Shape-based Adult Images Detection

Qing-Fang Zheng[1,3], Wei Zeng[2], Gao Wen[1,2,3], Wei-Qiang Wang[1,3]
[1]*Institute of Computing Technology, CAS, P.O.Box 2074, Beijing China, 100080*
[2]*Department of Computer Science and Technology, Harbin Institute of Technology, China*
[3]*Graduate School of Chinese Academy of Sciences, Beijing, China*
E-Mail: {qfzheng, wzeng, wgao, wqwang}@jdl.ac.cn

Abstract

This paper reports an investigation on adult images detection based on the shape features of skin regions. In order to accurately detect skin regions, we propose a skin detection method using multi-Bayes classifiers in the paper. Based on skin color detection results, shape features are extracted and fed into a boosted classifier to decide whether or not the skin regions represent a nude. We evaluate adult image detection performance using different boosted classifiers and different shape descriptors. Experimental results show that classification using boosted C4.5 classifier and combination of different shape descriptors outperforms other classification schemes.

1. Introduction

The steady growth of the Internet, the decreasing price of storage devices and people's increasing interest on images have been contributing to make the Internet an unprecedented large image library. Images now available on-line should be counted in billions. However, among these images, some are offensive and even illegal, for example, pornographic images. Exposure to the sea of pornography can lead to many social problems, including cyber-sex addition. It is now an urgently necessary task to prevent people, especially children， from accessing this type of harmful material.

The research reported here focuses on discriminating adult images from non-adult images based on the shape features of skin color regions in the images. We firstly perform skin detection to binarize images into skin regions and non-skin regions. Next, we extract shape features of these skin regions and use a boosted classifier to decide whether or not these skin regions represent a nude.

The rest of the paper is organized as follows: In section 2, we briefly review current adult image filtering techniques. Details of skin detection and shape classification are described in section 3 and section 4, respectively. Experimental results are shown in section 5 and section 6 concludes the paper.

2. Current adult image filtering techniques

Current adult image filtering techniques can be classified into three categories: keyword-based, blacklist-based and content-based. Keyword-based methods attempt to filter images by analyzing the text that names or surrounds them on a web page. However, many words that belong to the pornographer's lexicon also appear in web pages for education purpose. As a result, keyword-based methods may screen out benign images while admit salacious content. Blacklist-based method screens out images gleaned from blacklisted web addresses where pornography is deemed likely to turn up. But pornography has proved a faster target than such lists can catch.

Content-based techniques evaluate images by direct analysis of image content. The pioneer work was done by Forsyth et al [1]. Their approach combined tightly-tuned skin filter and smooth texture analysis for skin detection. After skin detection, the geometric analysis was used to group skin regions into human figure for human body detection. Wang et al. presented a system of screening objectionable images for practical applications [2]. Their method employed a combination of an icon filter, a graph-photo detector, a color histogram filter, a texture filter and a wavelet-based shape matching algorithm. The images that passed histogram analysis, texture analysis and shape matching were classified as the adult images. In Jones and Rehg's work [3], the adult images were recognized by the skin detector and the neural network classifier.

3. Skin detection

The task of skin detection is to binarize the input color images into skin regions and non-skin regions. Skin detection results can significantly affect the subsequent shape extraction and shape classification. For accurate skin detection, the skin detection method we used contains two main stages: skin pixels detection and skin region refinement.

3.1. Skin pixels detection using multi-Bayes classifiers

Skin detection is an important technique for identifying adult images because of the fact that there is a strong correlation between images with large skin patches and adult images. However, accurate skin detection is a non-trivial task. Skin color varies greatly between different human races and can changes greatly when illumination condition changes. Here we propose a skin pixel detection algorithm using the multi-Bayes classifiers. We use K-Means to pre-group sample images into different clusters according to their average brightness and average chromaticity. For each cluster, we build a Bayes skin classifier as described in [3]. In our method, pixels are represented by the color values RGB, average brightness L and average chromaticity T. The posterior probability of a skin pixel is computed as:

$$P(Skin \mid RGB, L, T) = \frac{P(RGB, L, T \mid Skin)P(Skin)}{P(RGB, L, T \mid Skin)P(Skin) + P(RGB, L, T \mid \neg Skin)P(\neg Skin)}$$

Average chromaticity T is represented by average chromaticity red and green.

$$T = [red_{avg}, green_{avg}],$$

Chromaticity red and green are calculated as:

$$red = R / B, \qquad green = G / B \ .$$

A pixel will be classified as skin pixel if

$$P(Skin \mid RGB, L, T) \geq \theta \qquad \theta \in [0,1] \ .$$

In practice, for an input image, we firstly calculate the average brightness and chromaticity of the image and choose the corresponding Bayes skin classifier accordingly. Then we use the skin classifier to perform skin pixel detection.

3.2 Skin regions refinement

The above mentioned skin detection method is pixel-based, as a result, some non-skin pixels with colors similar to skin may be classified as skin pixels. We repeat morphological operation erosion and dilation three times to refine the skin regions. Our structural element is a disk and the size of the disk used for each morphological operation is set automatically according to the size of the image.

4. Shape classification

After skin detection, we extract the shape features of the skin regions in the images and then we use Adaboost [7] to do classification. The goal is to tell whether or not the objects are nudes according to their appearance.

4.1. Shape descriptors

Shape is an important characteristic of an object. The goal of shape descriptors is to uniquely characterize the object shape. A good shape descriptor should minimize the within-class variance and maximize the between-class variance and be insensitive to noise. Three types of shape descriptors of objects are used in this paper, including:

(a) Three "simple" shape descriptors: eccentricity, compactness and rectangularity. Eccentricity is the length ratio between the major and minor axes of the objects. Compactness is the ratio between the length of object's boundary and the object's area. Rectangularity is the ratio of object area to the area of its bounding box. We refer this descriptor as D1.

(b) Seven normal moment invariants defined by Hu [4]. The seven moment invariants are independent of translation, scale, and rotation. Theoretical analysis has shown that the first invariant measures the total spread of the shape relative to its area square while the second invariant measures the degree of elongation of a best-fit ellipse on the shape. We refer this descriptor as D2.

(c) Zernike moments [5]. The kernel of Zernike moments is the set of orthogonal Zernike polynomials defined over the polar coordinate space inside a unit circle. Moments of different orders correspond to independent characteristics of the image. We refer this descriptor as D3.

Because these three shape descriptors are developed from different rationales, we also use the combination of these features to do classification. We argue that they will complement one another and provide better performance than only using each of them.

4.2. Classifiers

Boosting is a widely used scheme to combine multiple classifiers to increase the performance of a single classifier (weak classifier). We use AdaBoost [6] in this paper. AdaBoost algorithm begins with assigning equal weight to all instances in the training data. It then calls the learning algorithm to form a classifier for this data, and re-weights each instance according to the classifier's output. The weight of correctly classified instances is decreased, and that of misclassified ones is increased. In the next iteration, a classifier is built for the re-weighted data, which focuses on classifying the hard instances. The output of AdaBoost is a combination of weighted vote of each weak classifier. Since any classification function can potentially serve as a weak classifier, we try four weak classifiers in this paper: Decision Stump [7], C4.5 [8], SVM [9] and Multi-Layer Percetron (MLP) [10].We refer to these classifiers as C1, C2, C3 and C4 respectively in the following paper. Our goal is test the performances of these classifiers and choose the best one.

5. Experiment

We conduct two experiments in the performance evaluation: one for skin detection and the other for shape classification. Two performance measures used are true positive (TP) and false positive (FP). In skin detection, TP is defined as the ratio of the number of ground truth skin pixels detected to the total number of skin pixels and FP is the ratio of the number of non-skin pixels misclassified as skin pixels to the total number of non-skin pixels. In shape classification, TP is defined as the ratio of the number ground truth adult images identified to the total number of adult images and FP is the ratio of the number of non-adult images misclassified as adult images to the total number of non-adult images.

In skin detection evaluation, we use 1650 images for training and 138 adult images for test. Performance comparison between our multi-Bayes method and Jones's single Bayes skin detection method [3] is shown in Fig1. For offensive reasons, we don't show the full adult images, but parts of these images.

For shape classification, we use a dataset consisting of 897 adult images and 732 non-adult images. Non-adult images typically contain objects with colors very similar to human skin, such as desert, yellow flowers, lions. In these 1629 images, the ratios of the area of skin color regions to that of whole images are all above 25%. Some example images after skin detection are shown in Fig2. We use fivefold cross-validation to test the performance of shape classification. The 1629 images are split into five approximately equal partitions, and each in turn is used for testing while the remainder is used for training. The procedure repeats five times so that every image has been used for training and test. Performances of five classifications are averaged to yield a general performance measure. Table 1 1ist the performance of four boosted classifiers using different shape features. As a comparison, we also list the performance of corresponding classifiers without AdaBoost in Table 2.

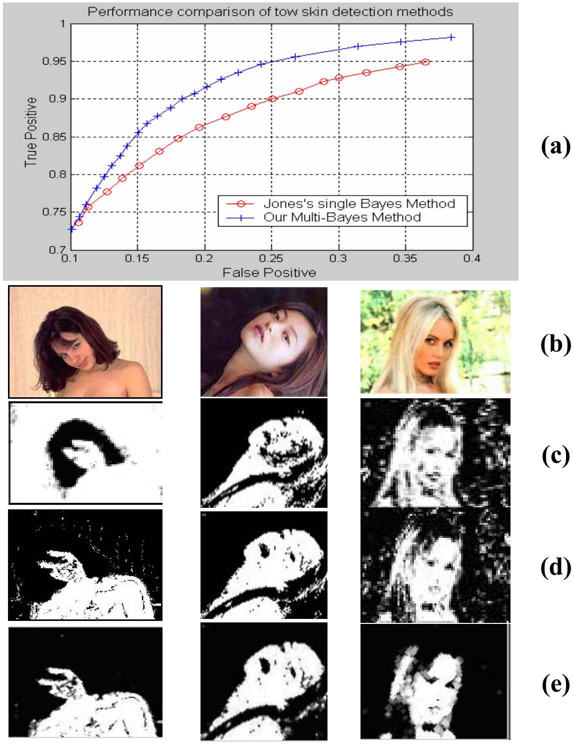

Fig1. Performance comparison between Jones's single Bayes skin detection method and ours. (a) ROC curves of two skin detection methods; (b) original images; (c) skin detection results by Jones's method; (d) skin detection results by our method; (e) refined results

From Table 1 and Table 2, we can see that combination of different shape descriptors generally provide better classification results than using only single shape descriptor. AdaBoost with C4.5 weak classifier using combination of three shape descriptors yields true positive of 89.2% and false positive of 15.3%, which outperforms the other schemes. Another interesting observation is that SVM classifiers with and without AdaBoost make no much difference in case the same shape descriptor is used. A possible explanation is that SVM and AdaBoost are both margin-maximizing techniques. Since margin has been maximized by a technique, it can not be further maximized by the other technique.

Fig2. Some example images after skin detection. Adult images are on the top row and non-adult images are on bottom row. Shapes of skin regions share some visual similarities in adult images. Non-adult images contain irregular skin regions

Table1. Performance comparison of different boosted classifiers using different shape features. ($\times 100\%$)

	D1		D2		D3		D2+D3		D1+D2+D3	
	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP
C1	90.6	42.5	87.1	30.6	84.1	37.7	86.7	29.3	86.9	29.6
C2	84.7	27.9	84.3	22.4	81.6	22.3	87.6	20.9	**89.2**	**15.3**
C3	76.4	28.5	90.5	39.3	85.2	34.7	90.2	29.2	90.0	26.8
C4	84.1	30.7	90.9	34.8	81.7	23.5	86.6	20.5	85.4	14.2

Table2. Performance comparison of different classifiers using different shape features. ($\times 100\%$)

	D1		D2		D3		D2+D3		D1+D2+D3	
	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP
C1	90.3	42.3	86.0	30.7	90.4	53.0	87.0	30.0	86.6	30.0
C2	81.8	24.7	80.8	24.2	77.3	32.1	81.1	22.9	83.6	18.2
C3	76.4	28.5	90.5	39.3	85.2	35.1	89.9	27.9	90.3	27.2
C4	85.8	33.0	89.6	35.4	79.4	26.9	86.6	20.5	83.4	19.3

6. Conclusion

In this paper, we report our investigation on using shape features to detect adult images. We use multi-Bayes skin color classifiers to detect skin regions and use boosted classifiers to decide whether or not the skin regions represent a nude. We demonstrate that using combination of different shape descriptors can enhance the performance of shape classification. We also show that AdaBoost with C4.5 weak classifier can achieve good adult image detection performance.

Acknowledgement

This work has been financed by the National Hi-Tech R&D Program (the 863 Pro-gram) under contract No.2003AA142140.

References

[1] Margaret Fleck, David A.Forsyth, Chris Bregler, “Finding naked people,” ECCV 1996, pp: 593-602.

[2] James Z. Wang, Jia Li, Gio Wiederhold, Oscar Firschein, “System for screening objectionable image” Computer Communications, pp: 1355-1360, 1998.

[3] Michael J. Jones, James M. Rehg, “Statistical Color Models with Applications to Skin Detection,” CVPR,1999

[4] M. K. Hu, “Visual pattern recognition by moment invariants”. IRE Transactions on Information Theory, IT-8 (1962), pp: 179-187.

[5] C.-W. Chong, P. Raveendran, R.Mukundan, “A comparative analysis of algorithms for fast computation of Zernike moments”, Pattern Recognition 36 (2003), pp: 731-742.

[6] Y. Freund, R.E. Schapire. “Experiments with a new boosting algorithm”, Proc. Thirteenth International Conference on Machine Learning, pp: 148-156, 1996.

[7] Wayne Iba, Pat Langley. “Induction of one-level decision trees”, Proc. of the Ninth International Conference on Machine Learning, pp: 233-240

[8] J.Ross Quinlan. “Improved use of continuous attributes in C4.5”, Journal of Artificial Intelligence, 1996, pp: 77-90.

[9] J. Platt, “Fast training of support vector machines using sequential minimal optimization”, Advances in Kernel Methods-Support Vector Learning, pp: 185-208, 1999.

[10] Christopher M. Bishop, “Neural networks for pattern recognition”, Oxford University Press, 1995.