ACCV2002: The 5th Asian Conference on Computer Vision, 23--25 January 2002, Melbourne, Australia

Adult Image Detection Method Base-on Skin Color Model and Support Vector Machine

Lijuan Duan, Guoqin Cui, Wen Gao and Hongming Zhang
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China 100080

Abstract

This paper demonstrates an automatic system to detect adult images based on compute vision and pattern recognition. First, images were filtered by skin color model. Then, the output images are classified by support vector machine. The experiment shows that 87.3% images contain naked people could be filtered accurately.

1. Introduction

With the development of WWW, the dramatically falling cost of data storage and the advancing in coding technology are generating dazzling array of photography, animation, graphics, sound and video. Now, having an appliance computer and Internet access has become as ordinary as having cable television. However, there is currently a large amount of adult images for free download on the web. Accessing objectionable media by children is increasingly a problem that many parents are concerned about.

Filter the adult images is very important for search engines to avoid offensive content on the web. Now, there are some ways to stop naked images arriving on computers, such as blocking unwanted sites, identifying images depicting naked or scantily dressed people. "X-Stop" is a tool to block pornographic sites. It provides the parents with necessary method to safeguard their children who are using Internet. Another way to identify pornographic images is by means of text analysis or computer vision. In order to find naked people, Forsyth and Fleck design software to detect naked people [1]. It begins by analyzing the color and texture of a photograph. When it finds matches for skin colors, it runs an algorithm that looks for cylindrical areas that might be corresponds to an arm or leg. It then seeks other flesh-colored cylinders, positioned at certain angles, which might confirm the presence of limbs.

This paper presents a new adult image detection method. It is based on compute vision and pattern recognition algorithms. Images were filtered by skin color model at first, then, they were classified by support vector machine.

The paper is organized as follows. A brief system overview is given in section2. In section 3, we put forward a skin color model. In section 4 we briefly introduce the support vector machine (SVM) algorithm and its' application in adult image detection. At last, the experiments and the conclusion are given.

2. System Framework

As illustrated in Figure 1, the system contains two parts: skin color detector and support vector machine.

The skin detector can be used as the basis for adult image detector, because it is a fact that there is a strong correlation between images with lager patches of skin and adult or pornographic images. But not all images with enough skin regions are relevant to pornographic, for example a face image. We use images, which are manually classified into adult and non-adult sets to train a support vector machine classifier. The support vector machine classifier outputs a number between 0 and 1 with 1 signifying an adult image.

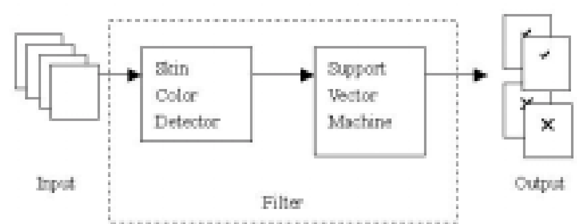


Figure 1. System Framework

3. Skin Color Model

Skin color can be a more powerful cue for detecting people in unconstrained imagery. There have been a number of researchers who have looked at using color information to detect skin. Jones and Rehg [2] constructed color model using histogram-learning techniques at *RGB* color space. It could attain about 85.8% correct detections with about 7.4% false positive. Yang and Auhuja [3] estimated probability density function of human skin color using a finite Gaussian mixture model whose parameters are estimated through the EM algorithm. Forsyth and Fleck [1] employed color and texture combination properties to obtain a mask for skin region. We introduce a new skin detection method based on human perception of colors. First, we label large collection of skin pixels as training examples. Then, we convert the pixels' value from *RGB* to *YUV* and *YIQ* respectively. Finally, we obtain the distributions of skin color pixels.

The *RGB* values are transformed into *YUV* values using the formulation.

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (3\text{-}1)$$

The chromaticity information is encoded in the U and V components. Hue and saturation are gotten by the following transformation.

$$Ch = \sqrt{|U|^2 + |V|^2} \quad (3\text{-}2)$$

$$\theta = \tan^{-1}(|V| / |U|) \quad (3\text{-}3)$$

θ represents hue, which is defined as the angle of vector in *YUV* color space. *Ch* represents saturation, which is defined as the mode of U and V. Proper hue thresholds are obtained according to the observation that the hues of most people's skin vary in the range from 100^0 to 150^0 according to Figure 2.

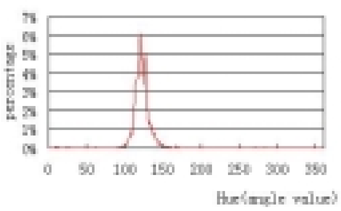


Figure 2. Skin color distribution at *YUV* color space

Like *YUV* color space, *YIQ* is the color primary system adopted by NTSC for color TV broadcasting. Conversion from *RGB* to *YIQ* may be accomplished using the color matrix:

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (3\text{-}4)$$

I is the red-orange axis, Q is roughly orthogonal to I. The less I value means the less blue-green and the more yellow. The flowing Figure 3 shows that most people's skin vary in the range from 20 to 90 in the term of I.

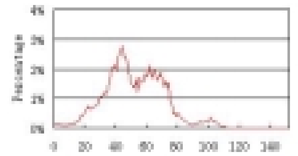
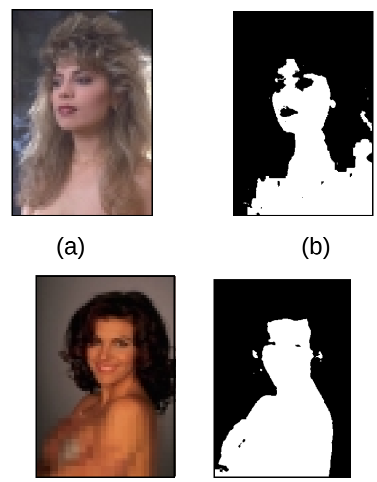


Figure 3. Skin color distribution at *YIQ* space

Through some experiments, we find that the combination of *YUV* and *YIQ* color space is more robust than each other. If a pixel satisfied $I \in [20,90] \cap \theta \in [100,150]$, it is possible be relevant to skin color. In order to find naked people, we also use other features, such as the percentage of pixels detected as skin, the size of the largest skin region. Based on these features, a segmentation procedure is performed on color images. Examples of such segmentation are given in Figure 4.



(a) (b)

(c) (d)

Figure 4. Skin color segmentation

(a)(c) Original images (b)(d) Segment images

Skin color model proposed in the section can effectively detect the skin color region from real images. Through experiments, which performed on 885 adult images and 760 non-adult images, we find that 90.2% adult images and 23.1% non-adult images have skin region which gets to 3,0000 skin color pixels and contains 10% area of each image. 94.4% adult images and 36.7% non-adult images have skin region that gets to 1,5000 skin color pixels and contains 10% area of each image

Skin color detector can find skin region coarsely. But determining which of a large set of pictures contain naked people can be seen as a problem in computer vision and pattern recognition. The challenges of computer vision will most likely remain for a decade or so to come. Searches capable of distinguishing clearly among nudes, marmalades and national flags are still an unrealized dream. Jones [2] used neural network to classify images into adult and non-adult images. Enlighten by Jones, we use support vector machine to detect adult image after skin detection.

4. Support Vector Machine

Support vector machine (SVM) was developed by Vapnik et al, which based on the Structural Risk Minimization (SRM) principle from statistical learning theory [6]. It can be applied to density estimation, classification, and regression problems [13]. Here we focus on the situation of classification. Let the training set be $\{(x_i, y_i)\}$, with each input $x_i \in D \subseteq R^N$ and $y_i \in \{1, -1\}$ is the class label of $x_i, i = 1, 2, \cdots, d$, d is the total number of the training data.

SVM map the training data x_i to a high-dimensional feature space Γ with mapping Φ $D \rightarrow \Gamma$ and separate the two classes (the label is 1 and -1) that are of data with a maximum margin hyperplane. Then the SVM learning algorithm finds a hyperplane (*w*, *b*) that makes:

$$Max \min_j \{w \cdot \Phi(x_j) + b\} \qquad (4\text{-}1)$$

Thus, SVM learns decision function is:

$$f(x) = sign(w \cdot \Phi(x) + b) \qquad (4\text{-}2)$$

where w is a weight vector and b is a threshold, $x \in R^N$ is a random variable .

It is easy to prove [11] that:

$$w = \sum_j \alpha_j y_j \Phi(x_j) \qquad (4\text{-}3)$$

where α_j are non-negative real numbers that maximize

$$\sum_j \alpha_j - \sum_{i,j} \alpha_i \alpha_j y_i y_j \Phi(x_i) \cdot \Phi(x_j) \qquad (4\text{-}4)$$

subject to

$$\sum_j \alpha_j y_j = 0, \alpha_j \geq 0 \qquad (4\text{-}5)$$

Thus, The decision function can be expressed as:

$$f(x) = sign(\sum_j \alpha_j y_j \Phi(x_j) \cdot \Phi(x) + b) \qquad (4\text{-}6)$$

A remarkable property of this equation is that only a subset of the points x_j will be associated with a non-zero α_j that can affect the hyperplane. These points are called *support vectors* and are the points that lie closest to the separating hyperplane. The sparseness of the α vector has several computational and learning theoretic consequences. It is important to note that neither the learning algorithm nor the decision function needs to represent explicitly the image of points in the feature space. To the mapping Φ , since (4-4) and (4-6) use only the dot products $\Phi(x_i) \cdot \Phi(x_j)$. Hence, if one were given a function

$$\mathrm{K}(x, z) = \Phi(x) \cdot \Phi(z) , \qquad (4\text{-}7)$$

one could get the hyper-plane in the feature space without ever explicitly performing the mapping. According to Mercer's condition [11], for each continuous positive definite function $\mathrm{K}(x, z)$ there exists a mapping Φ that makes (4-7), for all $x, z \in D$. The function $\mathrm{K}(x, z)$ is called the *kernel function*. The use of a kernel function allows the support vector machine to operate efficiently in a nonlinear high-dimensional feature spaces without being adversely affected by the dimensionality of that space. Indeed, it is possible to work with feature spaces of infinite dimension.

In our system we explore the use of a radial basis function (also called Gaussian kernel function), the correspondent nonlinear decision surface is

$$f(x) = sign(\sum_{i=1}^{d} y_i \alpha_i \exp(\frac{|x - x_i|^2}{2\sigma^2}) + b) \qquad (4\text{-}8)$$

Figure 5 and Figure 6 show the SVM training and classification process in our system. CCV (color coherence vector) and color histogram are extracted from each image in data preparing step.

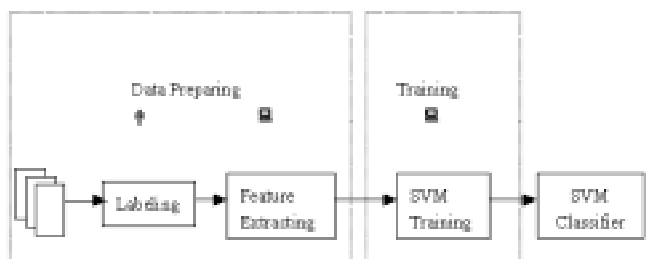


Figure 5. Training

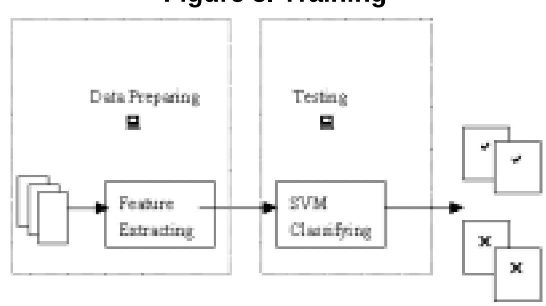


Figure 6. Testing

5. Experiments

We applied the proposed method to the real color images to test its robustness and validity. The system can handle the images at different lighting conditions and different size.

The performance of the system is tested using 312 target images of naked people and 710 assorted control images, containing some images of people but non of naked people. The target images are collected from the Internet and CD. The Most of the people in the images are Asians, Caucasian and European, a small number of Blacks. The control images are collected from the Internet, Corel Gallery, CD-ROM and family photos, which contain 210 images of clothed people, 100 mages of animal, 100 images of plant, 100 images of city, 100 images of landscape and 100 other images.

We fist use the skin color detector to filter. Of the 312 test and 710 control images processed, the skin filter marked 293 test images and 212 control images as containing skin region. Mistake by the skin model occur for several reasons. In some control images, there have desert, animal, wood etc. Figure 7 shows some non-adult images, which are detected having skin color. There are also some adult images that are not detected by the filter, such as a portrait shot or images under special illumination. Because of offensive reason, we don't list these images.

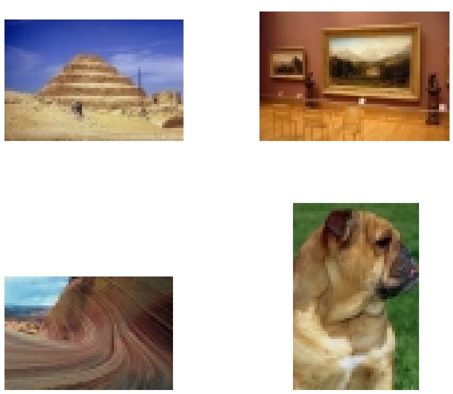


Figure 7. Typical non-adult images wrongly detected as containing skin region

The support vector machine runs on the output of the skin filter: 293 test images and 212 control images. The support vector machine marks 252 test images and 71 control images. 80.7% adult images could be detected by our method and 10% control images are labeled as adult images.

6. Conclusions

In this paper we have present an approach for the naked people detection in color images based-on skin color model and support vector machine. We have developed skin color model at YUV and YIQ color space to detect skin region coarsely. Many control images containing people will be detected. Support vector machine could classify most naked people from non-naked people.

7. References

[1] D. A. Forsyth. M. Fleek, and C. Bregler. Finding naked people. In Proc.Forth European Conference on Computer Vision. pp 593-602. 1996.

[2] Michael J. Jones, James M. Rehg. Statistical Color Models with Application to Skin Detection. CVPR,1999.

[3] Ming-Hsuan Yang and Narendra Ahuja. Gaussian Mixture Model for Human Skin Color and Its Applications in Image and Video Database. SPIE 1999.

[4] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery, Vol.2, pp.121-167. 1998.

[5] V.Vapnik. The nature of statistical learning theory. Springer, New York, 1995.

[6] V.Vapnik, Statistical learning theory. Chichester, GB:Wiley, 1998

[7] F.Riesz and B.Sz.-Nagy. Functional Analyses. (242-246).Ungar, New York, 1955.

[8] A.Snika abd V,Schölkopf. Atutorial on support vector regression.NeuroColt 2 TR 1998-03,1998.

[9] J.C.Platt. Fast training of support vector machines using sequential minimal optimization, Advances in Kernel Methods, MIT press, 271-284,1999

[10] S.K.Shevade, S.S.Keerthi, C.Bhattacharyya & K.R.K. Murthy, Improvements to SMO Algorithm for SVM Regression, Technical Report CD-99-16.

[11] V.N.Vapnik, Statistical Learning Theory. J.Wiley, forthcoming.

[12] Yuan Yaxiang and Sun Wenyu, Optimization Theory and Method (In Chinese), 422-454, Science Press1999.

[13] http://www.clopinet.com/isabelle/Projects/SVM/applist.html