image and Video Databases VII*, vol. 3656, pp. 458-466, 1999.

[13] D. Brown, I. Craw, and J. Lewthwaite, "A SOM based approach to skin detection with application in real time systems," in *Proc. of the British Machine Vision Conference*, 2001.

[14] K. Sobottka, and I. Pitas, "Segmentation and tracking of faces in color images", in *Proc. Automatic Face and Gesture Recognition*, pp. 236-241, 1996.

[15] M.-H. Yang, D. Roth, and N. Ahuja, "A SNoW-Based Face Detector," *Advances in Neural Information Processing Systems,* vol. 12, MIT Press, pp. 855-861, 2000.

[16] H. Wu, Q. Chen, and M. Yachida, "Face Detection From Color Images Using a Fuzzy Pattern Matching Method," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 6, pp. 557-563, 1999.

[17] A.M. Martinez, and R. Benavente, "The AR Face Database," *CVC Technical Report #24*, June 1998.

[18] G. Gomez, and E. Morales, "Automatic Feature Construction and a Simple Rule Induction Algorithm for Skin Detection," in *Proc. of the ICML Workshop on Machine Learning in Compute Vision*, pp. 31-38, 2002.