Proceedings of the 6th WSEAS International Conference on Multimedia Systems & Signal Processing, Hangzhou, China, April 16-18, 2006 (pp215-219)

Research on Automatic Classification Technology of Web Image Based on Fuzzy Pattern Recognition

YI XIE, XUN WANG
Institute of computer and information engineering
Zhejiang Gongshang University
Hangzhou, 310035
CHINA

Abstract: -The automatic classification technology of web image is important for control the web content. For handy image classification, it is difficult to satisfy people's demands because of the period, fee and effect. This paper presents a kind of web image automatic classification based on fuzzy pattern recognition, HSV color space and complexion model. The membership function is obtained to extract complexion characteristics firstly, and then the classification determinant rule of sexy image is given. The experiment results show the classification algorithm is effective for severity sexy image and attains anticipated target. This algorithm improved the currency character of classification algorithm because the hypothesis of classification model is a little.

Key-Words: – Web image classification, Fuzzy theory, Complexion model, Pattern recognition, Membership function.

1 Introduction

Web browser based on Internet has provided us with a hugeness global information space in the past tens years. But the environment of Internet is open and easy to lead the disorder of many web contents such as eroticism, violence and others harmful network contents. The search probability about pornography and sexual images is very high existing top of 20[1]. It is obvious that abuse of Internet eroticism is becoming more and more graveness. Some researches show that pornography and sexual images may make people addiction, at the same time, it will bring many severity negatives, especially to young person. The blight or hurt may be brought to people who indulge in pornography and sexual images.

Traditionally, the classification of web image is fulfilled by manual work, namely, after the Internet content is analyzed, an appropriate sort will be given. It is difficult to satisfy people's demands because of period, fee and effect for handy image classification. So it is important to bring forward a kind of effective web image automatic classification technology.

At present, there are many literatures on image classification. Aditya V., Mario A., Anil K.J., etc. have developed the method of extracting the eigenvalue of images for image classification [2].Using binary Bayesian classifiers; they attempt to capture high-level concepts from low-level image features under the constraint that the test image does belong to one of the classes. For the choice of classification algorithm, support vector machines (SVM) have been quite popular for the image classification task. [3,4,]. The SVM algorithm is able to take a large number of features for an image and combine them together to predict the category for that image. CART classifiers are used for the classification of images [5]. The paper presented a kind of image automatic classification based on fuzzy pattern recognition, HSV color space and complexion model. The algorithm is proposed. This algorithm improved the currency character of classification algorithm because the hypothesis of classification model is a little. The experiment results show the classification algorithm is effective for severity sexy image and attains anticipated target.

2 Fuzzy Classification Theory

The concepts, such as old man, taller and high temperature, are uncertain in some degree. So these are called the fuzzy phenomena. In order to analyze fuzzy phenomena with uncertain values using computer image processing, The fuzzy set is used to describe the phenomenon. When X is a limited aggregate if $X = \{x_1, x_2, x_3, ..., x_n\}$,then fuzzy aggregate A is expressed using (1).

$$A = \{(u_A(x_i), x_i) \mid i = 1,2,3,..., n\} \quad (1)$$

When one identify algorithm is used to analyze object x, there are a set of membership function $\sigma_{A1}(x), \sigma_{A2}(x), \Lambda\ , \sigma_{An}(x)$, which express the degree of object x belonging to $A_1, A_2, \Lambda\ , A_n$. When the membership function based on fuzzy model is established, the object x is estimated and indexed to a certain sort according some principle.

Fuzzy pattern recognition has the following three steps [6]:

(1) Selecting the characteristic targets of identified object. The characteristic targets which are in connection with recognition issue remarkable will be selected and the characteristic target values of object x measured, then characteristic vectors $x = (x_1, x_2, \Lambda\ , x_n)$ are calculated.

(2) Constructing the membership function in fuzzy model.

(3) Recognition and estimating. According membership principle, the object x is estimated and indexed to a certain pattern or sort.

3 Web Image Automatic Classification Based On Fuzzy Theory and Complexion Model

3.1 HSV Color System

HSV (hue, saturation, value) system is a color system format brought out by Munseu. In HSV system, Value expresses the intension or brightness. The intension value confirms the whole brightness, and regardless what the color is. The color images can be translated to black and white images through averaging the RGB value, in this way the color information is lost. Two parameters that contain color information are hue and saturation. The color ring describes the two parameters, shown in Fig.1. Hue expresses the kind of color, which is denoted in angle.

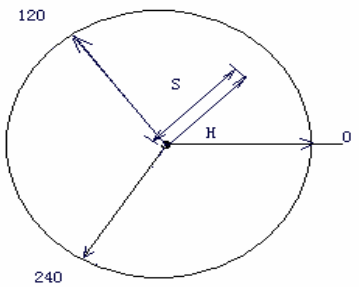


Fig.1 Color ring

In order to extract the complexion character, the RGB color space is converted to HSV color space usually because the HSV color model corresponds more closely to the human perception of color. The conversion equations are as (2), (3), (4) and (5).

$$\theta = COS^{-1}\left[\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right] \quad (2)$$

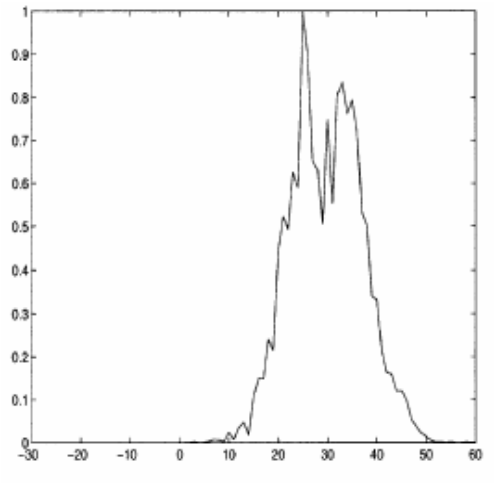
$$H = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G \leq B \end{cases} \quad (3)$$

$$S = \frac{Max(R,G,B) - Min(R,G,B)}{Max(R,G,B)} \quad (4)$$

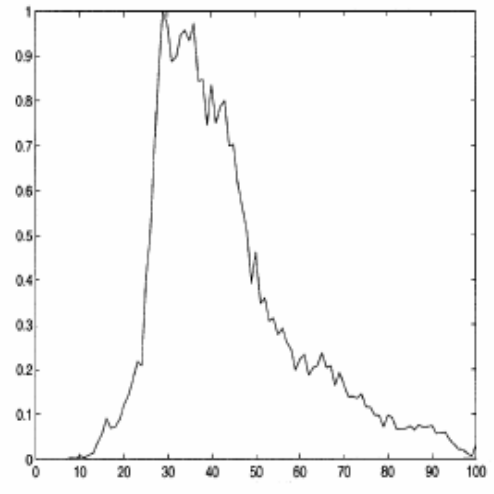
$$V = \frac{Max(R,G,B)}{255} \quad (5)$$

The hue (H) has a range of values between 0 and 360 degrees beginning with red at 0 degree. It gives us a measure of the spectral composition of a color. The saturation (S) is a ratio that ranges from 0 to 1. This component refers to the proportion of pure light of the dominant wavelength and indicates how far a color is from a gray of equal brightness. The value (V) also ranges between 0 and 1 and is a measure of the relative brightness. The normalized histograms [7] obtained form

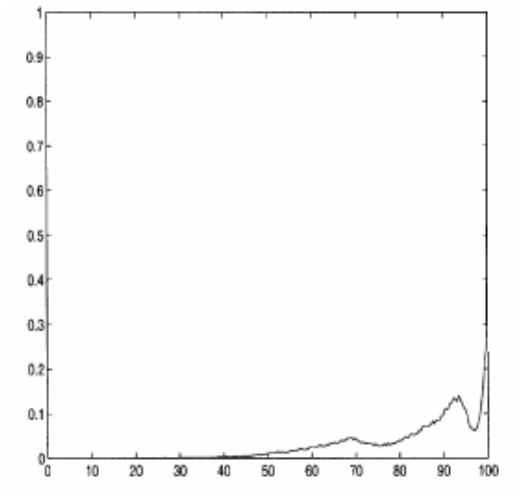
Asian is shown in Fig.2.



(a) Hue (Degrees)



(b) Saturation (%)



(c) Value (%)

Fig.2 Skin color distributions of Asian races in HSV

From Fig. 2, the hue value falls predominantly between 0 degree and 60 degree. The hue has the highest mean value (28.9 degree) and the lowest standard deviation (5.1 degree). So the hue is adopted to model the range of skin color.

3.2 Fuzzy Membership Function

In order to keep the complexity of the overall scheme to a minimum, a trapezoidal shape is selected as the membership function to quantify the degree to which the object fits the corresponding primitive. The general form of the function [7] is defined in (6) and is also shown schematically in Fig. 3.

$$u(x)=\begin{cases}\dfrac{x-c}{a-c} & c\le x\le a\\ 1 & a\le x\le b\\ \dfrac{d-x}{d-b} & b\le x\le d\\ 0 & otherwise\end{cases} \qquad 6$$

The hue characteristics of the skin region are used to form the membership function. The extent of the proposed Asian hue range is from 0° to 50° with the mean 28.9° and standard deviation 5.1°. Thus the membership function is defined as (7):

$$H=\begin{cases}\dfrac{x}{24} & 0\le x\le 24\\ 1 & 24\le x\le 34\\ \dfrac{60-x}{26} & 34\le x\le 60\end{cases} \qquad 7$$

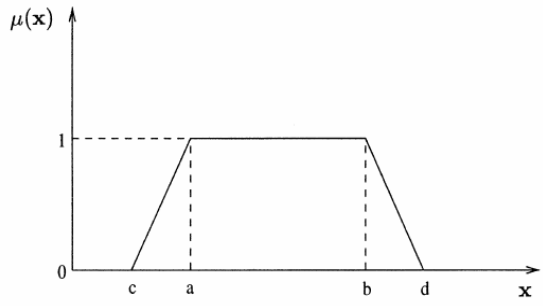


Fig. 3. General trapezoidal membership function.

The membership function can assume any value in the interval [0, 1]. A value of 0 in the definition above indicates that the event is impossible. On the contrary, the maximum membership value of 1 represents total

certainty. The intermediate values are used to quantify variable degrees of uncertainty.

Selecting (7) as the membership function to describe the complexion characteristics, according as the following reasons:

(1) The value of the membership function is between 0 and 1, which accords with the basic character of fuzzy function.
(2) Applying the function can describe the two kinds of fuzzy station of pixel in color image: complexion pixels and non- complexion pixels.
(3) Optimizing process of parameter make this algorithm has the better flexibility, which ensures the purpose of color character selecting.

First, pixel in an image for classification is converted from RGB color space to HSV. Then a certain H can be calculated by (7), and then extended into a closed region [H-2 H+2]. The complexion characteristics can be gotten by this closed region. If the character value accords with classification demands, it is the elementary complexion region, otherwise non- complexion region.

3.3 Classification Determinant Rule of Sexy Image

Through calculating, the elementary complexion region is gained. Because the difference of common clay's complexion is little, the standard deviation of the entire elementary complexion region is calculated. If the value of standard deviation accords with (8), the complexion region has the similar color. Moreover the wrong classification instance is avoided for the complexion difference of different clay is great.

$$\sigma \le 0.03 \qquad (8)$$

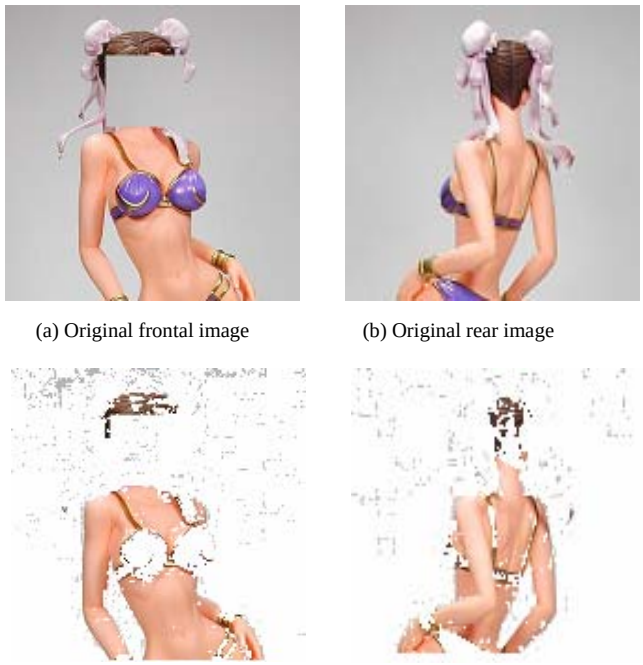
When the common clay's complexion region is gained, we can judge whether the image is sexy image or not. If the complexion proportion accords with (9), the image is judged as sexy image.

$$\frac{pixel\ amount\ of\ the\ sexy\ image}{pixel\ amount\ of\ all\ the\ image} > 5\% \qquad (9)$$

3.4 Classification Recognition Result of Web Image

Applying the above theory and method to classify the web image, draw the following result. In Fig.4, (a) and (b) are the original web images, (c) and (d) are the identified sexy images respective.

During the research on classification technology of the web image, we selected 200 frames of different images in any pose, nude or half- nude, feature. The classification result is shown in Table 1.



(a) Original frontal image (b) Original rear image

(c) Complexion region of frontal image (d)Complexion region of rear image

Fig.4 Recognition result

Table 1: The classification result of web image

Type / Index	The general sexy image	The severity sexy image
The ratio of classification	60	91

4 Results and Discussion

The experiment results show the classification algorithm is effective for severity sexy image and attains anticipated target. However, to general sexy image, because the image is influenced by beam easy, the common clay's complexion is changed in the image. Complexion which is irradiation in excess leans to white, whereas, non- irradiation complexion leans to dark. This deviates from complexion we defined. So the classification is lower. Considering the recognition speed, the timer is about 1 second in this paper. If it is used to analyze the image on-line, acceleration algorithm will be applied.

References

[1] The effects of pornography and sexual messages, National coalition for the protection of child & families, Cicinnati, Ohio, http://www.nationalcoalition.org.

[2] Aditya V., Mario A., Anil K.J., etc., image classification for content-based indexing, IEEE Transaction on image processing, Vol.10, No.1, 2001,pp.117-130.

[3] Olivier C., Patric H., Vladimir N.V., Support vector machines for histogram-based image classification, IEEE Transaction on neural networks, Vol.10, No.5, 1999 , pp.1055-1063.

[4] C. J. C. Burges, A Tutorial on Support Vector Machine for Pattern Recognition, Knowledge Discovery and Data Mining, Vol.2, No.2, 1998,pp.121-167.

[5] R.Schettini, C.Brambilla, G.Ciocca etc., A hierarchical classification strategy for digital documents, Pattern Recognition, Vol.35, 2002,pp.1759-1769.

[6] CHEN Qin, ZHANG Guo2xuan, WANG Xiaohua, Research on text automatic class if ication method based on fuzzy pattern recogn ition. Journal of Zhejiang University (Sciences Edit ion), Vol.27, No.3,2000,pp. 292- 295.

[7] N. Herodotou, K.N. Plataniotis, A.N. Venetsanopoulos, Automatic location and tracking of the facial region in color video sequences, Signal Processing: Image Communication Vol.14, 1999,pp.359-388.

[8] Chen-Te Ho, Rung-Ching Chen, a study of web pages classification based on image and test features, A paper submitted to Chaoyang University of technology (Taiwan) for the degree of Master of Science, 2004.10