The Naked Image Detection Based on Automatic White Balance Method

Yung-Ming Kuo[1], Jiann-Shu Lee[2] and Pau-Choo Chung[1]
[1] Department of Electrical Engineering, National Cheng Kung University
1, Ta-Hsueh Rd., Tainan,
[2] Department of Information and Learning Technology, National University of Tainan
33, Sec. 2, Su-Lin St. Tainan,
Taiwan.
Email: ymkuo@ee.ncku.edu.tw

Abstract

This paper presents a new naked image detection system. We propose the skin tone sampling algorithm based on automatic white balance method. The goal of white balance is to correct the image such that it looks as if it is taken in the canonical light. First, the uniform skin chromatic distribution is established manually from the skin images taken under the canonical light. After the target image is corrected by the white balance method based on the canonical light, the skin objects in the target image can be segmented by the uniform skin chromatic distribution. The texture feature, coarseness, is also utilized to acquire the more accurate skin regions. Then, the back propagation neural network is used to integrate several low-level and reliable geometrical constraints into a classifier to inspect the skin regions further. Finally, the mug shot exclusion procedure is applied to promote the system performance. The experimental results show our method is satisfactory for naked image detection.

Key Words

Naked image detection, face detection, automatic white balance, skin segmentation.

1. Introduction

Because of its anonymous and often anarchic structure, images that would be illegal to sell even in adult bookstores can be easily transferred to homes and schools through the Internet. That results in a serious problem that juveniles may intentionally or unintentionally see those adult images. Therefore, how to effectively block or filter out pornography has been attracting a rising concern in related research areas.

The mostly used approach to blocking smut from the Internet is based on contextual keyword pattern matching technology. It categorizes URLs by means of checking contexts of web pages and then traps the websites assorted as the obscene. Although this method can successfully filter out a mass of obscene websites, it cannot deal with images. This would result in missing detecting those obscene web sites containing adult images instead of smut text. Besides the threat coming from the web sites, plenty of the e-mail image attachments are naked. Hence, to prevent juveniles from contacting with pornographic contents from the Internet more thoroughly the development of adult image detection technology is desired.

Adult image detection using semantic matching is a hard task. It should deal with jointed objects of highly variable shapes, in a diverse range of poses, seen from many different views. Furthermore, lighting and background are uncontrolled, making segmentation very difficult. Bosson et al. [1] present a method to block the pornographic images. The authors compute the likelihood ratio for a quantized color space. Then, the blobs' features of the image, detected by likelihood histogram with 25^3 bins in RGB space, are computed and presented as a vector. Finally, the artificial neural network is utilized to classify the image is pornographic or not. Jedynak et al. [2] propose a statistical model for skin detection. The Maximum Entropy Model is used to infer the skin models from the data set. Then, the Bethe Tree approximation and Belief Propagation algorithm are utilized to approximate the probability for skin at pixel locations. H. Zheng et al. [3] propose a method to detect the adult images. The architecture is divided into two parts. The skin detecting model, similar to [2], is applied to detect the skin blocks in the image first. Next, the features of skin blocks in the testing image are fed into the Multi-Layer Perceptron Classifier to identify that is an adult image or not. In [4-5] Forsyth and Fleck proposed an automatic system for telling whether there are human nudes present in an image. That system marks skin-like pixels using color and texture properties. These skin regions are then fed to a specialized grouper, which attempts to group a human figure using geometric constraints on human structure. If the grouper finds a predefined structure, the system decides a human nude is present. The experimental results show that algorithm can extract 43% of the test images. It can be noted that neither of the above methods consider the inference coming from different light source and color altering. By our observation, a significant portion of naked pictures are often snapshotted under the lighting of different light source, such as warm lighting or cool lighting, to make skin look more attractive. However, the different light source owns the color temperature individually. The distribution of human skin in the color space would also deviate with the different color temperature. If the skin color model cannot tolerate the deviation, a lot of naked images would be omitted. On the contrary, if the skin color model accommodates the deviation, a lot of non-skin objects like wood, desert sand, rock, foods, and the skin or fur of animals in the background of image would be detected in the skin detection phase and deteriorates the system performance.

To develop a robust naked image detection system we emphasize two issues: the promotion of the skin detection accuracy and the use of low-level whereas reliable geometrical constraints on skin regions. For the skin detection part, we propose the skin tone sampling algorithm based on automatic white balance method. The goal of white balance is to correct the image such that it looks as if it is taken in the canonical light. First, the uniform skin chromatic distribution is established manually from the skin images taken under the canonical light. After the target image is corrected by the white balance method, the non-skin objects of the target image can be filtered out by the uniform skin chromatic distribution. As the performance of the naked image detection extremely depends on the accurate skin segmentation and there exist various non-skin materials with skin like chroma, these ambiguities should be excluded from as exhaustively as possible. An appropriate feature to play this role is the coarseness because the skin property is very smooth. Regions satisfying the above both conditions are called the smooth skin regions. Investigating a mass of adult images we induce several properties. First, for delighting viewers the protagonist area should be significant. Second, to harmonize with the frame the protagonist should appear around the center. Third, the aspect ratio of the human body should be in a reasonable range. To consider these properties simultaneously, the back propagation neural network is utilized to integrate them into a classifier to examine if the smooth skin region is a naked body. Subsequently, the face detection procedure is used to filter out the mug shot since it meets all the above requirements.

Because the target image has been corrected by the automatic white balance method based on the canonical light, it means the skin color deviation resulting from the special lighting can be accommodated without sacrificing the accuracy. The utilization of the roughness feature can further reject confusion from non-skin stuff thus our skin detection approach can get not only high detection rate but also high detection precision. In addition, because of the extraction of geometrical constraints about naked bodies from the low-level primitives, which are more reliable, instead of the high-level ones, the overall system can achieve satisfactory naked image detection performance. In section two, the skin tone sampling algorithm based on automatic white balance method is introduced. The post-processing employed to enhance the system accuracy is presented in section three. Section four demonstrates the experimental results. Finally, we make conclusions in section five.

2. Skin Tone Sampling Algorithm based on Automatic White Balance Method.

Automatic white balance is one of the most important keys for an image to achieve high quality. Human eyes have the ability to adjust their spectral response to overcome the different color temperature conditions and achieve the color consistency, but the cameras do not. When a white object is illuminated under a low color temperature, it will appear reddish in the stored image. Likewise, it will appear bluish under a high color temperature.

By our observation, most of the naked images utilize the characteristic of color temperature to alter skin color to make nude look more charming. However, skin color altering also makes the distribution of human skin color deviate in the color space. The Fig. 1 and Fig. 2 prove the results. The image taken under the 5500K for color temperature is shown as Fig. 1(a). The Fig. 1(b) and Fig. 1(c) show the images of Fig. 1(a) taken under the 3500K and 8500K. We obtain the three skin regions from Fig. 1(a), Fig. 1(b) and Fig. 1(c) individually. And, the area and position for the three regions are the same. Then, we choose the YCbCr to present the distribution of skin color. The Fig. 2 shows the results. In this figure, the red, green and blue zones show the skin color distributions of three skin regions for 3500K, 5500K and 8500K respectively. If the target image could be corrected by white balance method to the same canonical color temperature before segmenting the skin object, the drawback of skin color deviation would be warded off. Based on this consideration of color consistency, we proposed the skin tone sampling algorithm based on automatic white balance method.

2.1. Automatic White Balance Method

In this section, we will introduce the automatic white balance method. Generally, an automatic white balance method is divided into two steps. First is the estimation of the color temperature. The illumination condition is estimated by picking out gray color points from an image. The color of points has a little deviation under a non-canonical light source. Second, adjusting R channel gain and B channel gain to realize the automatic white balance adjustment. However, the crucial point is the estimation of color temperature, and its accuracy directly affects the second step's correctness. Existing automatic white balance methods can be divided into two categories. First are the global algorithms that use all pixels of an image for color temperature estimation. Second are the local algorithms because only these pixels which satisfy some special conditions are concerned.

The method adopted by our paper is based on the [6]. They proposed a novel method to decide the dynamic threshold to detect the reference white points in an image. According to our experiments, the correction effect of this method is excellent.

2.2. The Uniform Skin Tone

To separate the luminance component from the chromatic one, there are many color spaces can be chosen, such as *TSL*, *HIS*, *YUV*, *CIE Lab*, *CIE Luv*, and *YCbCr* etc. Because the *YCbCr* space performs this well we choose the *CbCr* components as the uniform chromatic distribution bases.

For the establishment of the uniform skin chromatic distribution, we firstly collect the training images with human skin that are taken under the canonical light. Second, we utilize the categories of human skin in the training images to establish the uniform skin chromatic distribution, and denoted as H. An example is shown in Fig. 3. To fit the skin distribution more completely, the principal component analysis (PCA) is applied to H. We can get two eigenvectors and then obtain the minimum elliptic skin chromatic distribution containing H like Fig. 4. The fitting result is denoted as H_e. The source's image is shown in Fig. 5. It is segmented out by setting the pixels with chromas in H_e to one and others are set to zero and the result is shown in Fig. 6. Then, the source's image is corrected by automatic white balance. It is shown in Fig. 7. We use the above method to segment out the skin objects and the result is shown in Fig. 8. According to the comparison between Fig. 6 and Fig. 8, the correct ratio of skin detection is indeed improvable. For the image of Fig. 8, we also get its binary image and denoted as S.

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. The roughness is calculated from the extreme number. First, we get an intensity image I from the test image. And, we create an extreme image X with the same size of I and each element is set to zero. For each pixel of I, if it is an extreme the corresponding element in X is set to one. Subsequently, the X is divided into blocks with size 8x8. If the extreme number of a block is less than a threshold value t, that block is set to one. Otherwise, that block is reset to zero. Therefore, the zero-block in X represents the rough region, whereas the one-block in X corresponds to the smooth region. The coarseness binary image corresponding to the Fig. 7 is demonstrated in Fig. 9. The smooth skin region *SSR* can be obtained by the multiplication of S and X.

3. The Post-processing

Investigating a mass of naked images we generalize several features. The first is that the body occupies a significant portion of the image. Second, the position of the body is almost near to the image center. Third, the body shape is usually not elongate too much. To consider these features simultaneously, the back propagation neural network is used to integrate them into a classifier. If the target image passes the classifier, the mug shot feature is checked. The post processes include the following four parts.

Part 1. Maximum Area Feature

To further separate the body from the skin like background in the *SSR* we apply the opening operation for the *SSR* and get the outcome *OSS*. We locate the body area by picking out the maximum object, denoted as *Max_O*, from the *OSS*. The area of *Max_O* is denoted as *AMO*. The occupation ratio of *AMO* in I is viewed as the first feature. The maximum skin object area corresponding to the Fig. 7 is shown in Fig. 10.

Part 2. Location Feature

The location of *Max_O* is utilized to check if the area is close to the image center, we used the gravity center as the feature of location.

Part 3. Shape Feature

To make the explanation more convenient we use two figures for examples. Applying the principal component analysis (PCA) to the *Max_O* in Fig. 11, we can get two eigenvectors. By projecting the *Max_O* to the two eigenvectors, we can obtain a minimum rectangle containing the *Max_O* like Fig. 12. Assume the lengths of two sides of the rectangle R are L and W (and $L \geqq W$). The aspect ratio Q can be computed by:

$$W\_avg = \text{Area}(Max\_O)/\text{L} \quad (1)$$

$$Q = L/W\_avg \quad (2)$$

The aspect ratio Q is viewed as the third feature.

Part 4. Mug Shot Verification

We refer to [7] for detecting face from the *Max_O*. If the ratio of the face area to the *Max_O* area is less than a threshold value, the test image is viewed as a naked image. Otherwise, the test image is considered not a naked image.

4. Experimental Results

Images were collected from the Internet. There are 150 human naked images and 300 non-naked images. The naked set comprises Caucasians, Blacks and Asians. The non-naked set comprises clothed people, nature scenes, buildings, wood, foods, rock, desert sand, and animals. For the naked test images, the detection rate is 81%. For the non-naked test images, the detection rate is 93%. The overall correction rate is 89%.

5. Conclusion

In this paper, we propose a new naked image detection algorithm. Our method enhances two issues. One is about how to overcome the chromatic deviation coming from the special lighting without at the cost of increasing false alarm. For this issue, we propose the skin tone sampling algorithm based on automatic white balance method to achieve it. The other issue is about how to use low-level whereas reliable geometrical constraints to further inspect the skin regions. For this issue, we derive three geometrical primitives including area, position, and shape. Then, the back propagation neural network is used to integrate them into a classifier. Finally, the mug shot exclusion procedure is utilized to further promote the system performance. The experimental results show our method is satisfactory for naked image detection.

6. References:

[1]. A. Bosson, G.C. Cawley, Y. Chian, R. Harvey, "Non-retrieval: blocking pornographic images," *Proc. Intl. Conf. on the Challenge of Image and Video Retrieval, Lecture Notes in Computer Science*, Springer-Verlag, London, 2383:50-60, 2002.

[2]. B. Jedynak, H. Zheng, M. Daoudi, "Statistical Models for Skin Detection," *IEEE Workshop on Statistical Analysis in Computer Vision*, in conjunction with CVPR 2003 Madison, Wisconsin, June 16-22, 2003.

[3]. H. Zheng, M. Daoudi and B. Jedynak, "Blocking Adult Images Based on Statistical Skin," *Electronic Letters on Computer Vision and Image Analysis*, Volume 4, Number 2, pp. 1-14, 2004

[4]. D.A. Forsyth and M.M. Fleck, "Body plans," *Conf. Computer Vision and Pattern Recognition*, pp. 678-683, 1997.

[5]. D. Forsyth and M. Fleck, "Automatic detection of human nudes," *International Journal of Computer Vision*, 32(1), pp. 63-77, 1999.

[6]. Ching-Chih Weng, Chen, H. and Chiou-Shann Fuh, "A novel automatic white balance method for digital still cameras," *IEEE International Symposium on Circuits and Systems*, Vol. 4, pp. 3801-3804, 2005.

[7]. Yao-Hong Tasi and Yea-Shuan Huang, "Fast Hierarchical Face Detection," *IEEE 2003 International Conference on Multimedia and Expo*, pp. 301-304, 2003.

(a) 5500K (b) 3500K (c) 8500K

Fig. 1: The images for the different color temperature.

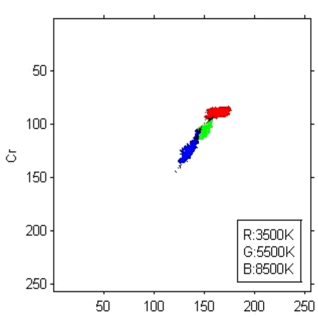

Fig. 2: The skin chromatic distributions for the different color temperature.

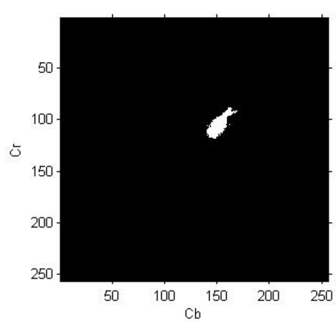

Fig. 3: The skin chromatic distributions

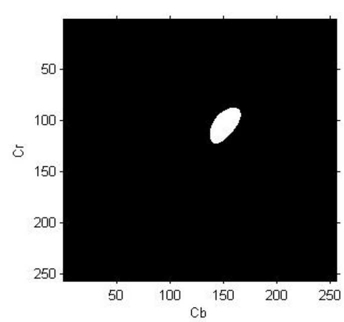

Fig. 4: The elliptic skin chromatic distributions

Fig. 5: The Source Image.

Fig. 6: The skin objects of Fig.5.

Fig. 7: The corrected Image.

Fig. 8: The skin objects of Fig. 7.

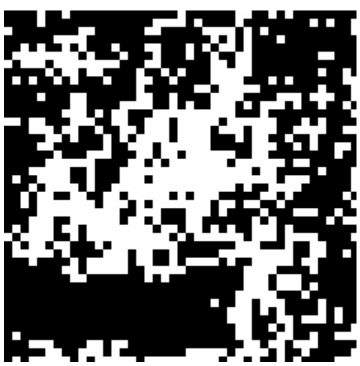

Fig. 9: The smooth Image.

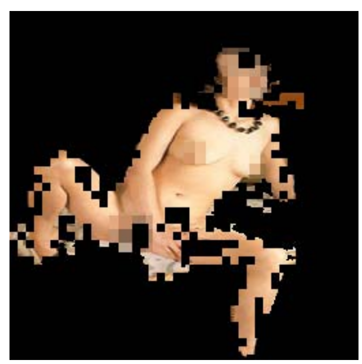

Fig. 10: The maximum skin object.

Fig. 11: The maximum object.

Fig. 12: The minimum rectangle containing the maximum object.