2—2 Analysis of Human Skin Color Images for a Large Set of Color Spaces and for Different Camera Systems

†Jean-Christophe Terrillon, †Arnaud Pilpré, †Yoshinori Niwa and ‡Kazuhiko Yamamoto
†Office of Regional Intensive Research Project (HOIP), Softopia Japan Foundation,
4-1-7 Kagano, Ogaki-City, Gifu 503-8569, Japan
{terrillon, pilpre, niwa}@hoip.jp
‡Faculty of Engineering, Gifu University, Yanagido, Gifu-City, Gifu 501-1193, Japan
yamamoto@info.gifu-u.ac.jp

Abstract

Human skin color is a powerful fundamental cue that can be used in particular, at an early stage, for the important applications of face and hand detection in color images [1] [2], and ultimately, for meaningful human-computer interactions. In this paper, we analyze the distribution of human skin for a large number of chrominance spaces and for skin images recorded with two different camera systems. By use of seven different criteria, we show that mainly the normalized r-g and CIE-xy spaces, or a space constructed with a suitable linear combination of normalized r, g and b values, are the most efficient for skin color-based image segmentation. In particular, they allow the use of a simple, single Gaussian skin chrominance model, and they yield the most robust skin distribution to a change of camera system.

1. Introduction

Image segmentation based on skin color (or skin pixel detection) is relatively robust to changes in illumination, in viewpoint, in scale, to shading, partial occlusions and to cluttered backgrounds as compared to the segmentation of gray-level images. Robustness of segmentation is generally achieved by separating the chrominance from the luminance in the original RGB color image, and by using only the chrominance for segmentation. This separation implies a dimensionality reduction by a suitable, linear or non-linear transformation from the 3-D RGB color space into a 2-D chrominance space. One important issue is the selection of an efficient chrominance space, which motivates an analysis of human skin color for different chrominance spaces: the performance of face and hand detection depends critically on the performance of segmentation, which in turn ultimately depends on the chrominance space that is used. In particular, for a given set of skin sample images or sample pixels that is collected for calibration before applying the segmentation, the space that is selected determines the shape of the skin chrominance distribution, which in turn determines the complexity of the skin chrominance model that is required in order to obtain a high quality of segmentation. The skin chrominance distribution also depends on the various skin groups that are considered (Asians, Caucasians, dark skin groups), the illumination conditions under which the color images were recorded, and on the camera system that is used to record the images. Finally, an important criterion that ultimately limits the quality of segmentation or of skin pixel detection is the degree of overlap, or discrimination, between the skin distribution and a distribution of "non-skin" pixels in a given chrominance space, which depends to some extent on the number of skin and non-skin pixels that are collected for calibration. To our knowledge, no in-depth comparative study of the efficiency of a large set of different chrominance spaces for skin color-based image segmentation has been performed, although we recently presented preliminary results for a relatively limited number of spaces [3].

In this paper, we analyze the distribution of human skin for a large set of 25 different color spaces (41 chrominance spaces), for facial skin images recorded with two different camera systems, and in terms of seven different criteria. The color spaces considered here that result from a linear transformation from the RGB space are the $I_1I_2I_3$ (Ohta's optimized color features [4]), $h_1h_2h_3$ (Wesolkowski's color space [5]), YCb_1Cr_1 (using the CIE standard illuminant C) and YCb_2Cr_2 (using the CIE standard illuminant D65) [6] [7], YES (a standard space developed by the Xerox company), YIQ and YUV spaces. The color spaces that result from a non-linear transformation form a second group, that is divided into 4 sub-groups: the normalized color spaces (r-g-b [8] [9], CIE-xyz [8] [9] for both standard C and D65 illuminants, and TSL [10]), the perceptually plausible color spaces (CIE-DSH [8], HSV and HSL [11]), the perceptually uniform color spaces (CIE-L*u*v* [8], CIE-L*a*b* [8], and Farnsworth's F-uv space [8] [12], for both standard C and D65 illuminants), and other color spaces ($C_1C_2C_3$, $l_1l_2l_3$, and $l_1'l_2'l_3'$ proposed as color invariants and used for viewpoint-invariant image retrieval and for color-based object recognition by Smeulders and Gevers in [13] [14], r.g and rg.b log-opponent space applied to color image indexing by Berens and Finlayson in [15], a'-b' space applied to the extraction of skin color areas in facial images by Kawato and Ohya in [16], mod-rgb space proposed by Tominaga in [17], P_1P_2 space used for the construction of the Fourier spectrum of the chromaticity by Vertan *et al.* in [18], and (R/G, R/B, G/B) and Yuv spaces). The conversions from the RGB space for both groups are shown in Tables 1-5. These Tables also show the boundaries of each space, as well as the dimensions used to calculate the discrete skin chrominance histogram for each space. For all chrominance spaces considered in this paper, the histogram dimensions are selected such that the histogram resolution is the same for all spaces, in order to ensure a valid comparative study.

Table 1. Linear Transformations from RGB color space.

COLOR SPACE	CONVERSION EQUATIONS	SPACE BOUNDARIES	HISTOGRAM No OF BINS
$I_1I_2I_3$ [4]	$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1 & 0 & -1 \\ -1/2 & 1 & -1/2 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$	$I_1 \in [0; 1.0]$, $I_2 \in [-1.0; 1.0]$, $I_3 \in [-1.0; 1.0]$	200 x 200
$h_1h_2h_3$ [5]	$\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$	$h_i \in [-1.0; 1.0]$, $i = 1, 2, 3$ $h_1+h_2+h_3 = 0$	200 x 200
YCb_1Cr_1 [6], [7]	$Y = 0.2989R+0.5866G+0.1145B$ $Cb_1 = \frac{0.8855B-0.2989R-0.5866G}{1.771}$ $Cr_1 = \frac{0.7011R-0.5866G-0.1145B}{1.4022}$	$\lvert Cb_1 \rvert \le 1/2$, $\lvert Cr_1 \rvert \le 1/2$	100 x 100
YCb_2Cr_2 [6], [7]	$Y = 0.2126R+0.7152G+0.072B$ $Cb_2 = \frac{0.9278B-0.2126R-0.7152G}{1.8556}$ $Cr_2 = \frac{0.7874R-0.7152G-0.0722B}{1.5748}$	$\lvert Cb_2 \rvert \le 1/2$, $\lvert Cr_2 \rvert \le 1/2$	100 x 100
YES	$\begin{pmatrix} Y \\ E \\ S \end{pmatrix} = \begin{pmatrix} 0.253 & 0.684 & 0.063 \\ 0.500 & -0.500 & 0.000 \\ 0.250 & 0.250 & -0.500 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$	$\lvert E \rvert \le 1/2$, $\lvert S \rvert \le 1/2$	100 x 100
YIQ	$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$	$\lvert I \rvert \le 0.596$, $\lvert Q \rvert \le 0.523$	120 x 120
YUV	$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$	$\lvert U \rvert \le 0.436$, $\lvert V \rvert \le 0.615$	125 x 125

Table 2. Non-linear conversions from RGB color space into normalized color spaces.

COLOR SPACE	CONVERSION EQUATIONS	SPACE BOUNDARIES	HISTOGRAM No OF BINS
rgb [8], [9]	$r = \frac{R}{R+G+B}$, $g = \frac{G}{R+G+B}$, $b = \frac{B}{R+G+B}$ If R=G=B=0, set r=g=b=1/3	$(r, g, b) \in [0; 1.0]$ $r + g + b = 1$	100 x 100
CIE-xyz (C Ill.) [8], [9]	$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.607 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$		
CIE-xyz (D65 Ill.) [8], [9]	$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.430574 & 0.341550 & 0.178325 \\ 0.222015 & 0.706655 & 0.071330 \\ 0.020183 & 0.129553 & 0.939180 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$		100 x 100
	$x = \frac{X}{X+Y+Z}$, $y = \frac{Y}{X+Y+Z}$, $z = \frac{Z}{X+Y+Z}$ If X=Y=Z=0, set x=y=z=1/3	$(x, y, z) \in [0; 1.0]$ $x + y + z = 1$	
TSL [10]	$g' = (g-1/3)$, $r' = (r-1/3)$ $T = \begin{cases} \frac{1}{2\pi}\arctan\left(\frac{r'}{g'}\right)+\frac{1}{4}, & g' > 0 \text{ and } r' \neq 0 \\ \frac{1}{2\pi}\arctan\left(\frac{r'}{g'}\right)+\frac{3}{4}, & g' < 0 \text{ and } r' \neq 0 \\ \frac{1}{4}, & g' > 0 \text{ and } r' = 0 \\ \frac{3}{4}, & g' < 0 \text{ and } r' = 0 \\ \frac{1}{2}, & g' = 0 \text{ and } r' > 0 \\ 0, & g' = 0 \text{ and } r' \le 0 \end{cases}$ $S = \left[\frac{9}{5}(r'^2 + g'^2)\right]^{1/2}$ $L = 0.299R + 0.587G + 0.114B$	$(T, S, L) \in [0; 1.0]$	100 x 100

Table 3. Non-linear conversions from RGB color space into perceptually plausible color spaces.

COLOR SPACE	CONVERSION EQUATIONS	SPACE BOUNDARIES	HISTOGRAM No OF BINS
CIE-DSH [8]	$D = \frac{1}{3}(R + G + B)$ $S = 1 - \frac{3[\min(R, G, B)]}{(R + G + B)}$	$(D, S, H) \in [0; 1.0]$	100 x 100
	$H = \begin{cases} \frac{1}{2\pi}\arccos\left\{\frac{1/2[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\} + \frac{1}{2}, & G > B \\ \frac{1}{2}, & G = B \\ \frac{1}{2} - \frac{1}{2\pi}\arccos\left\{\frac{1/2[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\}, & G < B \end{cases}$ if R=G=B, S=0, set H=0. If R=G=B=0, set S=H=0		
HSV	$H = \begin{cases} \frac{G - B}{\max(R,G,B) - \min(R,G,B)}, & R = \max(R,G,B) \\ 2 + \frac{B - R}{\max(R,G,B) - \min(R,G,B)}, & G = \max(R,G,B) \\ 4 + \frac{R - G}{\max(R,G,B) - \min(R,G,B)}, & B = \max(R,G,B) \end{cases}$ Normalize by setting H = H/6. If H < 0, set H = H + 1 $S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}$, $\max(R,G,B) \neq 0$ $V = \max(R, G, B)$ If max(R,G,B)=min(R,G,B), then R=G=B, S=0, set H=0 If max(R,G,B)=0, then R=G=B=0, set S=H=0	$(H, S, V) \in [0; 1.0]$	100 x 100
HSL [11]	$H = \begin{cases} \text{same as for HSV} \end{cases}$ $S = \begin{cases} \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B) + \min(R,G,B)}, & L \le 0.5 \\ \frac{\max(R,G,B) - \min(R,G,B)}{2 - [\max(R,G,B) + \min(R,G,B)]}, & L > 0.5 \end{cases}$ $L = \frac{\max(R,G,B) + \min(R,G,B)}{2}$ Same limiting conditions as for HSV - if R=G=B=1, set S=H=0	$(H, S, L) \in [0; 1.0]$	100 x 100

Table 4. Non-linear conversions from RGB color space into perceptually uniform color spaces.

COLOR SPACE	CONVERSION EQUATIONS	SPACE BOUNDARIES	HISTOGRAM No OF BINS
CIE-L*u*v* [8]	$L^* = \begin{cases} 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16, & \frac{Y}{Y_n} > 0.008856 \\ 903.3\left(\frac{Y}{Y_n}\right), & \frac{Y}{Y_n} \le 0.008856 \end{cases}$ $u^* = 13\,L^*(u' - u'_n)$, $v^* = 13\,L^*(v' - v'_n)$ $u' = \frac{4X}{X + 15Y + 3Z}$, $v' = \frac{9Y}{X + 15Y + 3Z}$ $u'_n = \frac{4X_n}{X_n + 15Y_n + 3Z_n}$, $v'_n = \frac{9Y_n}{X_n + 15Y_n + 3Z_n}$ (X_n, Y_n, Z_n) are the CIE (X, Y, Z) components of a reference illuminant, e.g. the C or D65 illuminants C Ill.: $\begin{pmatrix} X_n \\ Y_n \\ Z_n \end{pmatrix} = \begin{pmatrix} 0.982 \\ 1.000 \\ 1.183 \end{pmatrix}$ D65 Ill.: $\begin{pmatrix} X_n \\ Y_n \\ Z_n \end{pmatrix} = \begin{pmatrix} 0.950 \\ 1.000 \\ 1.089 \end{pmatrix}$ Achromatic point : $u^* = v^* = 0$;	$0 \le L^* \le 100$ Boundaries curved in u^*-v^* space	200 x 200 (skin) or 370 x 370 (non-skin)
CIE-L*a*b* [8]	$L^* = \begin{cases} \text{same as in CIE-L*u*v* space} \end{cases}$ $a^* = 500\left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right)\right]$, $b^* = 200\left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right)\right]$, $f(t) = \begin{cases} t^{1/3}, & t \le 0.008856 \\ 7.787t + \frac{16}{116}, & t > 0.008856 \end{cases}$ $(X_0, Y_0, Z_0) = (X_n, Y_n, Z_n)$	$0 \le L^* \le 100$ Boundaries curved in a^*-b^* space	200 x 200 (skin) or 370 x 370 (non-skin)
F-uv [8], [12]	$U = \frac{2X}{3}$, $V = Y$, $W = \frac{-X+3Y+Z}{2}$ $u_f = \frac{U}{U+V+W} = \frac{4X}{X+15Y+3Z} = \frac{4x}{-2x+12y+3}$ $v_f = \frac{V}{U+V+W} = \frac{6Y}{X+15Y+3Z} = \frac{6y}{-2x+12y+3}$	$u_f \in [0; 4.0]$ $v_f \in [0; 0.4]$	100 x 100

Table 5. Non-linear conversions from RGB color space into other color spaces.

COLOR SPACE	CONVERSION EQUATIONS	SPACE BOUNDARIES	HISTOGRAM No OF BINS
$C_1C_2C_3$ [13]	$C_1 = \frac{2}{\pi}\arctan\left\{\frac{R}{\max\{G, B\}}\right\}$, $C_2 = \frac{2}{\pi}\arctan\left\{\frac{G}{\max\{R, B\}}\right\}$, $C_3 = \frac{2}{\pi}\arctan\left\{\frac{B}{\max\{R, G\}}\right\}$ If R=G=B, C_i = 1/2. If R=G=B=0, set C_i = 0	$C_i \in [0; 1.0]$ $i = 1, 2, 3$	100 x 100
$I_1I_2I_3$ [13]	$I_1 = \frac{\lvert R-G\rvert}{\lvert R-G\rvert+\lvert R-B\rvert+\lvert G-B\rvert}$, $I_2 = \frac{\lvert R-B\rvert}{\lvert R-G\rvert+\lvert R-B\rvert+\lvert G-B\rvert}$, $I_3 = \frac{\lvert G-B\rvert}{\lvert R-G\rvert+\lvert R-B\rvert+\lvert G-B\rvert}$ If R=G=B or R=G=B=0, set I_i=1/3	$I_i \in [0; 1/2]$ $i = 1, 2, 3$ $I_1 + I_2 + I_3 = 1$	100 x 100
$I'_1I'_2I'_3$ [13], [14]	$I_1' = \frac{(R-G)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$, $I_2' = \frac{(R-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$, $I_3' = \frac{(G-B)^2}{(R-G)^2+(R-B)^2+(G-B)^2}$ If R=G=B or R=G=B=0, set I'_i=1/3	$I_i' \in [0; 2/3]$ $i = 1, 2, 3$ $I_1' + I_2' + I_3' = 1$	100 x 100
Ln-Chroma r_g, rg_b [15]	$r\_g = \ln\left(\frac{R}{G}\right) = \ln R - \ln G = R' - G'$ $rg\_b = \ln\left(\frac{RG}{B^2}\right) = \ln R + \ln G - 2\ln B = R' + G' - 2B'$	$r\_g \in]-\infty;+\infty[$ $rg\_b \in]-\infty;+\infty[$	300 x 300*
a' - b' [16]	$a' = r + \frac{g}{2}$, $b' = \frac{\sqrt{3}}{2}g$	$a' \in [0; 1.0]$ $b' \in [0; \frac{\sqrt{3}}{2}]$	100 x 100
mod-rgb [17]	Let $I = \sqrt{R^2+G^2+B^2}$, then $m\_r = R/I$, $m\_g = G/I$, $m\_b = B/I$ $(m\_r)^2 + (m\_g)^2 + (m\_b)^2 = 1$ If R=G=B=0, set $m\_r=m\_g=m\_b=\sqrt{3}/3$	$m\_r \in [0;1.0]$ $m\_g \in [0;1.0]$ $m\_b \in [0;1.0]$	100 x 100
P_1P_2 [18]	$P_1 = \frac{1}{\sqrt{2}}\frac{G-R}{R+G+B}$, $P_2 = \frac{1}{\sqrt{6}}\frac{2B-R-G}{R+G+B}$ If R=G=B=0, set $P_1=P_2=0$	$\lvert P_1\rvert \le 1/\sqrt{2}$ $-1/\sqrt{6} \le P_2 \le 2/\sqrt{6}$	150 x 150
$R_1R_2R_3$	$R_1 = G/R$, $R_2 = B/R$, $R_3 = B/G$ If R=G=B=0, set $R_1=R_2=R_3=0$	$R_i \in [0; \infty[$ $i = 1, 2, 3$	200 x 200*
Yuv	$u = \frac{U}{Y} = \frac{-0.147R-0.289G+0.436B}{0.299R+0.587G+0.114B} = \frac{-0.583r-0.725g+0.436}{0.185r+0.473g+0.114}$ $v = \frac{V}{Y} = \frac{0.615R-0.515G-0.100B}{0.299R+0.587G+0.114B} = \frac{0.715r-0.415g-0.100}{0.185r+0.473g+0.114}$	$u \in [\frac{-289}{587}, \frac{436}{114}]$ $v \in [\frac{-515}{587}, \frac{615}{299}]$	100 x 100 (skin) or 440 x 440 (non-skin)

Two separate sets of sample images used for the skin chrominance analysis are recorded with an inexpensive SGI camera, and with a high-quality SONY DXC-9000 camera system respectively. The seven criteria used for the analysis for each space are: 1) the robustness of the skin chrominance distribution with respect to the intrinsic variability of skin color (to three different skin groups), 2) its compactness, 3) its shape, 4) the degree of discrimination (or the overlap) between the skin and non-skin distributions, 5) the robustness (or "portability") of the skin distribution to a change of camera system, 6) the relative robustness of the skin distribution to changes in illumination conditions, and finally, 7) the computational cost of the transformation from the 24-bit RGB (NTSC) space into a given chrominance space.

2. Parameters used for the Skin Chrominance Analysis

We first define three different parameters that we use to perform a quantitative analysis of the skin chrominance.

1) The Kullback-Leibler Divergence (KLD) [3] is selected to estimate the goodness of fit of the skin chrominance distribution to a simple, single elliptical Gaussian. It is defined in the discrete case as :

$$KLD = \sum_{i=1}^{N}\sum_{i=1}^{M} S'_{ij} \ln\left(\frac{S'_{ij}}{G'_{ij}}\right) + \sum_{i=1}^{N}\sum_{i=1}^{M} G'_{ij} \ln\left(\frac{G'_{ij}}{S'_{ij}}\right) \quad (1)$$

where S'_{ij} is considered as the "true" distribution (the normalized skin histogram observed in a discrete chrominance space with M x N bins) and G'_{ij} as the "estimated" or "model" distribution (the normalized ideal Gaussian histogram calculated from the mean vector and from the covariance matrix of the skin distribution in the same discrete space). The KLD has the following properties: i) KLD ≥ 0 and ii) if KLD=0, then $S'_{ij} = G'_{ij}$. Hence, the lower the value of the KLD, the higher the goodness of fit to the single Gaussian model.

2) The Normalized Histogram Intersection (HIN) is a measure of the overlap between two different distributions, such as the skin and non-skin distributions. In the discrete case, it is defined as:

$$HIN = \sum_{i=1}^{N}\sum_{i=1}^{M} \min(S'_{ij}, NS'_{ij}) \quad (2)$$

where $NS'_{ij} = NS_{ij} / \sum_{i=1}^{N}\sum_{i=1}^{M} NS_{ij}$ is the normalized non-skin histogram calculated in the same discrete chrominance space as S'_{ij}, with M x N bins. The lower the value of the HIN, the higher the degree of discrimination between the two distributions.

3) The global shift S of a distribution can be calculated as:

$$S = \sqrt{(m_{x1} - m_{x2})^2 + (m_{y1} - m_{y2})^2} \quad (3)$$

where $\mathbf{m_x} = (m_{x1}, m_{x2})$ and $\mathbf{m_y} = (m_{y1}, m_{y2})$ are the mean vectors for the skin distributions in a given chrominance space (x, y) for cameras 1 and 2.

3. Skin Chrominance Analysis

3.1 Experimental set-up

Images of Asian and Caucasian subjects, and of subjects with dark skin color, were recorded under slowly varying illumination conditions in an office environment with both the SONY and the SGI camera systems. From the images obtained with the SONY camera, 65, 51 and 10 skin sample images of Asian, Caucasian, and dark skin-colored subjects respectively, were manually selected, yielding a total of 2.115x10E+05, 1.630x10E+05, and 2.580x10E+04 skin pixels for each respective skin group. When using the SGI camera, 111 skin sample images of both Asian and Caucasian subjects were manually selected, for a total of 1.515x10E+05 skin pixels. Also, 80 "non-skin" images were selected from various sources, mainly from the World Wide Web, producing a total of 2.6606x10E+06 non-skin pixels. For each image, the 24-bit RGB values are scaled between 0.0 and 1.0. Achromatic pixels (including black) were assigned suitable values adapted to the particular color space that is considered, as shown in Tables 1-5. For each

space yielding negative chrominance values, a shift was performed so that all values are positive, without any influence on the results of the chrominance analysis. Generally, the discrete, cumulative skin and non-skin histograms are calculated over an entire space, except for the CIE-L*u*v and CIE-L*a*b* spaces, whose boundaries are curved, and for the log-opponent and $R_1R_2R_3$ spaces, where the range (hence the histogram dimensions) is determined empirically, by observing the skin and non-skin distributions (we used a range of [-1.0; 2.0] along both the x and y axes for the log-opponent space, and of [0.0;2.0] for the $R_1R_2R_3$ space). The resolution of the skin and non-skin histograms is 0.01 unit for all spaces, except for the CIE-L*u*v* and CIE-L*a*b* spaces, where the resolution is 1.0 unit.

3.2 Robustness to the intrinsic variability of skin color and compactness of the skin distribution

As an example, Figure 1 shows the skin distribution separately for each of the three classes of subjects for the r-g, CIE-xy and H-S (HSV) spaces with the SONY camera. Table 6 shows the KLD and the HIN for the three skin groups for all spaces, when using the SONY camera. Table 7 shows, for each space and for both cameras, the area of the skin distribution relative to the area of the gamut of possible colors (owing to the particular boundaries of the CIE-L*u*v*, CIE-L*a*b*, r.g-rg.b and $R_1R_2R_3$ spaces, the area for these spaces is not shown). The gamut in all spaces with rectangular boundaries, except in the CIE-DSH, HSV and HSL spaces, fills only a part of the entire space defined by the space boundaries, and its geometry depends on the space that is considered, as Figure 2 illustrates by use of the non-skin distributions for a selected number of different spaces. The normalized CIE-xy and r-g spaces, as well as the a'-b', mod-gb and P_1-P_2 spaces yield the most robust distributions with respect to the intrinsic variability of skin color, because: 1) the KLD is consistently lower across the three skin groups than for the other spaces, and 2) the overlap between the skin groups varies typically within a relatively narrow range, between 45% and 64% for most spaces. The $l_1l_2l_3$, $l_1'l_2'l_3'$ and C_1-C_2 spaces yield the highest overlap, indicative of a higher robustness, but this advantage is offset by the large overlap between the skin and non-skin distributions in those spaces, as shown in Subsection 3.4. In almost all chrominance spaces, the distribution for the Asian subjects, who have an intermediate skin color, is the most compact, in terms of the relative area of the distribution, in particular in the r-g, CIE-xy, C_1-C_2, P_1-P_2, a'-b' and mod-rgb spaces.

3.3 Shape of the skin distribution

A few representative examples of the cumulative distribution for the Asian + Caucasian subjects, obtained with both cameras, are shown in Figure 3, for the r-g, CIE-xy, E-S and CIE-u*v* chrominance spaces. Visually, the skin distribution in the normalized spaces fits well to the single Gaussian model, whereas in the un-normalized spaces, its shape is generally complex and cannot be described well by a simple model. Table 8 shows the KLD for all the chrominance spaces and for both cameras. Since the KLD is consistently lowest for the normalized CIE-xy and r-g spaces, together with the C_2-C_3, a'-b',

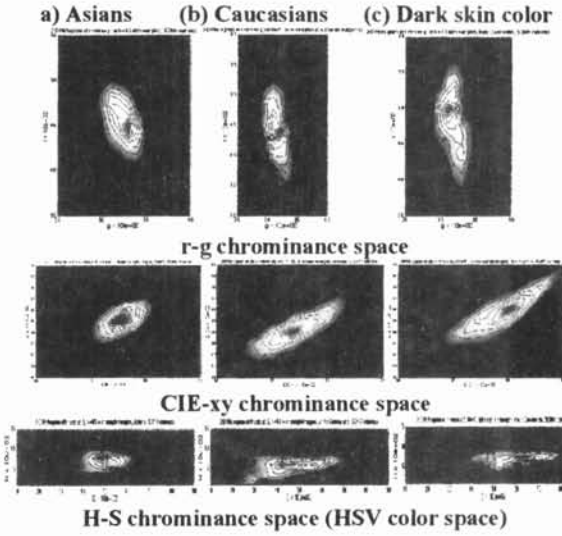

Figure 1. 2-D top view of the cumulative histograms in the normalized r-g (top), CIE-xy (middle) chrominance spaces and in the H-S (HSV) space (bottom) of skin sample images of (a) Asian, (b) Caucasian subjects, and (c) of subjects with dark skin color, recorded with the SONY camera. Here, only the relevant part of the histogram is shown.

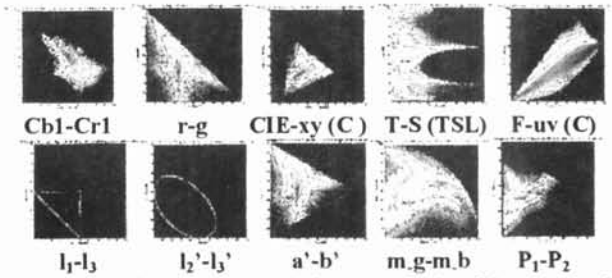

Figure 2. 2-D top view of the cumulative histograms in several selected chrominance spaces of 80 non-skin sample images collected from various sources.

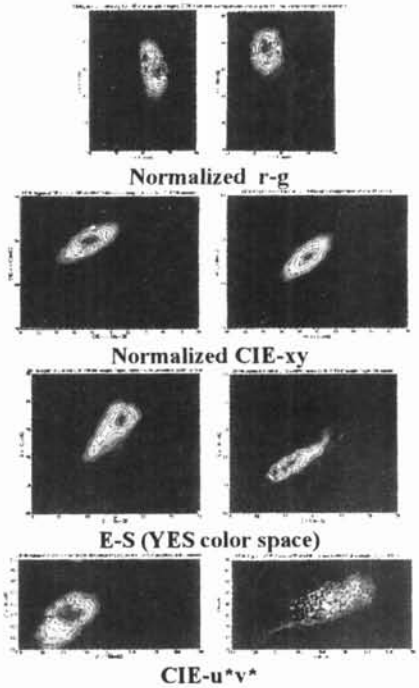

Figure 3. 2-D top view of the cumulative histograms in the r-g, CIE-xy, E-S (YES) and CIE-u*v* chrominance spaces of skin sample images of Asian + Caucasian subjects recorded with the SONY camera (left column) and with the SGI camera (right column).

Table 6. KLD and HIN for three different skin groups (Asians (A), Caucasians (C) and subjects with dark skin color (D)), for 41 chrominance spaces and with the SONY camera.

	KULLBACK-LEIBLER DIVERGENCE			OVERLAP (HIN) BETWEEN DIFFERENT SKIN GROUPS		
COLOR SPACE	A	C	D	A/C	A/D	C/D
I_2J_3	1.1771	0.9658	2.2285	0.5919	0.6259	0.5349
h_1h_2	1.3536	0.9798	2.2138	0.5934	0.6255	0.5363
h_2h_3	1.1237	0.9539	2.4292	0.5943	0.6244	0.5395
h_1h_3	1.1950	0.8987	2.7134	0.5935	0.6285	0.5355
Cb_1Cr_1	1.0569	0.7582	2.3620	0.6034	0.6464	0.5499
Cb_2Cr_2	1.0990	0.7198	2.1675	0.6012	0.6489	0.5511
E-S	1.1008	0.7440	1.6099	0.6025	0.6475	0.5528
I-Q	1.1613	0.6412	1.8226	0.6030	0.6415	0.5620
U-V	1.0696	0.7073	2.1329	0.6034	0.6495	0.5561
r-g	0.4878	0.6690	1.7057	0.5010	0.4739	0.5399
CIE-xy (C Ill.)	0.3771	0.5269	0.9205	0.5031	0.4882	0.5661
CIE-xy (D65 Ill.)	0.4323	0.4906	0.8467	0.5103	0.5066	0.5692
TSL	10.0736	4.6134	17.2312	0.4909	0.4681	0.5289
CIE-DSH	0.8066	8.7930	18.7132	0.4929	0.4688	0.5284
HSV	2.7220	20.0670	17.7311	0.4922	0.4675	0.5271
HSL	1.7272	18.6020	17.7355	0.4780	0.5537	0.6204
CIE-L*u*v* (C Ill.)	0.2653	0.4121	3.3201	0.5535	0.5456	0.5905
CIE-L*u*v* (D65 Ill.)	0.2828	0.2832	2.8929	0.5944	0.5275	0.6503
CIE-L*a*b* (C Ill.)	0.1979	0.6845	3.4265	0.5283	0.5308	0.5742
CIE-L*a*b* (D65 Ill.)	0.2137	0.3792	2.4141	0.5679	0.5111	0.6395
F-uv (C Ill.)	1.6057	4.8038	1.6153	0.5322	0.4002	0.3969
F-uv (D65 Ill.)	1.6348	4.0665	1.0664	0.5572	0.4156	0.3936
C_1C_2	23.6305	26.1568	29.4019	0.7683	0.5471	0.6909
C_2C_3	0.2057	2.0942	5.8982	0.4947	0.4690	0.5355
C_1C_3	0.6431	12.8966	16.1455	0.4947	0.4692	0.5366
I_1J_2	1.9433	20.1158	28.7567	0.6756	0.8200	0.7159
I_2J_3	1.9335	19.9077	28.7633	0.6757	0.8189	0.7188
I_1J_3	28.4289	21.3616	29.0135	0.6808	0.8330	0.7172
$I'_1J'_2$	20.6039	26.6590	30.5420	0.6823	0.8217	0.7145
$I'_2J'_3$	22.4415	28.3576	31.0432	0.6825	0.8331	0.7200
$I'_1J'_3$	20.8104	26.2118	29.8854	0.6821	0.8212	0.7175
r.g-rg.b	0.4316	4.3017	10.6025	0.4764	0.4285	0.4766
a' - b'	0.4282	0.6836	1.5669	0.4995	0.4820	0.5444
m.r-m.g	0.5026	0.5638	3.6269	0.4988	0.4709	0.5334
m.g-m.b	0.1916	0.6016	1.4755	0.4890	0.4658	0.5246
m.r-m.b	0.3913	1.8047	5.6624	0.4946	0.4714	0.5352
P_1P_2	0.1613	0.6085	1.6328	0.4928	0.4698	0.5331
R_1R_2	0.2508	0.6462	7.9108	0.4853	0.4575	0.5121
R_2R_3	0.3881	2.1389	11.7300	0.4850	0.4590	0.5115
R_1R_3	0.3101	1.2686	9.8428	0.4841	0.4535	0.5020
Yuv	0.9517	3.0968	3.7347	0.5109	0.3532	0.3559

mg-mb and P_1P_2 spaces, the skin distribution in those spaces can be modeled by a single Gaussian.

3.4 Discrimination between skin and non-skin distributions

Table 8 shows the HIN for Asian + Caucasian subjects, for all the chrominance spaces and for both camera systems. For both camera systems, the overlap between the skin and non-skin distributions is lowest for the r-g, CIE-xy, TSL, CIE-DSH, HSV, CIE-L*u*v*, CIE-L*a*b*, C_2-C_3, r.g-rg.b (ln-chroma), a'-b', mod-rgb, P_1P_2 and $R_1R_2R_3$ spaces. Hence, the discrimination capabilities between skin pixels and non-skin pixels are highest in these spaces. The lowest discrimination is found for the $I_1I_2I_3$ and $I_1'I_2'I_3'$ spaces.

Table 7. Area of the skin distributions for the three different skin groups relative to the area of the gamut of all possible colors in a given color space, for both camera systems.

	AREA OF SKIN DISTRIBUTIONS (% OF ALL SPACE)					
	SONY DXC-9000					SGI
COLOR SPACE	A	C	D	A+C	A+C+D	A+C
I_2J_3	6.279	8.396	5.975	9.518	10.388	18.040
h_1h_2	6.273	8.465	5.916	9.535	10.469	17.707
h_2h_3	6.346	8.528	5.948	9.682	10.563	17.644
h_1h_3	6.294	8.486	5.811	9.556	10.416	17.749
Cb_1Cr_1	6.499	8.945	6.499	10.092	11.124	23.012
Cb_2Cr_2	6.544	8.988	6.581	10.041	10.982	22.828
E-S	6.558	8.804	6.522	9.783	10.725	22.971
I-Q	6.481	8.844	6.447	9.882	10.813	23.022
U-V	6.505	9.166	6.505	10.209	11.107	23.077
r-g	3.260	8.900	10.000	9.080	13.140	12.000
CIE-xy (C Ill.)	4.078	12.234	12.707	12.411	17.317	14.894
CIE-xy (D65 Ill.)	4.278	12.752	13.156	12.833	17.918	17.111
TSL	6.188	15.103	18.450	15.648	24.074	16.005
CIE-DSH	5.380	12.600	15.420	13.120	20.390	14.050
HSV	6.470	14.730	18.500	15.420	25.160	15.430
HSL	7.240	15.380	15.370	15.910	23.100	20.020
F-uv (C Ill.)	12.622	19.804	20.131	20.131	26.551	26.659
F-uv (D65 Ill.)	12.349	19.329	20.269	19.597	25.638	28.993
C_1C_2	0.760	3.260	7.160	3.280	8.540	11.260
C_2C_3	6.680	13.760	12.940	15.600	18.040	20.620
C_1C_3	6.400	15.160	16.440	14.180	20.720	10.800
I_1J_2	27.419	53.226	56.129	53.226	58.710	69.355
I_2J_3	27.419	54.839	57.097	54.839	60.645	61.290
I_1J_3	25.807	53.548	55.484	53.871	60.000	65.807
$I'_1I'_2$	35.094	72.453	75.472	72.453	84.906	89.434
$I'_2I'_3$	30.566	67.547	73.585	68.302	82.264	84.151
$I'_1I'_3$	39.623	71.698	76.981	72.830	83.774	88.302
a' - b'	3.349	9.261	10.370	9.423	13.764	12.494
m.r-m.g	3.489	6.685	10.008	7.550	11.752	11.879
m.g-m.b	5.895	14.311	14.337	14.680	19.430	18.118
m.r-m.b	4.469	10.122	9.740	12.401	12.270	15.623
P_1P_2	2.991	7.979	8.718	8.153	11.571	10.347
Yuv	12.283	19.901	15.927	20.581	24.507	30.833

3.5 Robustness to a change of camera system

The robustness to a change of camera system can be measured as the change in the KLD, in the HIN and the global shift S of the distribution. As seen from Table 8, the change in the KLD is lowest for the r-g, CIE-xy, C_2-C_3, a'-b', mod-rgb and P_1P_2 spaces, while the overlap of the skin distributions between the two camera systems (HIN skin SONY/SGI) is intermediate to low for those spaces. The highest overlap is found for the $I_1I_2I_3$, and $I_1'I_2'I_3'$ but, as for the overlap between the three different skin groups, for those spaces this advantage is offset by a significant overlap between the skin and non-skin distributions, and also by very large values of the KLD for both cameras. The global shift S of the distribution is low to lowest for the above-mentioned 6 color spaces, and is also low for some of the spaces resulting from a linear transformation from RGB space.

3.6 Robustness to changes in illumination conditions and computational cost of the color space transformation

Finally, it is well known that a normalization of RGB values by (R+G+B) or of CIE-XYZ values by (X+Y+Z)

reduces the most the sensitivity of the skin distribution to changes in illumination, and a linear transformation from the RGB space, or a non-linear conversion into the normalized rgb coordinates and into the CIE-xyz space is not computationally intensive compared to that into other spaces. The mod-rgb space also provides a suitable normalization.

4. Conclusions

In conclusion, overall, in terms of seven different criteria, the normalized r-g and CIE-xy chrominance spaces, or spaces such as the a'-b' and P_1P_2 spaces that are constructed as a linear combination of normalized r, g and b values, offer the best tradeoff and appear consistently to be the most efficient for skin color-based image segmentation. In particular, the use of these normalized spaces obviates the necessity to apply a complex and computationally intensive skin chrominance model in order to obtain a high quality of segmentation, as is the case with most un-normalized spaces, in which the skin distribution is complex-shaped. The C_2-C_3 space, the mod-rgb space that also results from a suitable normalization, and to a lesser extent the $R_1R_2R_3$ space, are also good candidates. Owing to their particular geometry, the $l_1l_2l_3$ and $l_1'l_2'l_3'$ spaces are the least efficient for the specific problem of image segmentation based on skin color.

References

[1] J.-C. Terrillon, A. Pilpré, Y. Niwa, and K. Yamamoto. Robust face detection and hand posture recognition in color images for human-machine interaction. In *Proc. of the 16th International Conference on Pattern Recognition*, Québec City, Canada, August 2002. pp. 204-209.

[2] J.-C. Terrillon, M. David, and S. Akamatsu. Automatic detection of human faces in natural scene images by use of a skin color model and of invariant moments. In *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, 1998. pp. 112-117.

[3] J.-C. Terrillon, Y. Niwa and K. Yamamoto. On the selection of an efficient chrominance space for skin color-based image segmentation with an application to face detection. In *Proceedings of the 5th International Conference on Quality Control by Artificial Vision*, Le Creusot, France, May 2001. Cépaduès Éditions. Vol. 2, pp. 409-414.

[4] Y.-I. Ohta, T. Kanade, and T. Sakai. Color information for region segmentation. *Computer Graphics and Image Processing*, 13(3):222-241, July 1980.

[5] S. Wesolkowski, M. E. Jernigan, and R. D. Dony. Comparison of color image edge detectors in multiple color spaces. In *Proceedings of the International Conference on Image Processing*, Vancouver, Canada, September 2000.

[6] CIE Colorimetry. Official recommendations of the International Commission on illumination, Publication CIE No. 15.2, Second Edition, Central Bureau of the Commission Internationale de L'Éclairage, Vienna, Austria (1986).

[7] ITU-R Recommendation BT.709, Basic Parameter Values for the HDTV Standard for the Studio and for International Programme Exchange , [formerly CCIR Rec. 709], ITU, 1211 Geneva 20, Switzerland (1990).

[8] G. Wyszecki, and W. S. Styles. Color science: concepts and methods, quantitative data and formulae, 2nd edition, John Wiley, New York, 1982

[9] R. W. G. Hunt. Measuring Colour. 2nd edition, Ellis Horwood (1991).

[10] J.-C. Terrillon, M. David, and S. Akamatsu. Detection of human faces in complex scene images by use of a skin color model and of invariant Fourier-Mellin moments. In *Proceedings of the 14th International Conference on Pattern Recognition*, Brisbane, Australia, August 1998. Volume 2, pp. 1350-1355.

[11] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. Second Edition in C Computer Graphics Principles and Practice. Addison-Wesley, New York, 1996.

[12] H. Wu, Q. Chen, and M. Yachida. Face detection from color images using a fuzzy pattern matching method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-21(6):557-563, 1999.

Table 8. KLD, HIN, and global shift of the skin distribution for Asian + Caucasian subjects, for 41 chrominance spaces, and for both the SONY and SGI cameras.

	KULLBACK-LEIBLER DIVERGENCE		HIN SKIN/NON-SKIN		HIN SKIN SONY/ SGI	GLOBAL SHIFT SONY/ SGI
COLOR SPACE	SONY DXC-9000	SGI INDIGO	SONY DXC-9000	SGI INDIGO		
I_2I_3	0.8765	4.6131	0.1889	0.1740	0.2031	0.1143
h_1h_2	0.9063	4.9569	0.1917	0.1781	0.2034	0.0936
h_2h_3	0.8218	5.7059	0.1907	0.1760	0.2013	0.1418
h_1h_3	0.7288	4.9885	0.1916	0.1763	0.2020	0.1086
Cb_1Cr_1	0.6454	1.5237	0.1967	0.1978	0.2214	0.0510
Cb_2Cr_2	0.6084	1.4554	0.1952	0.1933	0.2251	0.0503
E-S	0.7658	1.4701	0.1965	0.2008	0.2267	0.0505
I-Q	0.5926	1.2589	0.1975	0.2017	0.2314	0.0464
U-V	0.5598	1.6010	0.1988	0.1950	0.2253	0.0464
r-g	0.4482	0.2681	0.1682	0.1057	0.1487	0.0561
CIE-xy (C Ill.)	0.3768	0.2183	0.1768	0.1139	0.1784	0.0213
CIE-xy (D65 Ill.)	0.3796	0.2351	0.1817	0.1139	0.1822	0.0205
TSL	6.2430	0.6211	0.1694	0.1041	0.1466	0.0754
CIE-DSH	4.6092	9.8105	0.1635	0.1045	0.1502	0.0734
HSV	15.4888	24.5459	0.1633	0.1006	0.1383	0.2357
HSL	13.3774	24.5932	0.1949	0.1391	0.2400	0.2354
CIE-L*u*v* (C Ill.)	0.1917	1.0639	0.1715	0.1377	0.2385	12.2472*
CIE-L*u*v* (D65 Ill.)	0.1828	0.8657	0.1667	0.1291	0.1599	13.8944*
CIE-L*a*b* (C Ill.)	0.2282	0.5040	0.1740	0.1347	0.2606	9.3797*
CIE-L*a*b* (D65 Ill.)	0.1901	0.3795	0.1705	0.1238	0.1793	10.4512*
F-uv (C Ill.)	2.6016	0.8623	0.2274	0.1588	0.0950	0.1235
F-uv (D65 Ill.)	2.1259	0.8294	0.2406	0.1869	0.1007	0.1261
C_1C_2	37.7590	14.1395	0.2188	0.1575	0.1959	0.0850
C_2C_3	0.2470	0.2890	0.1618	0.0991	0.1406	0.0850
C_1C_3	1.0441	8.0846	0.1741	0.1099	0.1608	0.0319
I_1I_2	14.9248	27.1080	0.4379	0.2962	0.2742	0.1265
I_2I_3	14.8468	27.0137	0.4337	0.2959	0.2751	0.1022
I_1I_3	17.2852	27.0444	0.4341	0.2959	0.2751	0.1586
$I'_1I'_2$	25.7310	28.0667	0.4301	0.2955	0.2729	0.2217
$I'_2I'_3$	27.5210	28.8758	0.4415	0.2948	0.2730	0.1374
$I'_1I'_3$	24.7953	28.4646	0.4307	0.2947	0.2715	0.1980
r_g-rg_b	1.2728	3.6653	0.1539	0.0737	0.0943	0.2985
a' - b'	0.4483	0.2485	0.1669	0.1053	0.1531	0.0406
m_r-m_g	0.4288	0.4315	0.1654	0.1046	0.1392	0.0935
m_g-m_b	0.2839	0.2790	0.1639	0.1016	0.1362	0.0835
m_r-m_b	0.5898	0.5921	0.1659	0.1042	0.1453	0.0468
P_1P_2	0.2502	0.2208	0.1661	0.1042	0.1424	0.0575
R_1R_2	0.2876	0.7896	0.1606	0.0940	0.1293	0.1466
R_2R_3	0.6511	1.7565	0.1601	0.0914	0.1274	0.1815
R_1R_3	0.3868	1.3170	0.1582	0.0908	0.1225	0.2326
Yuv	1.4184	3.8764	0.2381	0.2578	0.2864	0.2510

[13] T. Gevers and A. W. M. Smeulders. Color-based object recognition. *Pattern Recognition*, 32(3):453-464, 1999.

[14] T. Gevers and A. W. M. Smeulders. Content-based image retrieval by viewpoint-invariant color indexing. *Image and Vision computing*, 17:475-488, 1999.

[15] J. Berens and G. D. Finlayson. Log-opponent chromaticity coding of colour space. In *Proceedings of the 15th International Conference on Pattern Recognition*, Barcelona, Spain, September 2000. pp. 206-211.

[16] S. Kawato and J. Ohya. Real-time detection of nodding and head shaking by directly detecting and tracking the between eyes. In *Proceedings of the 4th International Conference on Face and Gesture Recognition*, Grenoble, France, March 2000. pp. 40-45.

[17] S. Tominaga. Illuminant estimation of natural scenes from color images. In *Proceedings of the International Conference on Color in Graphics and Image Processing*, Saint-Etienne, France, 2000. Cépaduès Éditions. pp. 35-40.

[18] C. Vertan, M. Ciuc, and N. Boujemaa. On the introduction of a chrominance spectrum and its applications. In *Proceedings of the International Conference on Color in Graphics and Image Processing*, Saint-Etienne, France, 2000. Cépaduès Éditions. pp. 214-218.