Appearance-Based Nude Image Detection

Chi-Yoon Jeong
ETRI
161 Gajeong-dong, Yuseong-gu, Daejeon, 305-350
Republic of Korea
iamready@etri.re.kr

Jong-Sung Kim and Ki-Sang Hong
Dept. of EEE, POSTECH
San 31, Hojadong, Namgu, Pohang, 790-784
Republic of Korea
{kimjs,hongks}@postech.ac.kr

Abstract

In this paper, we propose an appearance-based nude image detection system. The proposed system is novel in that shape information is used to classify the nude images, and detect small nude images in a large background image. First, the proposed system finds skin regions using texture characteristics of the human skin, which 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 as a high level feature to classify the nude images. The image feature vector (resized skin likelihood image) is used as an input to a nonlinear-SVM. Experimental results show that the proposed system can achieve an excellent classification performance. Additionally, the proposed system can detect small nude images in a large image.

1. Introduction

The Internet is one of the greatest inventions of all time, but has also become a playground for pornographers. Aggressive marketing tactics and tricks are used to target and prey upon innocent children. Therefore, a filtering system is necessary for blocking nude images. Most of the filtering systems block the access to objectionable sites by comparing IP address/URLs lists. However, the internet is very dynamic and web sites are always changing, these methods have a limitation on refreshing lists. Therefore, an image analysis technique is needed in order to classify objectionable images and block accessing to objectionable sites.

Forsyth and Fleck [1] viewed the human body as an assembly of cylindrical parts and used a geometric grouper to inspect the presence of naked human structures. Jones and Rehg [2] locate skin pixels using statistical skin/non-skin color models based on histogram and then compute aggregate features for nude image detection. Wang et al.[3] proposed a nude image detection system using Daubechies's wavelets, normalized central moments, and a color histogram. Bosson et al.[4] finds skin blobs, and then computes the area, centroid, length of the axes of an ellipse, eccentricity, solidity, and extent of skin blobs. These methods use aggregate features computed from the binary skin region image. These features mostly represent the percentage of skin within an image. Therefore, it is not sufficient for high accuracy nude image detection. To increase the accuracy of nude image detection, the high level features should contain the shape information of the nude images. Therefore, we proposed an image feature vector as a new high level feature. This feature vector contains shape and skin color information, so it will be a good feature to separate nude images from non-nude images. Also, using an image feature vector, the proposed system can detect small nude images in a large background image.

The remainder of this paper is organized as follows. In section 2, the proposed system is described. In section 3, it is explained how to make a image feature vector. Picture in Picture problem is described in section 4 and experimental results are given in section 5. In section 6, we conclude this work.

2. Nude image detection system

The proposed system contains three major modules as in figure 1. The face detection and skin filter are used to classify the non-nude images. The image feature vector is generated from the image passing skin filter and then used as an input to a nonlinear-SVM which determines whether images are "nude" or not.

2.1. face detection

The images, containing many faces or close-up faces, as illustrated in figure 2, can be classified as nude images. To resolve this problem, the proposed system uses the face detection algorithm of Viola and Jones[5]. After performing face detection, if the total face area or the largest face area

0-7695-2128-2/04 $20.00 (C) 2004 IEEE

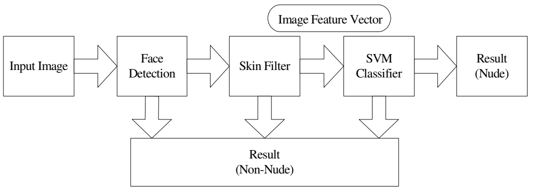

Figure 1. Proposed nude detection system

is greater than the specified threshold, the image is classified as a non-nude image. Face detection results are shown in figure 2.

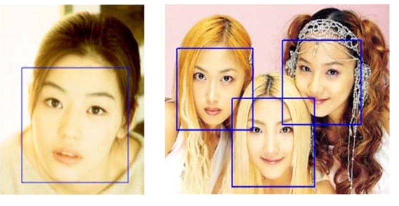

Figure 2. Face Detection results

2.2. Skin Filter

Since there is a strong correlation between the percentage of skin and a nude image, identifying skin pixels is important. RGB color space is used to identify the skin pixels. Skin color and non-skin color are modelled using Gaussian mixture models,

$$P(\mathbf{x}) = \sum_{i=1}^{N} \omega_i \cdot \frac{1}{(2\pi)^{\frac{3}{2}} \mid \Sigma_i \mid^{\frac{1}{2}}} \exp^{-\frac{1}{2}(\mathbf{x}-\mu_i)^T(\Sigma_i)^{-1}(\mathbf{x}-\mu_i)}$$

where $\mathbf{x}$ is the RGB color vector, N is the number of Gaussian models, ω_i is the mixing parameter, μ_i, Σ_i is the mean and covariance matrix of Gaussian models. Two separate 16 mixture of the Gaussian models for skin color and non-skin color are used.[2]

The skin likelihood ratio map is generated using the probability of skin and non-skin color;

$$L(\mathbf{x}) = \frac{P_{skin}(\mathbf{x})}{P_{non-skin}(\mathbf{x})}$$

The original image and skin ratio map are shown in figure 3 (a) and (b), respectively.

The adaptive skin segmentation method[6], one of the region-based skin detection methods, is used to find skin regions. Skin regions have a homogeneous property, which is the texture characteristic of human skins. But pixel-based methods do not use skin homogeneity. Therefore, region-based skin detection methods can locate skin regions more exactly than pixel-based methods.

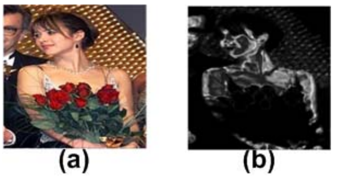

Figure 3. (a) Original image (b) skin likelihood ratio map

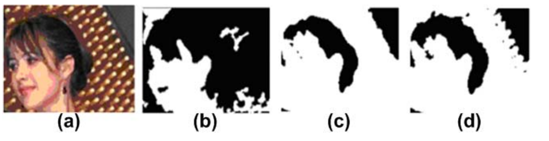

Figure 4. Skin detection results (a) Original image (b)Adaptive skin segmentation (c)Pixel-based method (th=1.0) (d)Pixel-based method(th=2.0)

The original image and adaptive skin segmentation result are shown in Figure 4. Pixel-based skin detection results are shown in (c), (d) with different threshold values. Pixel-based results have false skin regions which are non-homogeneous regions. But adaptive skin segmentation results show that most false skin regions are not detected. Using adaptive skin segmentation, the proposed system can detect skin regions more accurately.

2.3. SVM classifier

SVM map the data to a predetermined very high dimensional space via a kernel function and find the hyperplane that maximizes the margin between the two classes. So SVM classification is extremely efficient and robust. In this paper, SVM [7] is used to characterize the appearance of nude and non-nude images. The High level feature calculated from skin region image is used as an input to an SVM.

3. Image Feature Vector

Many researchers attempt to find high level features that can separate nude images and non-nude images. The high level features used by other researchers [2][4] contain only skin region information, for example, the number of skin blobs, the area of skin blobs, and the number of colors. So the non-nude images that contain many skin regions are classified as nude images and the nude images that contain several small skin regions are classified as non-nude images. To solve this problem, the shape information of the nude images is required. It is not an easy task to characterize the shape information of the nude images. Because the nude images may have complicated backgrounds. However, in 5(a), we recognize an appearance similarity between the skin region images of the nude images. An appearance difference between nude images and non-nude images also exists, as illustrated in Figure 5.

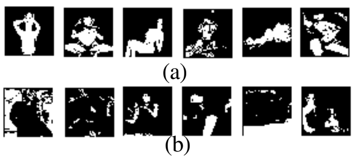

Figure 5. Skin likelihood images (a) nude images (b) non-nude images

To use the appearance information, we propose the skin likelihood image that has the shape and skin color information of the image. From the skin region image as figure 6 (c), skin likelihood image is generated using the following equation,

$$S(x,y) = \begin{cases} 0 & \text{if } L(x,y) \leq s_0 \\ \frac{255}{(s_1 - s_0)}(L(x,y) - s_0) & \text{if } s_0 < L(x,y) < s_1 \\ 255 & \text{if } L(x,y) \geq s_1 \end{cases}$$

where $L(x,y)$ is the skin likelihood ratio at (x,y), s_0,s_1 is the minimum, maximum threshold for skin ratio. A skin likelihood image is shown in figure 6 (d). We used image feature vectors which resized skin likelihood images to classify the nude images. Resizing is used to reduce the dimension of high level features. Image feature vectors used as an input to a SVM classifier. In figure 6, black areas are marked on purpose.

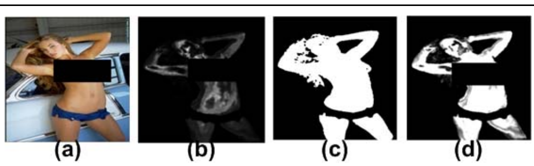

Figure 6. (a) Original images (b) Skin likelihood ratio map (c) Skin region image (d) Skin Likelihood image

4. Picture in Picture Problem

Picture in picture (called PIP) means that large images contain the small nude images as figure 7 (a). In PIP problem, the nude images has small skin regions relative to original image, other methods cannot solve this problem. If the PIP problem can be viewed as the object detection problem, we can solve. The proposed system can detect small nude images using image feature vectors that have the appearance information of the nude/non-nude images.

In the skin region image, illustrated in figure 7 (b), small nude images have the appearance of nude images. Image feature vectors are generated from each skin blob in figure 7 (b). The green rectangular box means that it is a suspicious region. Each image feature vector is used as an input of nonlinear-SVM to classify the small nude image. To reduce the false positive rate, bootstrapping[8]is used.

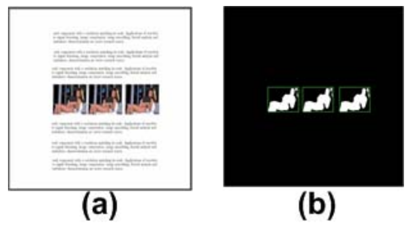

Figure 7. (a) PIP image (b) Skin region image

5. Experimental results

Nude/non-nude images are gathered from various web sites. The total number of nude images is about 4600, and the total number of non-nude images is about 13000. Non-nude images contain animals, people, buildings, cars, scenery, and so on. Accuracy, precision, and recall are used as the performance measures. Accuracy is defined as the ratio of the number of correctly classified images to the total number of test images. Precision is defined as the ratio of the number of nude image identified to the total number of images that classified as nude images. Recall is defined as the ratio of the number of nude image identified to the total number of nude images.

Images passing the face detection and skin filter are used as the training and test set for nonlinear-SVM. The training set contains 2325 nude images, and 2252 non-nude images. The test set contains 2341 nude images, and 2172 non-nude images. The image feature vectors are generated from the training set and then used as the input of the nonlinear-SVM. Radial basis function $\exp -(\gamma \|a-b\|^2)$ is used as an SVM kernel function.

SVM classification result is shown in Table 1. Five features proposed by Bosson[4] are used as SVM feature vec-

High level featurs	Accuracy	Precision	Recall
Image Feature Vector(γ=0.0005)	94.37%	94.16%	95.04%
5 Features(γ=0.005)	80.41%	79.80%	83.34%

Table 1. SVM classification result

tors for the performance comparison. The result show that, the proposed features have a good separability. The overall classification result is shown in Table 2. Most non-nude images are classified as non-nude images by the face detection and skin filter stage. Correctly classified images by proposed system are shown in figure 8 and 9. These images cannot classify correctly by the 5 features method. This result shows that appearance-based nude image detection system has a high performance. Although the proposed high level features have high dimensions, it takes about 15ms to evaluate the SVM. It takes about 1 sec. to perform the proposed system. The PIP detection result is shown in figure 10. Each blue rectangular box represents the small nude image detected. This result show that the PIP problem can be solved.

	Skin Filter Face Detector(Non)	SVM (Nude)	SVM (Non)
Nude(2413)	72	2225	116
Non(6525)	4353	138	2034
Accuracy: 96.35% , Precision: 94.16 % , Recall: 92.20 %			

Table 2. Overall classification result

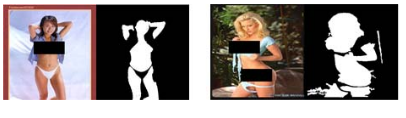

Figure 8. Correctly classified nude images

Figure 9. Correctly classified non-nude images

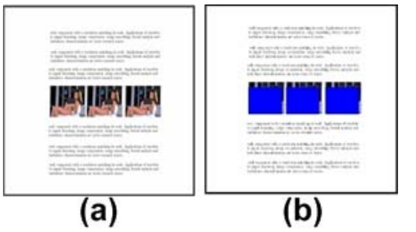

Figure 10. PIP detection result

6. Conclusion

In this paper, we proposed the appearance-based nude image detection system. It is novel in that shape information used to classify the nude images. In addition, texture characteristics of the human skin are used to find skin regions more exactly. The PIP problem can be solved using the proposed method. Experiment results show that proposed method has an excellent performance in classifying nude images and detecting small nude images in large images.

Simple skin color model is used to skin color detection. However, because nude detection system require a precise skin region images, more research on skin color model is necessary.

References

[1] M. Fleck, D. Forsyth, and C. Bregler. Finding Naked People. *In European Conf. on Computer Vision*, vol. II:592–602, 1996.

[2] M. J. Jones and J. M. Rehg. Statistical Color Model With Application to Skin Detection. *In Technical Report CRL*, 98/11:1–12, 1998.

[3] J. Z. Wang, G Wiederhold and O. Firschein. System for Screening Objection-able Images. *Computer Communications*, vol.21:1355–1600, 1998.

[4] A. Bosson, G. C. Cawley, Y. Chan and R. Harvey Non-Retrieval:Blocking Pornographic Images. *In International Conf. on Image and Video Retrieval*, 2002.

[5] P. A. Viola and M. J. Jones. Robust real-time object detection. *In Technical report COMPAQ Cambridge Research Laboratory*, Feb. 2001.

[6] S. Phung, A. Bouzerdoum, and D. Chai Adaptive Skin segmentation in color images *In International Conf. on Acoustics Speech and Signal Processing*, 2003

[7] B. Scholkopf and C. Burges and A. Smola Advances in Kernel Methods - Support Vector Learning *MIT-Press*, Cambridge, USA, 1998.

[8] T. Poggio and K.K. Sung Example-based learning for view-based human face detection. *In Proc. of the ARPA Image Understanding Workshop*, 1994