ADULT IMAGES DETECTION BASED ON COLOR CORRELOGRAM

YIZHI LIU[1,2,3,4], SHOUXUN LIN[1,2], SHENG TANG[1,2], YONGDONG ZHANG[1,2]

[1]Virtual Reality Laboratory, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2]Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, CAS, Beijing, China
[3]Graduate School of the Chinese Academy of Sciences, Beijing, China
[4]Institute of Computer Science and Engineering, Science and Technology College of Hunan, Xiangtan, China
E-mail: liuyizhi@ict.ac.cn, sxlin@ict.ac.cn

Abstract:

Skin-color detection is fast and powerful for adult images detection, but the true positive (TP) and the false positive (FP) are both high. To detect adult images accurately and avoid benign images being misclassified as much as possible, we propose a three-step process followed with Gaussian mixture model (GMM) detector, color correlogram (CC) detector and face detector. The objective of the last two steps is mainly to reduce FP. Different from previous systems, we adopt CC descriptor which combined the spatial correlation of colors and the global distribution of local spatial correlation of colors. Experiments strongly indicate that above 90% benign images can be detected accurately by CC detector and so FP is reduced effectively.

Keywords:

Adult images detection; color correlogram; skin-color; Gaussian mixture model; texture

1. Introduction

With the rapid proliferation of adult videos and images on the Internet, many social problems become more and more serious, such as cyber-sex addiction. There is a definite need for tools like filtering software to prevent people, especially children, from adult images. However, in some cases, users would not enjoy benign videos and images being filtered because they are misclassified as adult content.

Skin-color detection methods are proved to be fast and powerful. Although high true positive (TP) of adult images is obtained, false positive (FP) is also high so that lots of benign images misclassified, such as bikini images and human faces. Here, TP is defined as the ratio of the number ground truth adult images detected to the total number of adult images, and FP is defined as the ratio of the number of the misclassified benign images to the total number of adult images. Obviously, the sum of TP and FP is equal to 1. Therefore, it is one of the most important issues in adult images detection to improve TP indirectly by reducing FP to meet the needs of different applications.

Our goal is to detect adult images accurately and to avoid benign images being misclassified as much as possible. Obviously, low FP is especially stressed. To reduce FP at the premise of filtering adult images, this paper presents a new approach to separate benign images from adult images by analyzing color correlogram (CC) of images and using face detector after skin-color detection based on Gaussian mixture model (GMM).

The rest of the paper is organized as follows: Related work and CC descriptor are introduced respectively in section 2 and 3. In section 4, our system and details of CC detector are described. Experimental results are shown in section 5 and section 6 concludes the paper.

2. Related work

Different texture analysis techniques were used to reduce FP. Forsyth et al. combined tightly-tuned skin filter and smooth texture analysis to detect whether there are naked people present in an image [1]. After skin detection, the geometric analysis was used to group skin regions into human figure for human body detection. It takes about 6 minutes per image for the figure grouper [2]. Wang et al. presented a system of screening objectionable images for practical applications named WIPE [2]. The images that passed histogram analysis, texture analysis and shape matching were classified as the adult images. But the texture analysis in the WIPE is rather simple. In Image Guarder system [3], a combination of skin feature, texture feature and shape feature are used. The textures in the system are simple, including texture contrast and texture coarseness. The features used by Google's system include simple shape (lines), color, texture and face etc. [4]. However, the two texture features are all skin-dependent. The two texture features outside of skin blobs are not

computed, let alone comparing the texture features in skin blobs and those outside of skin blobs.

There are a few researches focused on separating benign images from adult images. [5] presented an approach to separate benign images from adult images by analyzing a set of semantically high-level features obtained from face detection and adaptive skin detection. [6] utilized the information about the body shape and face to distinguish the benign bikini photos from pornographic photos.

3. CC descriptor

Texture, as a low-level feature for the human visual system, reveals scene depth and surface orientation. Moreover, it describes properties like smoothness, coarseness, and regularity of a region. Most texture analysis techniques can only deal with gray-scale images. However, color is of importance in texture recognition [7]. A color histogram, a method of first order statistical measures, describes the global color distribution in an image. But it does not include any spatial information, and is therefore liable to FP. This problem is especially acute for large databases. Moreover, the histogram is not robust to large appearance changes.

CC is a method of second order statistical measures. The highlights of CC [8] are: (1) it includes the spatial correlation of colors; (2) it can be used to describe the global distribution of local spatial correlation of colors; (3) it robustly tolerates large changes in appearance and shape caused by changes in viewing positions, camera zooms, etc.; (4) it is easy to compute; and (5) the size of the feature is fairly small. According to the Broek's research using six color spaces combined with five different quantization schemes, the HSV 162 bins configuration performs best in combination with CC and using more bins does not improve performance [9]. Furthermore, the gray level is always ignored during distilling the color feature. Since the computation time is proportional to the length of the feature vector, the decrease of the gray level lightens the computational cost [10]. Therefore, we presented a method based on CC descriptor, in combination with HSV 162 bins and four gray levels. Details of the CC feature distilling algorithm can be found in [8].

The basic principles of SVM have been developed by Vapnik in 1995 [11]. SVM supports classification and regression tasks based on the concept of optimal separator. Basically, SVM projects the feature space into high dimension space to find a hyper-plane, which theoretically best separates samples belonging to different classes. We adopted LIBSVM, which is a mature library for support vector classification and regression, in our system.

4. The procedure of our system

The procedure of our system is shown in Fig.1. Our goal is to detect adult images accurately and to avoid benign images being misclassified as much as possible.

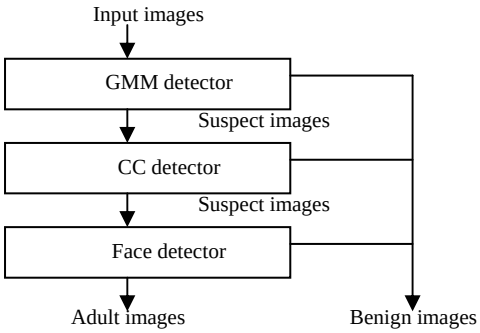

Fig.1 The procedure of our system

4.1. Three-step procedure

Since skin-color detection has been proved to be a powerful cue, we firstly adopt the skin-color detection based on the GMM. After analyzing the results, we find that two types of benign images are always misclassified as adult images because the background color is similar to skin-color but there is few bare body (such as the images in the first line of Fig.2) or skin regions are conspicuous (such as the images in the second line of Fig.2) in the images.

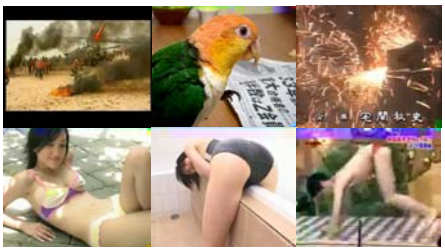

Fig.2 The false positives samples after GMM detector

CC detector is different from previous systems, that is, CC of images is extracted from the results of skin-color detection based on GMM and then fed into a SVM classifier to decide whether or not these images contain adult content. The objective is to effectively reduce FP and consequently improve the detection accuracy. Furthermore, CC is easy to compute and the size of the feature is fairly small.

The objective of the third step is to discriminate the images containing conspicuous faces further when the background color is similar to skin-color. Face detection is now being extensively studied and some effective and efficient face detectors have been introduced in the literature. We chose the face detector in the OpenCV toolkits.

4.2. CC detector

CC detector includes CC descriptor and LIBSVM classifier. In order to reduce FP well and improve the precision in adult images detection, there is something worth noting during training process.

The training data consist of thousands of images collected from the Internet. After detected by GMM detector, the results were annotated manually and divided into adult images and benign images. The first thing needed to notice is how to judge adult images. To us, only when pornographic parts of human bodies or pornographic postures are exposed obviously in the images, the images are regarded as containing adult content.

It is also worth noting how to choose adult images as the positive training set and how to choose benign images as the negative training set for CC detector. After eliminating degraded images from the results detected by GMM detector, we added the misclassified adult images to the negative training set and added the misclassified benign images to the positive training set as much as possible.

5. Experiments

We captured four groups of images from the Internet at different time and implemented two experiments for performance evaluation: one for the skin-color detection based on GMM, the other for CC detector.

Table1. Performance of skin-color detection based on GMM detector

Test Group	Adult Images Detected by GMM Detector (frame)	Adult Images (frame)	Benign Images (frame)	TP	FP
GroupA	38	19	19	50.00%	50.00%
GroupB	721	264	457	36.62%	63.38%
GroupC	9915	2860	7055	28.85%	71.15%
GroupD	13689	6272	7417	45.82%	54.18%

Two performance measures used are TP and FP. The definition of TP is the ratio of the number ground truth adult images detected to the total number of adult images. And FP is defined as the ratio of the number of the misclassified benign images to the total number of adult images.

In the experiment on skin-color detection based on GMM, we can observe that FP is higher than 50% (shown in Table 1). And the fact that TP of adult images detection is lower than 50% is relative to the two following factors. One is related to the testing data, the other is our strict criterion evaluating adult images.

We annotated 12253 frames manually as the training database (including 2181 adult images and 10072 benign images), which were all detected as adult images by the skin-color detection based on GMM. CC of these images is extracted and used to train a LIBSVM classifier. SVM is a popular and powerful technique for binary classification. In the training process, some common techniques, such as scaling feature values and cross-validation based model selection, were applied.

Table2. Effects on FP in CC detector

Test Group	Adult Images Detected by GMM Detector (frame)	Benign Images Misclassified by GMM Detector (frame)	Benign Images Misclassified by CC Detector (frame)	Accuracy of Detecting Benign Images	Processing Time (second per frame)	FP
GroupA	38	19	1	94.74%	0.95	2.63%
GroupB	721	457	60	86.87%	0.55	8.32%
GroupC	9915	7055	159	97.75%	0.76	1.60%
GroupD	13689	7417	1209	83.70%	0.79	8.83%
Average	24363	14948	1429	90.44%	0.76	5.87%

Fig.3 The benign images detected by CC detector

Fig.4 The benign images misclassified by CC detector

According to the results in Table2, we may find that CC detector is powerful to detect benign images and the

accuracy is generally higher than 90%. That is to say, the number of benign images misclassified as adult images in CC detector is far lower than that in GMM detector. Therefore, the FP of CC detector is lower than 10%. However, there are negative effects on the recall ratio of adult images and the consumed time.

After analyzing the results of CC detector, we observe that many images, the background color is similar to skin-color but there is few bare body (such as the images in the first line of Fig.2) or skin regions are conspicuous but the regions in skin-color are different from the background color (such as the images in the second line of Fig.2 and those in Fig.3), are detected accurately. Nevertheless, the images containing conspicuous skin-color regions, in which the background color is very similar to the background color (such as the images in Fig.4), are still misclassified as adult images. So we adopted face detector as the third step.

Our experiments strongly indicate that CC detector is able to reduce FP of adult images detection because it can effectively recognize benign images whose background color is similar to skin-color. Both the spatial correlation of colors and the global distribution of local spatial correlation of colors are considered in CC detector. That is the key point to success for our system.

6. Conclusions

To detect adult images accurately and avoid benign images being misclassified as much as possible, the proposed approach follows a three-step process, including GMM detector, CC detector and face detector. The main objective of the last two steps is to reduce FP.

CC is both a kind of color feature and a kind of texture feature. CC descriptor describes the global distribution of local spatial correlation of colors and the spatial correlation of colors. In principles, adult images detection based on CC and skin-color detection would complement each other because the former can reduce FP effectively. Experiments strongly indicate the preceding viewpoints.

Acknowledgements

This work was supported in part by the National Basic Research Program of China (973 Program, 2007CB311100), the National High Technology and Research Development Program of China (863 Program, 2007AA01Z416), the National Nature Science Foundation of China (60773056), the Beijing New Star Project on Science & Technology (2007B071), the Knowledge Innovation Project of The Institute of Computing Technology, Chinese Academy of Sciences (20076031), the National Nature Science Foundation of China (60873165), and the General Project of the Education Department of Hunan Province (06C348).

References

[1] M. M. Fleck, D. A. Forsyth, and C. Bregler, “Finding Naked People”, Proc. European Conf. on Computer Vision, 1996, 2, pp. 593-602.

[2] J. Z. Wang, J. Li, G. Wiederhold, and O. Firschein, “System for Screening Objectionable Images”, Computer Communication, 1998, 21(15), pp. 1355-1360.

[3] W. Zeng, W. Gao, T. Zhang, and Y. Liu, “Image Guarder: An Intelligent Detector for Adult Images”, Asian Conference on Computer Vision, 2004, pp. 198-203.

[4] H. A. Rowley, J. Yushi, and Shumeet Baluja, “Large Scale Image-Based Adult-Content Filtering”, Proceedings of the First International Conference on Computer Vision Theory and Applications, 2006, pp. 290-296.

[5] Qing-Fang Zheng, Ming-Ji Zhang, and Wei-Qiang Wang, “A Hybrid Approach to Detect Adult Web Images”, 5th Pacific Rim Conference on Multimedia, 2004, pp. 609-616.

[6] Yushi Wang, Weiqiang Wang, and Wen Gao, “Research on the Discrimination of Pornographic and Bikini Images”, Proceedings of the Seventh IEEE International Symposium on Multimedia, 2005.

[7] T. Mäenpää, and M. Pietikäinen, “Classification with Color and Texture: Jointly or Separately?”, Pattern Recognition, 2004, 37(8), pp. 1629-1640.

[8] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, “Image Indexing Using Color Correlograms”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 762-768.

[9] E. L. van den Broek, and E. M. van Rikxoort, “Evaluation of Color Representation for Texture Analysis”, Proceedings of the Sixteenth Belgium-Netherlands Artificial Intelligence Conference, 2004, pp. 35-42.

[10] I. Kunttu, L. Lepistö, J. Rauhamaa, and A. Visa, “Image Correlogram in Image Database Indexing and Retrieval”, Proceedings of 4th European Workshop on Image Analysis for Multimedia Interactive Services, 2003, pp. 88-91.

[11] V. N. Vapnik, The Nature of Statistical Learning Theory, Springer Verlag, 1995.