Adult image content filtering: A statistical method based on Multi-Color Skin Modeling

Conference Paper · November 2010

DOI: 10.1109/ISSPIT.2010.5711812 · Source: DBLP

2 authors:

Mahmoud Mofaddel
Sohag University

Samy Bakheet
Sohag University

Adult Image Content Filtering: A Statistical Method Based on Multi-Color Skin Modeling

Mahmoud A. Mofaddel and Samy Sadek
Department of Mathematics and Computer Science
Faculty of Science, Sohag University, 82524 Sohag, Egypt
mmofaddel@hotmail.com, samy.bakheet@ovgu.de

***Abstract*—Automatic skin detection is a key enabler of various imaging applications, such as face detection, human tracking, and adult content filtering. In 1996, the first paper on identifying nude pictures was published. Since then, different researchers argue different color models to be the best choice for skin detection. But, to the best our knowledge, no significant work has been reported previously that attempted to use more than one color model and evaluate the performance for recognizing adult contents. In this paper, a simple statistical framework for recognizing adult images based on an MCSM (Multi-Color Skin Model) is described. From a high level, our approach works in two steps. First, skin regions in an input image are detected using the MCSM. Then these suspected regions are fed into a specialized geometrical analyzer that attempts to assemble a human figure using simple geometric shapes derived from human body structure. Quantitative evaluation shows that our method compares favorably with the state-of-the-art methods in terms of detection rate and false alarm, while reducing the computational complexity by a factor of 1/6 with respect to the Forsyth's method.**

***Index Terms*—Skin detection, adult content, object recognition, content-based retrieval.**

I. Introduction

Over the course of the past two decades protecting children from harmful materials on the Internet such as adult contents has received a great deal of attention and is still a matter of great concern for many researchers in the field of computer vision and pattern recognition. The pioneering work on identifying nude pictures is that of Forsyth *et al.* [1]. In this approach, the images containing large areas of skin are first identified. Then by using a specially defined human structure, the images containing human figures are recognized as nude images. In their approach, only the images that are at least covered in one third by the skin area are fed to a geometric analyzer. While the images used in [1] are 128×192, it takes about six minutes for the method to process a single image. Though, the final recall of the method is less than 50%. After that, several new contributions have been made in this field [2], [3], [4], [5], [6], [7]. Generally speaking, skin detection plays an important role in various applications related to computer vision and pattern recognition such as searching and filtering image content on the web, face detection and tracking. Much work has been done in this regard. References [8], [9], [10], [11], [12] discuss detection of human skin and the effect of different cameras, light-setting, human race and color spaces on the recognition process. Furthermore, references [1], [13] use texture information as a component in the skin detection. As a third component in object recognition, many researchers have looked at the shape of the object as a last stage in the information gathering process [14], [15], [16]. Arentz and Olstad [17] have proposed a method for helping to identify adult web sites by using image-contents as a mean to detect erotic material. The average detection error rates for offensive sites were 14.1%, as compared to 9.8% for non-offensive as indicated in their paper. In [18] Veltkamp and Hagedoorn have written a survey of different state-of-the-art shape matching methods. With a set of features describing the image, such as color and shape for segmented objects, it is possible to build fully system for adult image detection.

The remainder of this paper is structured as follows. In Section II, the proposed method is described in detail. Experimental setup and results are given in Section III. Finally, conclusions are drawn in the last section with some discussion on future work.

II. Suggested Methodology

Indeed, it is intuitively plausible that it would be very difficult to find a precise definition of "skin" based on attributes of color alone, which can correctly identify skin pixels and non-skin pixels in all cases since there are too many outside factors such as lighting conditions that might change the apparent color of skin, and of course, different people have different colore skin. In addition, objects in the background may have the same color as a person's skin and there is no clear way of telling the difference using these methods.

In order to deal with or tackle such problems, it seems reasonable to take an approach that uses a set of color model to get a new skin detection algorithm that gives higher accuracy. It was stated that using a set of color spaces is a good idea to precisely extract a more defined skin region. In the implementation of the proposed MCSM, there are some main steps viz. (1)apply skin filter to identify the skin regions and apply threshold if necessarily, (2) apply median filter to get rid of impulsive salt-pepper noise and (3) The skin-colored regions of areas smaller than predefined threshold, are deleted. These regions are too small to be counted as human skin. An overview of the MCSM is shown in Fig. 1.

978-1-4244-9991-5/11/$26.00 ©2011 IEEE

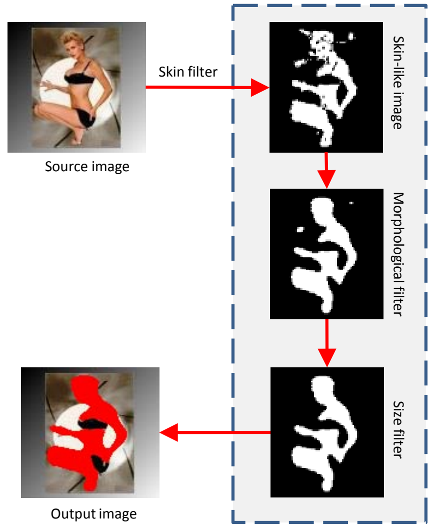

Fig. 1. An overview of multi-color skin modeling.

A. Multi-color Skin Modeling

As stated before, a robust skin detector is the primary need of many fields of computer vision, including face detection, gesture recognition, and adult content filtering. In the following subsections, a statistical multi-color model for skin detection that makes use of RGB, $Normalized\ RGB$, $YCrCb$ and HSI color spaces is described.

1) Skin Filter in RGB Space: Indeed one of the simplest methods for detecting skin pixels is to use an explicitly defined skin region. The simplicity of these methods has attracted (and still does) many researchers [19], [20]. Normally, in these methods, the pixels for skin region can be detected by defining explicitly (through a number of rules) the boundaries skin cluster in RGB color space which provides a fast means of skin detection. For example:

$$\begin{aligned} &\text{if } r > r_m, g > g_m, b > b_m, r-b > \alpha, r-g > \beta \\ &\text{then } skin(r,g,b) = 1 \\ &\text{else } skin(r,g,b) = 0 \end{aligned} \tag{1}$$

where r_m, g_m, b_m, α and β are constants that are estimated from the training data. The obvious advantage of this algorithm is simplicity of skin detection rules that leads to construction of a very rapid classifier. The main difficulty to achieve high recognition rates with this approach is the need to find precise decision rules empirically.

2) Skin Filter Using RGB Channels Ratio: It was observed, that pixels belonging to skin region regularly contains a significant level of red. Using this observation, certain values of the two ratios: $g = G/R$ and $b = B/R$ can be used as skin presence indicator [21]. Let the thresholds be chosen as (g_1, g_2), (b_1, b_2) then a pixel is classified to have skin tone if $g \in (g_1, g_2)$ and $b \in (b_1, b_2)$.

3) Skin Filter in Normalized RGB Space: The skin detection algorithm used here is based on normalized RGB, or chromaticity space. The chromaticities are defined as

$$\begin{aligned} r &= \frac{R}{R+G+B} \\ g &= \frac{G}{R+G+B} \end{aligned} \tag{2}$$

The reason for using this color space is due to evidences that the human skin color is more compactly represented in it than it is in other color spaces, such as RGB, HSI, SCT and YQQ [22]. Here we use the chromaticities r and g to describe the color. Skin color distribution can be modeled by an elliptical Gaussian joint probability density function (pdf), defined as:

$$p_{skin}(x) = \frac{1}{2\pi|\Sigma_s|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_s)^T \Sigma_s^{-1} (x-\mu_s)} \tag{3}$$

Here, x is a color vector and μ_s and Σ_s are the distribution parameters (mean vector and covariance matrix respectively). The model parameters are estimated from the training data using the following formulas,

$$\mu_s = \frac{1}{n}\sum_{k=1}^{n} x_k \tag{4}$$

$$\Sigma_s = \frac{1}{n}\sum_{k=1}^{n} (x_k - \mu_s)^T (x_k - \mu_s) \tag{5}$$

where n is the total number of samples and x_k is the vector representing the k sample. The probability can be used directly as the measure of how "skin-like" the color is [23]. Gaussian skin modeling has been also employed in [24], [22], [11].

4) Skin Filter in $YCrCb$ Space: In this color space, the two chroma components Cr, and Cb can be efficiency used to define explicitly skin region. The thresholds be chosen as (Cr_{max}, Cr_{min}) and (Cb_{max}, Cb_{min}) , a pixel is classified as skin pixel if the values (Cr, Cb) fall within the thresholds.

5) Skin Filter in HSI Space: Skin filter in HSI space The value of Hue and saturation (H, S) are adequate to segment the skin region from non-skin region. To detect skin region, the values of H, and S should be explicitly determined through defining their intervals as shown below:

$$\begin{aligned} &\text{if } H \in (H_{min}, H_{max}) \wedge S \in (S_{min}, S_{max}) \\ &\text{then } skin(r,g,b) = 1 \\ &\text{else } skin(r,g,b) = 0 \end{aligned} \tag{6}$$

The values of $H_{min}, H_{max}, S_{min}, S_{max}$ are experimentally determined using the training data.

B. Adult Image Content Detection

One interesting application of skin detection is as part of a larger system for detecting adult content in photos. An adult content detector that worked reliably to detect adult images could be a valuable tool for image search services in digital libraries [25], as well as for image categorization. The main goal of the proposed adult content detection system is to determine whether or not an input image contains an adult content by feeding the output of the skin detector to

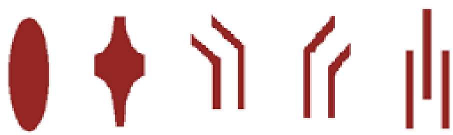

Fig. 2. Main primitive shapes used to represent human body parts.

a geometric analyzer, which attempts to find human figure using geometric constraints on human figure structure. The geometric analyzer based on the fact that the human figure can be viewed as an assembly of nearly cylindrical parts, where the individual geometry of these parts is constrained by the geometry of the skeleton. So the proposed system models a human as a set of primitives that form the human parts. The primitives used by the geometric analyzer are shown in Fig. 2. The steps involved in our system can be summarized as follows,

1) *Step1 (Skin Detection):* The first main step in building a system for adult images detection is the skin-colored regions recognition. This stage can be accomplished by the proposed MCSM described previously at the beginning of this section.
2) *Step2 (Geometrical Analysis):* In this stage, the skin regions are fed to a specialized geometric analyzer, which attempts to find a human figure using simple geometric shapes derived from human body structure. Then, a decision is made to determine whether or not the image region contains an adult content. Finally, images containing sufficiently adult contents are masked or blocked. Fig. 3 shows a simple flowchart of these two steps of the adult content detection.

By taking advantage of the fact that adult images are often sized to frame a standing or reclining figure, some features can be computed from the output image of the skin detector. These features play an important role in increasing overall recognition rate of the proposed system, which include: (1) edge of connected components of skin, (2) height and width of the largest region of skin, (3) percentage of pixels detected as skin, and (4) number of connected components of skin These features can all be computed in a single pass, before feeding the image into the geometric analyzer that allows the adult content detection system to be extremely fast compared with the Forysth's system counterpart [20].

III. Experimental results

Due to the lack of any standard image databases for testing and comparison of adult content detection systems, the proposed system has been tested using our own images database. This database has 562 test adult images and 1580 assorted control control images, containing some images of people but none of adult images. All images were taken of type RGB format 8 bits / pixels in each color channel. The

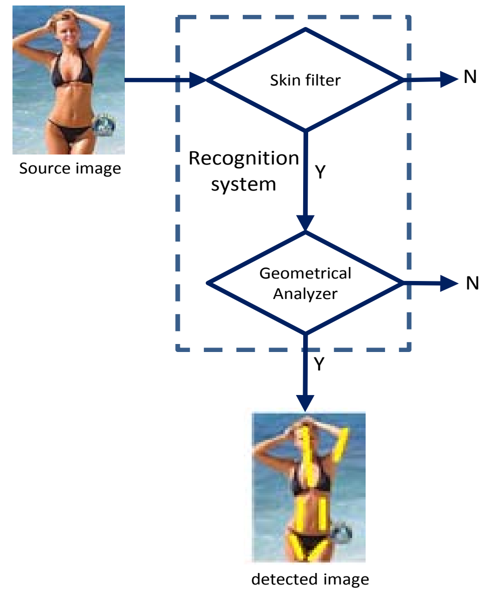

Fig. 3. A simple flowchart diagram showing the main steps of the approach.

test images were collected from the internet. They show a very wide range of postures. Some depict several people including naked or scantily-dressed. Some depict only small parts of the bodies of one or more people. Most people in the images are Caucasian; a small number are Blacks or Asians. To evaluate the results of the proposed adult content detector two different metrics are used. TP (true positive) is the number of adult images identified correctly divided by the number of all test images. FP (false positive) is the number of non-adult images identified as adult images divided by the number of all control images. Receiver operating characteristics (ROC) curve that shows the relationship between correct detections and false detections of the proposed adult content detector as a function of the detection threshold is shown in Fig. 4.

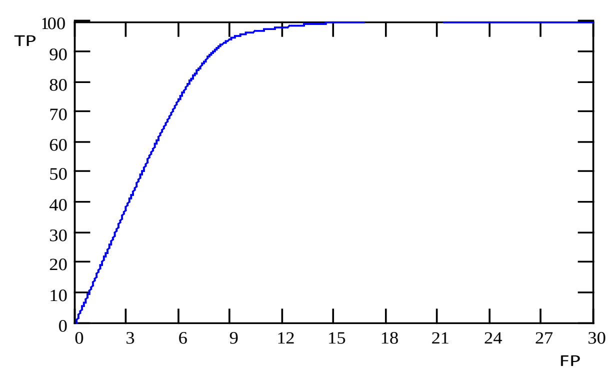

Fig. 4. ROC curve of the proposed adult content detector.

Fig. 5 shows the capability of the proposed system to detect a sample of clothed images in the test set. These images correctly classified as containing adult content. Mistakes by the proposed pornography detector may occur for several reasons.

Fig. 5. Performance obtained by the proposed adult content detector: (i) source images; (ii) result of skin filter; (iii) result of geometric analyzer.

TABLE I
COMPARISON WITH OTHER STATE-OF-THE-ART SYSTEMS

System	TP	FP
Proposed system	**89.3%**	**7.4%**
Arentz-Olstad [17]	85.9%	9.8%
Jones-Rehg [26]	86.7%	9%
Forsyth-Fleck [1]	43%	4.2%

In some images, adult contents are too small to be detected. In others, most or all of the skin area is desaturated, so it fails the skin detector. Some control images pass the skin detector because they contain people, particularly several close-up portrait shots. Other control images contain material whose color closely resembles that of human skin; particularly wood, sand, and skin or fur of certain animals. The proposed recognition system can be variously configured. In one configuration of the system, it can successfully identify 89.3% of the test images (true detection), but only 7.4% of the control images (false alarm). Table I provides a quantitative comparison between our system and those of other investigators in the literature. As can be clearly seen from the table, our method yields encouraging results and is relatively superior compared to the other methods. Such results have the potential to compare favorably with those reported in the existing literature in terms of detection rate and false alarm rate. Furthermore, our system is relatively fast; first-stage,i.e. skin detection, takes a trivial amount of time while the geometric analyzer processes pictures at a rate of less than a minute per picture. On the other hand the method developed by Forsyth and Fleck [20] takes about 6 minutes on a workstation for the figure grouper to process a suspect image passed by the skin filter.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a statistical method to detect and identify adult contents in color images using a combination of a simple visual color cue and geometric human figure characteristics. The proposed system can be flexibly configured to meet the users needs. The results obtained showed that our method performs comparably well as or better than the existing methods in terms of detection rate and false alarm rate, while it is simple and relatively rapid. The future work will be along two major axes. The first will be the further improvement of the method by using more elaborated feature sets; for example, by adding texture features, while the second will be to examine the robustness of the method by performing more experiments on real-word video data.

ACKNOWLEDGMENT

The doctoral scholarship awarded by University of Sohag (Egypt) to the second author is gratefully acknowledged.

REFERENCES

[1] David A. Forsyth and Margaret M. Fleck ”Identifying nude pictures”, IEEE Workshop on the Applications of Computer Vision, pp.103–108, 1996.

[2] S. Sadek, A. Al-Hamadi, B. Michaelis, and U. Sayed, ”A Robust Neural System for Objectionable Image Recognition,” Second Int. Conference on Machine Vision (ICMV2009), Dubai, UAE , pp. 32–36, 2009.

[3] H. Zhu, S. Zhou, J. Wang and Z. Yin, ”An algorithm of pornographic image detection,” Fourth Int. Conference on Image and Graphics (ICIG 2007), pp. 801–804, , 2007.

[4] L.-L. Cao, X.-L. Li, H.-H. Yu and Z.-K. Liu,”Naked people retrieval based on adaboost learning,” First International Conference on Machine Learning and Cybernetics, Beijing, China, pp. 1133–1138, 2002.

[5] R. Schettini, C. Brambilla, C. Cusano and G. Ciocca, ”On the detection of pornographic digital images,” Visual Communications and Image Processing, vol. 5150, pp. 2105–2113, 2003.

[6] Y.-C. Lin, H.-W. Tseng and C.-S. Fuh, ”Pornography detection using support vector machine,” 16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP 2003), Kinmen, ROC, pp. 123–130, 2003.

[7] T. Deselaers, L. Pimenidis and H. Ney, ”Bag-of-Visual-Words Models for Adult Image Classification and Filtering,” ICPR, 2009.

[8] E. Angelopoulou ”Understanding the color of human skin”, In Proc. SPIE Conf. On Human Vision and Electronic Imaging VI (SPIE) 2001, SPIE vol. 4299, SPIE Press, pp.243–251, 2001.

[9] L. Bretzner, I. Laptev and T. Lindeberg ”Hand gesture recognition using multi-scale color features, hierarchical models and particle filtering”, 2002. http://www.pernilleravn.dk/381/downloads/BreLapLin-FG02.pdf

[10] R.-L. Hsu, M. Abdel-Mottaleb and A. Jain ”Face detection in color images”, IEEE Trans. Patt. Anal. Mach. Intell., pp. 696–706, 2002.

[11] H. Kruppa, M.A. Bauer, and B. Schiele ”Skin Patch Detection in Real-World Images”, DAGM Symp., Zurich, Switzerland, pp. 109–116, 2002.

[12] M. C. Shin, K. I. Chang and L. V. Tsap ”Does color space transformation make any difference on skin detection?,” IEEE Workshop on Applications of Computer Vision, Orlando, FL, pp. 275–279, (2002).

[13] J.M. Rehg, K.P. Murphy and P.W. Fieguth ”Vision-based speaker detection using Bayesian networks,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Ft. Collins, Colorado, pp. 110–116, 1999.

[14] Qing-Fang Zheng, Wei Zeng, Wei-Qiang Wang, Wen Gao ”Shape-based Adult Image Detection”, Int. J. Image Graphics 6(1):115–124, 2006.

[15] J.S. Park and D.H. Chang, ”2-d invariant descriptors for shape-based image retrieval”, Pattern Recognition Letters, 2001.

[16] M. Safar, C. Shahabi and X. Sun, ”Image retrieval by shape: a comparative study”, IEEE International Conf. on Multimedia and Expo (I), pp. 141–144, 2000.

[17] W. A. Arentz and B. Olstad ”Classifying offensive sites based on image content”, Computer Vision and Image Understanding, pp. 295–310, 2004.

[18] R. Veltkamp. and M. Hagedoorn, ”State-of-the-art in shape matching”, Technical Report, UU-CS-1999-27, Utrecht University, the Netherlands, 1999.

[19] P. Peer, J. Kovac and F. Solina ”Human skin color clustering for face detection”, EUROCON 2003 - International Conference on Computer as a Tool, 2003.

[20] D. A. Forsyth and M. Fleck, ”Automatic detection of human nudes,” International Journal of Computer Vision, Vol. 32(1), pp. 63–77, 1999.

[21] J. Brand and J. Mason, ”A comparative assessment of three approaches to pixel level human skin-detection”, in: Proc. of the International Conference on Pattern Recognition, vol. 1, pp. 1056–1059, 2000.

[22] E. Littmann and H. Ritter ”Adaptive color segmentation. A comparison of neural and statistical methods”, IEEE Trans. Neural Networks, pp. 175–185, 1997.

[23] B. Menser and M. Wien ”Segmentation and tracking of facial regions in color image sequences”, in: Proc. SPIE Visual Communications and Image Processing, pp. 731–740, 2000.

[24] U. Sayed, S. Sadek and B. Michaelis ”Two Phases Neural Network-Based System for Pornographic Image Classification”, In Proceedings of 5th International Conference of Sciences of Electronic, Technologies of Information and Telecommunications (SETIT2009), pp. 1–6, 2009.

[25] J. Z. Wang, J. Li, G. Wiederhold and O. Firschein. ”Classifying objectionable websites based on image content”, Notes in Computer Science, Special issue on interactive distributed multimedia systems and telecommunication services, 21/15:113–124, 1998.

[26] M. Jones and J. Rehg, ”Statistical Color Models with Application to Skin Detection,” [online] Available: http://www.hpl.hp.com/techreports/Compaq-DEC/ CRL-98-11.pdf. (30 September 2004)