

HUMAN SKIN COLOR DETECTION AND APPLICATION TO ADULT IMAGE DETECTION

Ryszard S. Choraś

*Institute of Telecommunications, University of Technology & Life Sciences
Kaliskiego Street 7, 85-796 Bydgoszcz, Poland*

Keywords: Skin color, YC_bC_r color space, skin likelihood image, adult images.

Abstract: In this paper, we aimed at the detection of adult image. The methods of detection mainly focus on the detection/identification of skin region. Skin detection is of the paramount importance in the detection of adult images. Our algorithm is designed to detect human skin color in YC_bC_r color space. The proposed system finds skin regions and then generates the skin likelihood image. Since the skin likelihood image contains shape information as well as skin color information, we used the skin likelihood image to classify the adult images.

1 INTRODUCTION

The rapid development in the field of digital media has exposed us to huge amounts of non-textual information such as audio, video, and images. The retrieval and classification techniques have been and are continuously being developed and improved to facilitate the exchange of relevant information.

But images can be able to contributing nude content. Effective filtering and blocking these images is of paramount importance in an wide computer networks (e.g. Internet). An image analysis technique is needed in order to classify adult (nude, objectionable) images and further block accessing to objectionable sites.

The detection of images containing nudity and pornography is based on the identification of human skin and usually based on color with texture as a addition feature (Fleck, M.M., Forsyth, D.A., Bregler, C., 1996),(Jones, M.J., Rehg, J.M., 1998),(Chan, Y., Harvey, R., Bagham, J., 2000).

There is an extensive literature (J.Z. Wang, J. Li, G. Wiederhold, O. Firschein, 1998), (A. Bosson, G. C. Cawley, Y. Chan and R. Harvey, 2002) on the detection of naked images by using features such as color histograms, texture measures and shape measures. Bosson et al. finds skin blobs, and then computes the area, centroid, length of the axes of an ellipse, eccentricity, solidity, and extent of skin blobs. The skin color model used by Fleck et al. (Fleck, M.M., Forsyth, D.A., Bregler, C., 1996) consists of a manually specified region in a log-opponent color space. Detected regions of skin pixels form the input

to a geometric filter based on skeletal structure.

Skin-color detection has been employed in many applications such as face detection, gesture recognition, human tracking, pornographic image filtering, etc.

Skin color has proven to be a useful and robust cue for face detection, localization and tracking. Numerous techniques for skin color modeling and recognition have been proposed during several past years. Most existing skin segmentation techniques involve the classification of individual image pixels into skin and non- skin categories on the basis of pixel color.

The main goal of skin color detection is to build a decision rule that will discriminate between skin and non-skin pixels. Identifying skin colored pixels involves finding the range of values for which most skin pixels would fall in a given color space. Various color spaces are used for processing digital images. For some purposes, one color space may be more appropriate than others.

Jiao et al. (F. Jiao, W. Gao, L. Duan, G. Cui, 2001) presented an adult image detection method. They first use the skin color model to detect naked skin areas roughly. Then the Gabor filter are applied to remove those non-skin pixels. Some simple features are extracted to detect adult images. Several representative features induced from the naked images are used to verify these skin areas.

2 COLOR

Each image is represented using three primaries of the color space chosen. Most digital images are stored in RGB color space. RGB color space is represented with red (R), green (G), and blue (B) primaries and is an additive system. RGB color space is not perceptually uniform, which implies that two colors with larger distance can be perceptually more similar than another two colors with smaller distance, or simply put, the color distance in RGB space does not represent perceptual color distance.

RGB is one of the most widely used colorspace, however, high correlation between channels, significant perceptual non-uniformity, mixing of chrominance and luminance data make RGB not a very favorable choice for color analysis and colorbased recognition algorithms.

Normalized RGB is a representation, that is easily obtained from the RGB values by a simple normalization procedure: $r = \frac{R}{R+G+B}$; $g = \frac{G}{R+G+B}$; $b = \frac{B}{R+G+B}$. The three normalized components r , g and b are called pure colors; they contain no information about the luminance. Because ($r+g+b = 1$), the third component does not hold any significant information and can be omitted, reducing the space dimensionality. It is enough to use only two components r and g to completely describe the skin color space.

YC_rC_b is an encoded nonlinear RGB signal for image compression work. Color is represented by luminance, computed from nonlinear RGB (Poynton, 1995), constructed as a weighted sum of the RGB values, and two color difference values C_r and C_b that are formed by subtracting luminance from RGB red and blue components.

YC_bC_r color space has been defined in response to increasing demands for digital algorithms in handling video information, and has since become a widely used model in a digital video. It belongs to the family of television transmission color spaces. These color spaces separate RGB into luminance and chrominance information.

$$\begin{aligned} Y &= 0,299R + 0,587G + 0,114B \\ C_r &= 0,713(R - Y) \\ C_b &= 0,564(B - Y) \end{aligned} \quad (1)$$

HSI, HSV, HSL - Hue, Saturation, Intensity (Value, Lightness), color spaces describe color with intuitive values. Hue defines the dominant color (such as red, green, purple and yellow) of an area, saturation measures the colorfulness of an area in proportion to its brightness (Poynton, 1995). The "intensity", "lightness" or "value" is related to the color lu-

minance. The intuitiveness of the color space components and explicit discrimination between luminance and chrominance properties made these color spaces popular in the works on skin color segmentation (Zarit, B.D., Super, B.J., Quek, F.K.H., 2002), (Sigal, L., Sclaroff, S., Athitsos, V., 2000).

$$\begin{aligned} H &= \arcsin\left(\frac{C_r - 128}{128 \cdot S}\right); V = \frac{Y}{256} \\ S &= \frac{\sqrt{(C_r - 128)^2 + (C_b - 128)^2}}{128} \end{aligned} \quad (2)$$

3 SKIN-COLOR MODEL

3.1 Explicitly Defined Skin Region

The statistical skin-color model is generated by means of a supervised training, using a set of skin-color regions, obtained from a color human body database. Such images were obtained from people of different races, ages and gender, with varying illumination conditions.

One method to build a skin classifier is to define explicitly (through a number of rules) the boundaries skin cluster in some color space. For example (Peer, P., Kovac, J., Solina, F., 2003) (R, G, B) is classified as skin if:

$$\begin{aligned} R &> 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and} \\ \max\{R, G, B\} - \min\{R, G, B\} &> 15 \text{ and} \\ |R - G| &> 15 \text{ and } R > G \text{ and } R > B. \end{aligned}$$

We have found that skin-color region can be identified by the presence of a certain set of chrominance (ie C_r and C_b) values that is narrowly and consistently distributed in the YC_rC_b color space. We denote RC_r and RC_b as the respective ranges of C_r and C_b values that correspond to skin color, which subsequently define our skin-color reference map. The ranges that we found to be the most suitable for all the input images that we have tested are $RC_r = [133, 173]$ and $RC_b = [77, 127]$. This map has been proven, in our experiments, to be very robust against different types of skin color (Figure 1).

With this skin-color reference map, the color segmentation can now begin. Since we are utilizing only the color information, the segmentation requires only the chrominance component of the input image. The output of the color segmentation, is a skin bitmap SM described as

$$SM(x, y) = \begin{cases} 1 & \text{if } C_r(x, y) \in RC_r \cap C_b(x, y) \in RC_b \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

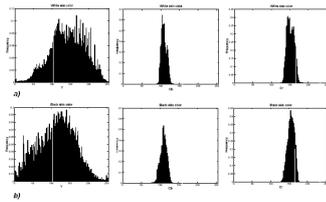


Figure 1: The histogram of Y, C_b, C_r components of white (a) and black (b) skin colors.

The output pixel at point (x, y) is classified as skin-color and set to 1 if both the C_r and C_b values at that point fall inside their respective ranges, RC_r and RC_b . Otherwise, the pixel is classified as non-skin-color and set to 0 (Figure 2).

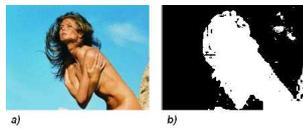


Figure 2: Original image (a) and skin map (b).

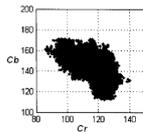


Figure 3: Skin color distribution.

The skin color model based on C_b and C_r values can provide good coverage of all human races. Despite their different appearances, these color types belong to the same cluster in $C_b C_r$ plane (Figure 3). We classify the popular skin colors of naked images, including various lighting conditions, into several categories. The corresponding compactly chroma histogram of each category is compiled in advance.

3.2 Feature Extraction

The performance of the naked image detection extremely depends on the accurate skin segmentation. Since the skin property is very smooth, we utilize the roughness feature to further reject confusion from non skin stuff with skin like chroma.

3.2.1 Texture Feature from Gabor Wavelet Transform

For further improved skin detection blobs we propose skin color filter based on texture features. The Gabor filter will be eliminated non-skin pixels.

Gabor wavelet based texture is robust to orientation and illumination change, It is a powerful tool to

extract texture features. Gabor functions are Gaussians modulated by complex sinusoids. In two dimensions they take the form:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right] \quad (4)$$

$$G(u, v) = \exp\left\{-\frac{1}{2}\left[\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right]\right\} \quad (5)$$

$$\text{where } \sigma_u = \frac{1}{2\pi\sigma_x}, \sigma_v = \frac{1}{2\pi\sigma_y}.$$

Gabor wavelets can be obtained by appropriate dilations and rotations of $g(x, y)$ through the generating function:

$$g_{mn}(x, y) = a^{-m}g(x', y'); \quad a > 1 \quad m = 0, 1, \dots, S - 1 \quad (6)$$

$$x' = a^{-m}(x \cos \theta + y \sin \theta), \quad y' = a^{-m}(-x \sin \theta + y \cos \theta) \quad (7)$$

$$\theta = \frac{n\pi}{K} \quad k = \text{integer}.$$

Where k is the number of orientations, given an image $I(x, y)$, its Gabor wavelets transform is calculated as:

$$W_{mn}(x, y) = \sum \sum I(x_1, y_1)g_{mn}^*(x - x_1, y - y_1) \quad (8)$$

Where g^* is the complex conjugate of g , the texture feature is given as:

$$T(x, y) = \sqrt{\left(\sum \sum W_{mn}^2(x, y)\right)} \quad (9)$$

In our experiment, we use three scales ($m = 0, 1, 2$) and four orientations ($n = 0, 1, 2, 3$) (Figure 4).

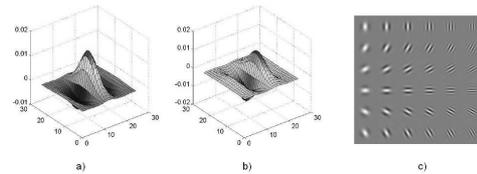


Figure 4: Real (a) and imaginary (b) parts of Gabor wavelets and Gabor kernels with different orientations (c).

Then texture image is computed from equation 9, the texture mask image is obtained as:

$$TM(xy) = \begin{cases} 1 & \text{if } T(x, y) \leq th \\ 0 & \text{if } T(x, y) > th \end{cases} \quad (10)$$

where th is the texture threshold. Since human skin is smooth, its texture feature value is relatively low. Using texture mask, pixels reassemble skin in color but has high texture feature value will be filtered.

For each obtained skin color region (blob) we define a "rectangular box" with the aspect ratio ar defined as

$$ar = \frac{|x_{right} - x_{left}|}{|y_{top} - y_{down}|} \quad (11)$$

where $x_{right}, x_{left}, y_{top}, y_{down}$ indicate the smallest and largest x- and y-coordinates, respectively.

The rectangular box contain a skin color region as well as several non-skin color region. The skin color ratio scr is defined as

$$scr = \frac{\text{The number of skin color pixels}}{\text{Area rectangular box}} \quad (12)$$

Next, we apply the ellipse model for skin body blobs (Figure 5). The central position (x_c, y_c) of the blob are estimated as $x_c = \frac{M_{10}}{M_{00}}$; $y_c = \frac{M_{01}}{M_{00}}$ where M_{00}, M_{10}, M_{01} are the first order moment calculated from the blob pixels.

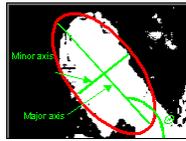


Figure 5: Blob ellipse model.

The ellipse model for a candidate skin body region is defined by area, major and minor axis of the ellipse with following equations

$$axis_{major} = \sqrt{6(p+r + \sqrt{q^2 + (p-r)^2})} \quad (13)$$

$$axis_{minor} = \sqrt{6(p+r - \sqrt{q^2 + (p-r)^2})} \quad (14)$$

where $p = \frac{M_{20}}{M_{00}} - x_c^2$, $q = 2(\frac{M_{20}}{M_{00}} - x_c y_c)$, $r = \frac{M_{02}}{M_{00}} - y_c^2$.

The parameters of the ellipse are used to decide the body skin area.

The nudity detection algorithm works in the following manner:

1. Calculate the corresponding $YCbCr$ values from the RGB values. Label each pixel as skin or non-skin and identify connected skin pixels to form skin regions.
2. Use the Gabor filter to improve skin regions and eliminate non-skin pixels.

3. Define a "rectangular box" for the largest skin blobs and calculate ar , scr . Fit the skin blobs using a simple geometric shape such as ellipse. Calculate the parameters of the ellipse.
4. If the percentage of skin pixels relative to the image size is less than 15 percent, the image is not nude. If the ranges of ar and scr are $ar = [0.35, 0.85]$ and $scr = [0.75, 0.85]$ the image is nude.

4 EXPERIMENTAL RESULTS AND CONCLUSIONS

In this paper, we use the color information of image to detected the regions skin in nude image. Because the shape and texture information of image is useful information, in classifying nude image is necessary.

The detection rate and false alarm rate are expressed as a percentage, which represent in equation 15

$$DR = \frac{TP}{TP + FN} ; FAR = \frac{FP}{TP + FP} \quad (15)$$

where TP = True Positive, FP = False Positive and FN = False Negative.

The detection rates and false alarm rates for the testing set are respectively $DR = 93\%$ and $FAR = 7\%$. These results show that the proposed segmentation algorithm could provide the face segmentation more efficient and it also has less affect from they fast or slow moving object in video sequences.

REFERENCES

- A. Bosson, G. C. Cawley, Y. Chan and R. Harvey (2002). Non-retrieval: blocking pornographic images. In *Proceedings of the International Conference on the Challenge of Image and Video Retrieval*, volume 2383 of *Lecture Notes in Computer Science*, pages 50–60. Springer, London.
- Chan, Y., Harvey, R., Bagham, J. (2000). Using colour features to block dubious images. In *Proc. Eusipco*.
- F. Jiao, W. Gao, L. Duan, G. Cui (2001). Detecting adult image using multiple features. In *Proceedings of IEEE International Conference on Info-tech and Info-net*, pages 378–383.
- Fleck, M.M., Forsyth, D.A., Bregler, C. (1996). *European Conference on Computer Vision*, volume II, chapter Finding naked people, pages 593–602. Springer-Verlag.

- Jones, M.J., Rehg, J.M. (1998). Statistical color models with application to skin detection. Technical Report CRL 98/11 11, Compaq Cambridge Research Laboratory.
- J.Z. Wang, J. Li, G. Wiederhold, O. Firschein (1998). System for screening objectionable images. *Computer Communications*, 21(15):1355–1360.
- Peer, P., Kovac, J., Solina, F. (2003). Human skin colour clustering for face detection. In *EUROCON 2003 - International Conference on Computer as a Tool*.
- Poynton, C. A. (1995). Frequently asked questions about colour. In <ftp://www.inforamp.net/pub/users/poynton/doc/colour/>.
- Sigal, L., Sclaroff, S., Athitsos, V. (2000). Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, volume 2, pages 152–159.
- Zarit, B.D., Super, B.J., Quek, F.K.H. (2002). Comparison of five color models in skin pixel classification. In *ICCV99 Intl Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems*, pages 58–63.



SciTeP Press
Science and Technology Publications