

Skin Detection for Adult Image Identification

Jung-Jae Yu , Seung-Wan Han

Video Surveillance Research Section

Electronics and Telecommunications Research Institute, Korea

jungjae@etri.re.kr, hansw@etri.re.kr

Abstract— This work is aimed at the estimation of the skin region in pornographic images to identify and block the pornographic contents. We propose a novel method to estimate the skin region using simple operation in HSV color space and additional post-processing to reduce the noise effects. Our method does not require any pre-process such as object detection. It is fast and its result is good enough to be used for filtering out the clear non-objectionable images before main identification process. So the proposed method can be used as pre-processing for more accurate poronographic image identification process.

Keywords— skin, color, detection, estimation, image, adult, pornography, naked

I. INTRODUCTION

This work is aimed at the estimation of skin regions in pornographic images to identify and block the pornographic contents. Skin information is of paramount importance in the pornographic image identification. If we could measure the ratio of the skin region from the input image, we can use this information to exclude distinctive non-objectionable images before the main identification block. For example, if the ratio of the skin region from a image is very low, we can conclude that the possibility of the image to be objectionable is very low. So the ration of skin region can be used in the pre-filtering system for objectionable candidate images. This pre-filtering system system will conserve computation time to deal with massive input images and enable the indentification system to concentrate on the objectionable candidate images. Fig. 1 shows this notation.

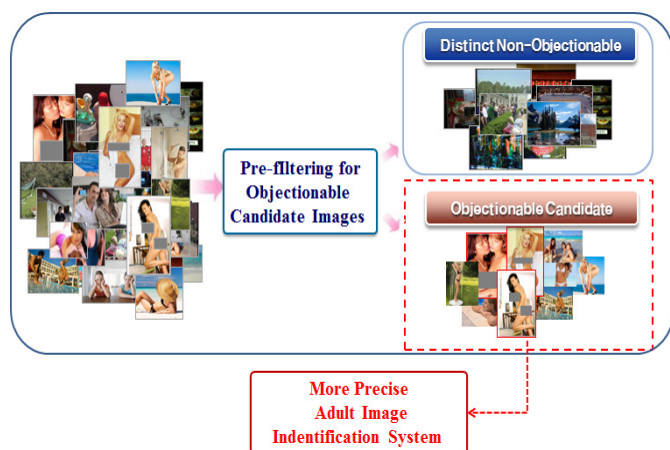


Figure 1. The pre-filtering system for objectionable candidate images

But it is difficult to extract accurate skin regions from various input images. The changes of illumination condition and various races make it difficult to apply pre-determined color models to various input images. Object detection methods such as face detection or eye detection have been employed to estimate the color model adjusted to the current input image. But these approaches increase the computation time and cannot be used for input images which do not include any faces. So we propose a novel method to estimate the skin region using simple operation in HSV color space and additional post-processing to reduce the noise effects. Our method does not require any pre-process such as object detection and is so fast that it is suitable as a pre-processing of more accurate porographic image identification systems.

II. PREVIOUS WORKS

Jones[1] proposed a statistical method to extract skin pixels from the input image. He used MLE approach to differentiate the skin pixels with non-skin pixels. For his approach, many training images are required for estimating the likelihood function. Lee[2] used a face-detection process before extracting skin pixels to estimate the color distribution of skin pixels from the input image. The color information of faces are useful for estimating the color distribution of skin pixels but additional computation time is consumed for the face detection process. Shih[3] used a simple operation in YCbCr color space to extract skin pixels. His method is very fast but the accuracy of the result is not sufficient for the adult image identification. Ghomsheh[4] estimated the color density function of exponential shape in HSV color space. His method requires a curve fitting process and is not robust in images including various color components in the back ground region.

III. PROPOSED METHOD

Proposed method is composed of three steps. The first step is converting the RGB raw image to a HSV color image and extract skin candidate pixels by simple threshold operation. We used a simple threshold operation to extract skip candidate pixels in Hue components. Estimated skin candidate pixels are expressed as an alpha image in which the intensity is 255 for skin pixels and the intensity is 0 for non-skin pixels.

The second and third step are post-processing steps to remove incorrectly detected pixels which have color components similar to skin color but should be belonged to a non-skin region. In the second step, an edge density map for

the input image is computed and non-skin pixels are removed from the skin-candidate pixels. To compute an edge density map, the input is converted to a gray image and simple edge operation such as sobel edge is used to get the edge image. And a sum image of the edge image is computed and finally edge density image is computed according to the equation (1). In equation (1), $E(i,j)$ is the edge image and $E_{sum}(u,v)$ is the sum density image and $E_{dns}(u,v)$ is the final edge density image.

$$E_{sum}(u,v) = \sum_{i=0}^{u-1} \sum_{j=0}^{v-1} E(i,j)$$

$$E_{dns}(u,v) = \frac{1}{N} \sum_{i=u-w}^{u+w} \sum_{j=v-w}^{v+w} E(i,j)$$

$$= \frac{1}{N} (E_{sum}(u+w, v+w) + E_{sum}(u-w, v-w) - E_{sum}(u+w, v-w) - E_{sum}(u-w, v+w)), \quad (1)$$

when $N = (2w+1)^2$

We use a premise that the edge density is low in skin regions and the pixels with very high edge density are likely to belong to a background region. So we remove the pixels with very high edge density from the skin candidate pixels extracted at the previous step. Fig. 2 shows the background removing process in step 2.

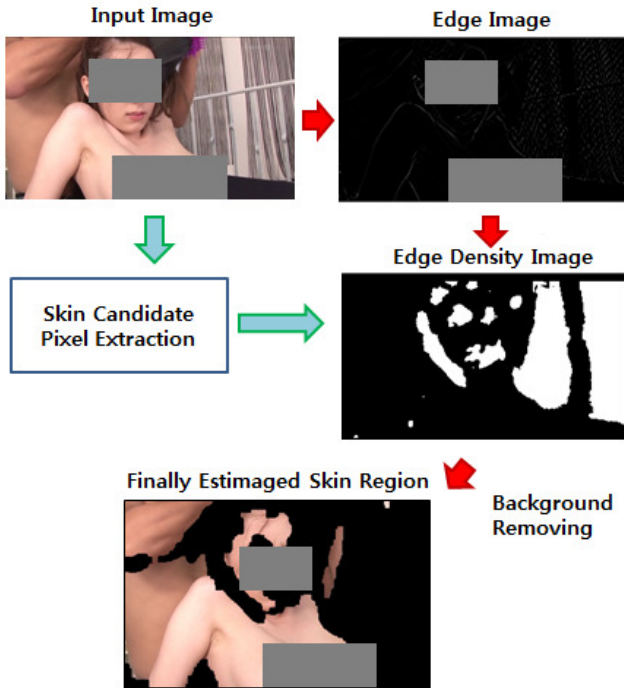


Figure 2. Background removing process using an edge density map

In Fig. 3, we compared the result of proposed background removing method using edge density image and that of simple method seperating the input image into blocks and compute average edge in each block.

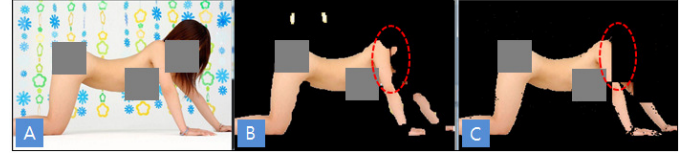


Figure 3. An input image (A), the result of proposed background removing method using edge density image (B), the result of simple method seperating the input image into blocks and compute average edge in each block (C)

The third step includes morphology operations to reduce the effect of pepper and salt noise. We apply closing operation and opening operation to the estimated alpha image and the final alpha image for skin region is estimated.

The last step is to decide the possibility of the input image according to the ratio of the skin region. We used training images composed of an objectionable image class and a non-objectionable image class. Means and standard deviations of the ratio of the skin region are computed in these two classes. The distribution of the ratio is assumed to be gaussian, and an optimal threshold value to distinguish the two classes is estimated by the equation (2). In this equation, a coefficient K is used to impose a larger penalty for a false negative decision than the penalty for a false positive decision because Recall rate is more important than Precision rate for a pre-filtering system.

$$Error = K * \int_{-\infty}^{th} P(x|Obj) dx + \int_{th}^{\infty} P(x|NonObj) dx \quad (2)$$

$$th_k = \min_{th} (Error)$$

Using this threshold value, the input images are classified into two classes, a objectionable candidate class and a distinctive non-objectionable class. Fig. 4 shows this process.

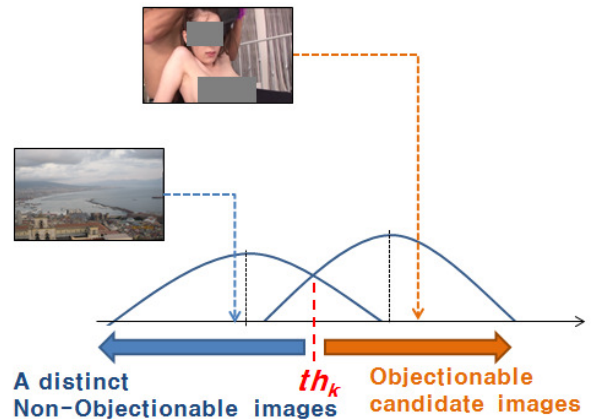


Figure 4. The decision of the objectionable candidate images

IV. EXPERIMENTAL RESULTS

Fig. 5 and Fig. 6 shows the result of skin region estimation using the proposed method. The left image is the input image and the right image is the result of skin region extraction after above three steps.

Fig. 5 shows the result of objectionable input images and Fig. 6 shows the result of non-objectionable images. Obviously, the ratios of skin regions from the objectionable input images are larger than those from the non-objectionable images.

We used the proposed method for the pre-filtering to filter out the clear non-pornographic images before the main identification process. We used a threshold method about the ratio of estimated skin region and regarded the input image as a pornographic image if the skin ratio is bigger than the estimated threshold value th_k in equation (2).

Fig 7 shows the ROC curve for this simple pre-filtering method. Recall is the true positive ratio for the pornographic images and FPR means the false positive ration. For the test, we used 500 objectionable images captured from pornographic videos and 500 non-objectionable images captured from normal movies.

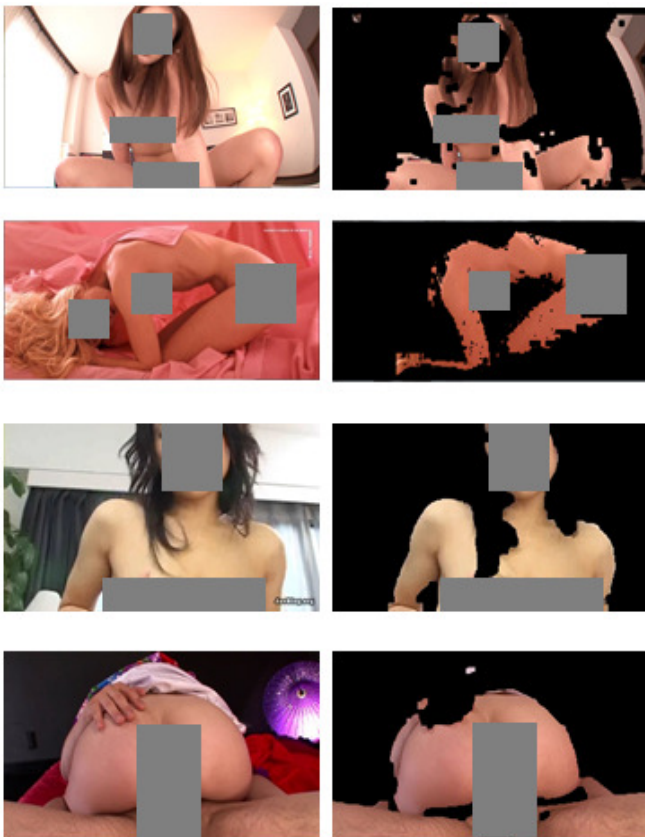


Figure 5. The left column images are the objectionable input images and the right column images are the result of the skin region extraction.



Figure 6. The left column images are the non-objectionable input images and the right column images are the result of the skin region extraction.

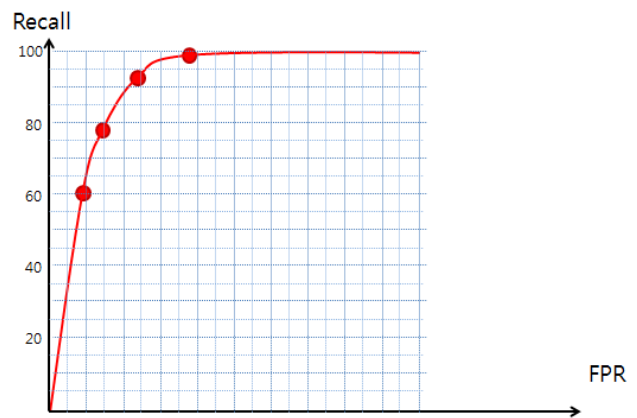


Figure 7. ROC curve of objectionable candidate image estimation using the proposed skin region estimation and simple threshold for decision.

ACKNOWLEDGMENT

This research was funded by the MSIP(Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

REFERENCES

- [1] M. J. Jones, J. M. Rehg, "Statistical color models with application to skin detection," *Computer Vision and Pattern Recognition*, vol. 1, 1999.
- [2] Jiann-Shu Lee, Yung-Ming Kuo, Pau-Choo Chung, E-Liang Chen, "Naked image detection based on adaptive and extensible skin color model," *Pattern Recognition*, vol. 40, Issue 8, pp2261 - 2270, Aug. 2007
- [3] Jau-Ling Shih, Chang-Hsing Lee, Chang-Shen Yang, "An adult image identification system employing image retrieval technique," *Pattern Recognition Letters*, vol. 28, Issue 16, pp2367 - 2374, Dec.
- [4] A. N. Ghomsheh, A. Talebpour, "A New Skin Detection Approach for Adult Image Identification," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4(21), pp 4535-4545, 2012



Jung-Jae Yu received a B.S. degree in electronic engineering from Yonsei University, Rep. of Korea, in 2003 and a Master degree in electronic engineering from KAIST, Rep. of Korea, in 2005. Since 2005 he is a senior researcher in Electronics and Telecommunications Research Institute (ETRI). His research interests include computer vision, pattern recognition, machine learning and 3D reconstruction



Seugn-Wan Han received a B. S. degree in Computer Science from Chonnam National University, Rep. of Korea, in 1994 and received Ph.D. degree in Computer and Statistics from Chonnam National University, Rep. of Korea, in 2001. He was a adjunct professor at University of Science & Technology (UST), Rep. of Korea from 2005 to 2007. Since 2001 he is a principal researcher in Electronics and Telecommunications Research Institute (ETRI). His research interests include image processing, machine learning and security technology.