

OPTIMUM COLOR SPACES FOR SKIN DETECTION

Alberto Albiol[†], Luis Torres[‡], Edward J. Delp[§]

Politechnic University of Valencia, Spain[†]
Politechnic University of Catalonia, Spain[‡]
Purdue University, USA[§]

ABSTRACT

The objective of this paper is to show that for every color space there exists an optimum skin detector scheme such that the performance of all these skin detectors schemes is the same. To that end, a theoretical proof is provided and experiments are presented which show that the separability of the skin and no skin classes is independent of the color space chosen.

1. INTRODUCTION

Skin detection can be defined as the process of selecting which pixels of a given image correspond to human skin. Skin detection is useful in, for example, face detection and face tracking for security and video indexing applications, model-based video coding, etc. [1, 2, 3]. There are some difficulties when detecting skin pixels. Skin color is affected by ambient light which is unknown in many situations; different cameras produce different colors, even from the same person, under the same illumination conditions; and finally, skin colors change from person to person. Several color spaces have been proposed in the literature for skin detection applications. YCbCr has been widely used since the skin pixels form a compact cluster in the Cb-Cr plane. As YCbCr is also used in video coding and then no transcoding is needed, this color space has been used in skin detection applications where the video sequence is compressed [1, 2]. The HSV color space has been proposed in [4, 5] because it is more related to human color perception. In [6] two components of the normalized RGB color space (rg) have been proposed to minimize luminance dependence. And finally CIE Lu*v* has been used in [7]. However, it is still not clear which is the color space where the skin detection performance is the best. In addition, it is said many times in the literature that the use of a specific color space improves the performance of the skin detectors, [6, 8] which is not exactly true. The objective of this paper is to show that for all color spaces their corresponding optimum skin detectors

This work was partially supported by the grants TIC 98-0442, TIC 98-0335 of the Spanish Government and Programa Hispano-Norteamericano de Cooperación Científica y Tecnológica

have the same performance since the separability of the skin and no skin classes is independent of the color space chosen. That is, the color space does not have any influence, provided that the optimum skin detector for that color space is used. A mathematical proof is given in section 3 and examples are provided to verify this statement in section 4.

2. THEORETICAL BACKGROUND

In this section some definitions are provided which are useful to proof our statement.

Definition: Let \mathcal{I} be an image, p a pixel of the image \mathcal{I} and \mathbf{x}_p a vector containing the values of the pixel p in the color space \mathcal{C} . A function $D(\mathbf{x}_p)$ is called a skin color detector in \mathcal{C} if:

$$D(\mathbf{x}_p) = \begin{cases} 1 & \text{if } \mathbf{x}_p \text{ is considered skin} \\ 0 & \text{if } \mathbf{x}_p \text{ is considered no skin} \end{cases} \quad (1)$$

Definition: Given a set of images \mathcal{H} and a skin color detector $D(\mathbf{x}_p)$, the performance of $D(\mathbf{x}_p)$ can be evaluated with the following two parameters:

$$P_D = \frac{\sum_{p \in S} D(\mathbf{x}_p)}{P_S} \quad P_{FA} = \frac{\sum_{p \in N} D(\mathbf{x}_p)}{P_N} \quad (2)$$

where P_D is the detection rate, P_{FA} is the false alarm rate and S and N are the subsets of skin and non-skin pixels in \mathcal{H} with populations P_S and P_N respectively.

Definition: A skin color detector $D(\mathbf{x}_p)$ is said to be optimum when it has the highest detection rate P_D for a given fixed false alarm P_{FA} .

3. STATEMENT

Let $D(\mathbf{x}_p)$ be an optimum skin detector defined in \mathcal{C} and $T(\mathbf{x}_p)$ be an invertible function that transforms the pixels values \mathbf{x}_p from the color space \mathcal{C} into a vector \mathbf{x}'_p in a different a color space \mathcal{C}' , then there exists another optimum skin detector $D'(\mathbf{x}'_p)$ in \mathcal{C}' , with the same detection rate and false alarm rate, given by:

$$D'(\mathbf{x}'_p) = D(T^{-1}(\mathbf{x}'_p)) \quad (3)$$

Color space	Number of bins/channel
HSV	$180 \times 50 \times 33$
RGB	$85 \times 85 \times 85$
YCbCr	$85 \times 128 \times 128$
CbCr	128×128

Table 1. Number of bins/channel used to estimate $p(\mathbf{x}_p/H_0)$ and $p(\mathbf{x}_p/H_1)$ for each color space.

This means that we can always find a skin color detector with the same performance no matter the specific color space chosen, provided that there exists an invertible transformation between the compared color spaces.

Proof: Since $\mathbf{x}_p = T^{-1}(\mathbf{x}'_p)$, the detection and false alarm rates for $D(\mathbf{x}_p)$ and $D'(\mathbf{x}'_p)$ are exactly the same. If there exists another skin detector $D''(\mathbf{x}'_p)$ in \mathcal{C}' which has a higher detection rate a new skin detector $D'''(\mathbf{x}_p)$ in \mathcal{C} can be constructed:

$$D'''(\mathbf{x}_p) = D''(T(\mathbf{x}_p)) \quad (4)$$

which would have a detection rate $P_D > P'_D$ in \mathcal{C} , which is wrong by hypothesis.

4. RESULTS AND EXPERIMENTS

In order to prove the theoretical statement of the previous section three skin detectors have been designed for the RGB, YCbCr and HSV color spaces using the Neyman-Pearson test [9]. To that end, the skin detection is formulated as a classical detection problem where two hypothesis are tested. The two possible hypothesis for the pixel values of \mathbf{x}_p are:

$$H_0 = \mathbf{x}_p \text{ is skin} \quad H_1 = \mathbf{x}_p \text{ is not skin} \quad (5)$$

In any decision problem to make a choice implies some costs. The Neyman-Pearson test is useful in situations where these costs are difficult to quantify, since it works directly with the false alarm and detection probabilities. One important property of the Neyman-Pearson test is that it is optimum in the following sense: given a fixed false alarm it provides the detector with the maximum detection rate. Let $p(\mathbf{x}_p/H_0)$ and $p(\mathbf{x}_p/H_1)$ be the conditional probability density functions given the classes H_0 and H_1 respectively. The decision rule using the Neyman-Pearson test is formulated as follows:

$$\lambda(\mathbf{x}_p) = \frac{p(\mathbf{x}_p/H_0)}{p(\mathbf{x}_p/H_1)} \underset{H_1}{\overset{H_0}{\gtrless}} \lambda \quad (6)$$

where $\lambda(\mathbf{x}_p)$ is the likelihood ratio and λ is a suitable threshold chosen to fulfill the constraint of a fixed false alarm. In our experiments $p(\mathbf{x}_p/H_0)$ and $p(\mathbf{x}_p/H_1)$ have been estimated from the normalized histograms obtained from more

than 200 CIF images. These images were randomly selected from the ViBE video database [10]. One important issue in order to estimate $p(\mathbf{x}_p/H_0)$ and $p(\mathbf{x}_p/H_1)$ is the number of bins/channel of the histograms. This number is a trade off between poor accuracy for few bins and over-fitting for too many bins. Table 1 shows the number of bins/channel which have given the best results in our experiments for each color space. Figure 1 shows the ROC (Receiver Operating Characteristic) curves obtained for each skin detector. It can be seen that the ROC curves corresponding to the three dimensional color spaces are practically the same. In [11] 6822 images were used in order to estimate the conditional probabilities, however the results that they obtained are very close to ours which proves that our dataset is representative enough. Therefore it can be seen that the performance of the optimum skin detectors is independent of the color space. The small differences are consequence of the quantization of the color space. One important restriction is that the transformation between the color spaces $T(\mathbf{x}_p)$ has to be invertible. Figure 1 also shows the ROC curve for the CbCr color space. It can be noticed that the performance is lower since the transformation from any three dimensional color space to the bidimensional CbCr color space is not invertible. In [12] more examples are presented using other non-invertible color space transformations. As a conclusion, in this paper we have proven that if an optimum skin detector is designed for every color space, then their performance will be the same.

5. REFERENCES

- [1] A. Albiol, L. Torres, C.A. Bouman, and E. J. Delp, "A simple and efficient face detection algorithm for video database applications," in *Proceedings of the IEEE International Conference on Image Processing*, Vancouver, Canada, September 2000, vol. 2, pp. 239–242.
- [2] H. Wang and S-F. Chang, "A highly efficient system for automatic face region detection in mpeg video," *IEEE Transactions on circuits and system for video technology*, vol. 7, no. 4, pp. 615–628, August 1997.
- [3] D. Chai and K. N. Ngan, "Foreground background video coding using H.261," in *SPIE Visual Communications and Image Processing*, San Jose, CA, January 1998, pp. 434–445.
- [4] K. Sobottka and I. Pitas, "Face localization and facial feature extracion based on shape and color information," in *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, September 1996, vol. 3, pp. 236–241.

- [5] D. Saxe and R. Foulds, "Towards robust skin identification in video images," in *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, Killinton, Vermont, October 1996, pp. 379–384.
- [6] J. G. Wang and E. Sung, "Frontal-view face detection and facial feature extraction using color and morphological operators," *Pattern recognition letters*, vol. 20, no. 10, pp. 1053–1068, October 1999.
- [7] Ming-Hsuan Yang and Narendra Ahuja, "Detecting human faces in color images," in *Proceedings of the International Conference on Image Processing*, Chicago, IL, October 4-7 1998, pp. 127–130.
- [8] C. H. Lee, J. S. Kim, and K. H. Park, "Automatic human face location in a complex background using motion and color information," *Pattern recognition*, vol. 29, no. 11, pp. 1877–1889, 1996.
- [9] M. D. Srinath, P. K. Rajasekaran, and R. Viswanathan, *Introduction to statistical signal processing with applications*, Prentice-Hall, 1996.
- [10] J. Chen, C. M. Taskiran, A. Albiol, C. A. Bouman, and E. J. Delp, "Vibe: A video indexing and browsing environment," in *IEEE- SPIE Conference on Multimedia Storage and Archiving Systems IV*, Boston (USA), September 1999.
- [11] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," Technical report, Cambridge Res. Lab., Compaq Comput. Corp., 1998.
- [12] J. Brand and J. S. Mason, "A comparative assessment of three approaches to pixel-level human skin-detection," in *Proceedings of the International Conference on Pattern Recognition*, Barcelona, Spain, September 2000, pp. 1056–1059.

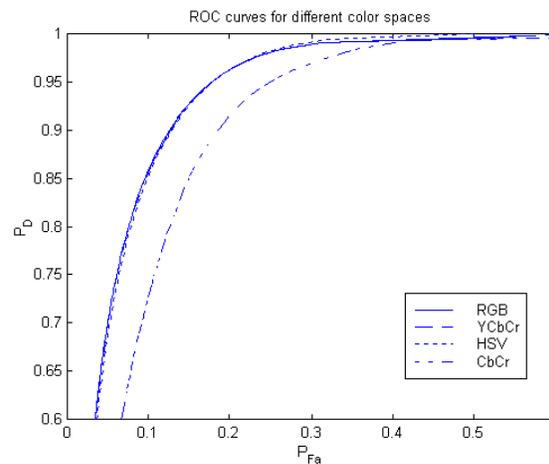


Fig. 1. ROC curves for different color spaces