Categorizing Traditional Chinese Painting Images

Shuqiang Jiang[1,2] and Tiejun Huang[2]

[1] Digital Media Lab, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
[2] Research Center of Digital Media, Graduate School of of Sciences, Beijing 100039, China
{sqjiang, tjhuang}@jdl.ac.cn

Abstract. Traditional Chinese painting ("Guohua") is the gem of Chinese traditional arts. More and more Guohua images are digitized and exhibited on the Internet. Effectively browsing and retrieving them is an important problem need to be addressed. This paper proposes a method to categorize them into Gongbi and Xieyi schools, which are two basic types of traditional Chinese paintings. A new low-level feature called edge-size histogram is proposed and used to achieve such a high level classification. Autocorrelation texture feature is also used. Our method based on SVM classifier achieves a classification accuracy of over 94% on a 3688 traditional Chinese painting database.

1 Introduction

With the advances of network and computing technology, many organizations have a large digital images content available for online access. Various museums are constructing digital archives of art paintings and preserve the original artifacts. More and more artists attempt to exhibit and sell their productions on the Internet. Thus it is possible to access and appreciate art pieces in digitized format. Effective indexing, browsing and retrieving art images are important problems need to be addressed not only for computer scientists and art communities but also for common art fanners. Guohua dates back to the Neolithic Age, some 6,000 years ago. As an important part of the East Asian cultural heritage, it is highly regarded for its theory, expression, and techniques throughout the world. Guohua is generally classified into two styles: Xieyi (freehand strokes) and Gongbi ("skilled brush"). The Xieyi School is marked by exaggerated forms and freehand brush work. The essence of landscapes, figures and other subjects are rendered with a minimum of expressive ink (Fig.1 (a)(b)). In contrast, the brushwork in Gongbi paintings is fine and visually complex, it is characterized by close attention to detail and fine brushwork (Fig.1 (c)(d)).

In the literature, automatically image understanding and retrieval use content-based method. A variety of techniques and systems have been developed such as QBIC, Photobook, VisualSEEK, and WebSEEK [1]. The drawback of

K. Aizawa, Y. Nakamura, and S. Satoh (Eds.): PCM 2004, LNCS 3331, pp. 1–8, 2004.

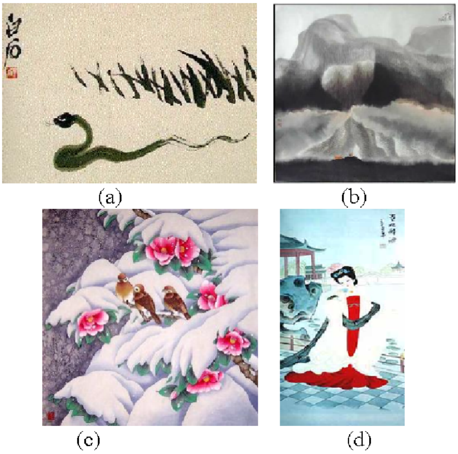


Fig. 1. Examples of Gongbi and Xieyi

this method is that low-level features used by them always could not be interpreted to high-level concepts that are commonly used by human. To overcome the drawback, semantic-sensitive image retrieval techniques have been introduced. Image semantic classification is a form of semantic image understanding. Its goal is to assign the image to semantic class, thus assisting image retrieval and related processing. Authors in [2][3] investigated on classifying indoor and outdoor scenes. Vailaya el al. [4] addressed the problem of classifying city versus landscape and further group landscape images into sunset, mountain and forest. The authors in [5][6] gave approaches to identify natural photographs versus artificial graphs generated by computer tools. Other examples of image semantic classification include face detection and objectionable image identification [7].

Processing on digitized art images is becoming an important research topic. The DELOS-NSF [8] working group discusses problems of retrieving art images and bridging the semantic gap, it points out that this area is still in the early stages of research. Li and Wang [9] use multi-resolution HMM method to characterize different drawing styles of artists. References [10] and [11] give techniques to identify Canvas painting and Traditional Chinese painting images respectively. A.D. Bimbo et al. [12] investigate on the problem of retrieving painting images using color semantics derived from the Itten color sphere.

As described in [8], constructing computer algorithms to learn to classify paintings of different styles is a promising step to bridging the semantic gap. Some works have been done in this domain [9,10,11]. Gongbi and Xieyi are two basic schools of traditional Chinese paintings. To characterize these two types of paintings may be of much help in analyzing traditional Chinese paintings and in retrieving them. In this paper, we propose a method to achieve this goal based on the investigations of visual differences of these two types of Guohua. This has not been explored before. Edge and texture features accompanied with SVM classifiers are used to achieve this high-level classification.

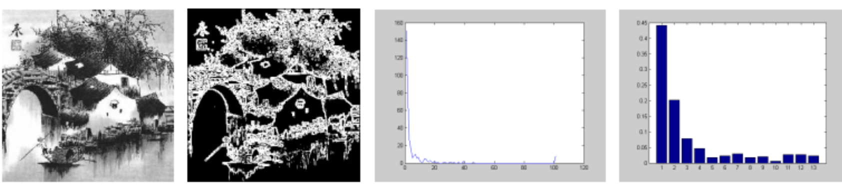
The rest of the paper is as follows. Image features are introduced in section 2 and implementation issues in section 3.Experimental results are given in section 4. And section 5 concludes the paper.

2 Image Features

Two distinguishable low-level texture and color features are employed. Edge-size histogram is first introduced in this paper. It measures the sparseness and granularity of edges of an image. An autocorrelation feature is a traditional texture feature that reflects coarseness of an image.

2.1 Edge-Size Histogram

Edge is regarded as an important feature to represent the content of the image. It conveys a large amount of visual information and human eyes are known to be sensitive to edge features for image perception. Edge histogram descriptor for MPEG-7 is well used in image matching [13]. Shim et al. [14] integrate color histogram and edge histogram for image retrieval. Edge-size histogram introduced in this paper is different from the above two method. It measures consistency and granularity of image edges. We give the formal description below.



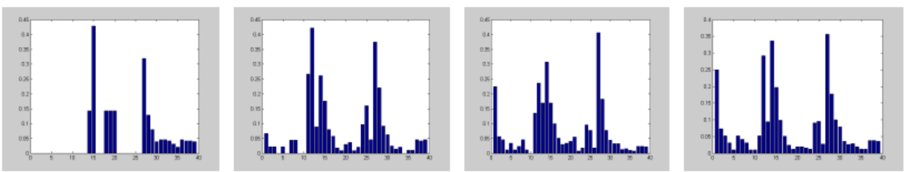
(a) (b) (c) (d)

Fig. 2. Edge size histogram of an example image

Let I be a gray level image, *p(x,y)* be pixels in *I*. Edge detection is first performed using Sobel detector and generating k number of edges:$\{e_1, e_2, \ldots, e_k\}$. Let n_i be the size of the edge e_i: $n_i = |\{p|p \in e_i\}|$, where $|\bullet|$ denote the number of elements.The edge-size histogram is defined to have 13 dimensions. From 1 to 13, each dimension accumulates the number of edges that have the size of:{1, 2,. . . , 10, [11,20], [21,100], [101,∞]},thus generating the vector$[\mathrm{EH}_j]_{13}$. Edge-size histogram is computed by quantization of the above vector:$ESH_j = EH_j/k, j \in [0, 13]$.. To compute this kind of feature, image should first be resized to have same number of pixels. Figure 2 gives an example to compute this feature. Figure 2 (a) is the original gray-level image. 2(b) is the result of edge detection; there are totally 342 edges. Figure 2(c) is the edge size numbers from 1 to 101, the 101^{st} dimension is the numbers that edges have the size larger than 100. Figure 2(d) is the final edge-size histogram.

As we know, color, saturation and luminance are three factors that artists used to create their productions. Thus we use HSL color space to represent images in the implementation. Edge-size histogram is computed on each of the 3 channels and a 39-bin feature is obtained.

It could be observed that Gongbi images generally have more detailed edges than Xieyi images, this is because the former is characterized by simple and bold strokes intended to represent the exaggerated likenesses of the objects, while the latter by fine brushwork and close attention to detail. The following figure shows this difference demonstrated by edge-size images. Figure 3 (a) (b) is the edge-size histogram of Xieyi paintings in figure 1 (a)(b). On the hue channel, small edges is less than that of figure 3(c)(d), which are edge-size histogram of two Gongbi images in Figure 1(c)(d). From data analysis, the other two channels may also be of some help to differentiate these two schools of traditional Chinese paintings.



(a) (b) (c) (d)

Fig. 3. Structure of the System in Fig.1

2.2 Autocorrelation Texture Features

Autocorrelation [15] measures the coarseness of an image by evaluating the linear spatial relationships between texture primitives. Large primitives give rise to coarse texture and small primitives give rise to fine texture. If the primitives are large, the autocorrelation function decreases slowly with increasing distance whereas it decreases rapidly if texture consists of small primitives. Autocorrelation function of an image is described as:

$$C_{ff}\left(p,q\right) = \frac{mn}{(m-p)(n-q)} \frac{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} f(i,j)f(i+p,j+q)}{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} f^2(i,j)}$$

(p,q) varied from (2,2) to (10,10) in a step of two, totally giving of 25 features.

It is conceivable that Gongbi images generally have finer textures than that of Xieyi. While in some cases, the margin part of Gongbi images is rather large, so the center part of the image is segmented to compute autocorrelation features

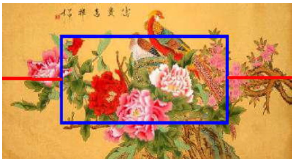


Fig. 4. Center part of an image

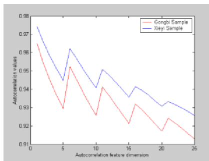


Fig. 5. Average autocorrelation feature values of typical Gongbi and Xieyi images

as illustrated in Figure 4. Xieyi images have larger feature values compared to Gongbi images. Figure 5 shows average result on 30 Gongbi and 30 Xieyi images selected from training set.

3 Implementation Issues

3.1 Classifier

Support vector machine is a two-class classification approach to learn linear or non-linear decision boundaries [16]. Given a set of points, which belong to either of two classes, SVM finds the hyper-plane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyper-plane. This is equivalent to performing structural risk minimization to achieve good generalization. Assuming l examples from two classes,

$(x_1, C_1)(x_2, C_2) \ldots (x_l, C_l)$, $x_i \in R^N$, $C_i \in \{-1,+1\}$

finding the optimal hyper-plane implies solving a constrained optimization problem using quadratic programming. The optimization criterion is the width of the margin between the classes. The discriminate hyper-plane is defined as:

$$g(x) = \sum_{i=1}^{l} a_i C_i k(x, x_i) + a_0$$

where $k(x, x_i)$ is a kernel function, x_i are so-called support vectors determined from the training data, C_i is the class indicator associated with each x_i, and a_i are constants which are also determined by training. Constructing the optimal hyper-plane is equivalent to finding all the nonzero a_i.The sign of g(x) indicates the membership of x.

3.2 Dataset

The image database used in this experiment consists of 3688 traditional Chinese painting images collected from various sources and different artists in different periods. 1799 are Gongbi paintings and 1889 are Xieyi paintings. All these images are used as the test set. The training set includes 117 Gongbi images and 118 Xieyi images.

4 Experimental Results

We conducted a variety of experiments to measure classification performance. Overall classification accuracy criterion is used to evaluate the performance. Let K_g and K_x denote total number of Gongbi and Xieyi images respectively; C_g and $C_x$denote the number of correctly identified Gongbi and Xieyi images respectively. Thus detection precisions are described as follows:

P(G)=Precision (Gongbi) $= C_g \ / \ K_g$
P(X)= Precision (Xieyi) $= C_x \ / \ K_x$
P(O)= Precision (Overall) =(P(G)+ P(X))/2

Table 1. Result of classification method

	ESH	AC	ESH+AC
P(G)	85.2854%	77.2686%	95.5553%
P(A)	79.2523%	78.9172%	93.6644%
P(O)	82.2689%	78.0929%	94.6098%

Table 2. Comparison of different classifiers

	SVM	C4.5 Decision Tree	Naive Bayesian
P(G)	95.5553%	88.1601%	92.1623%
P(A)	93.6644%	83.748%	90.0476%
P(O)	94.6098%	85.9541%	91.105%

Table 1 shows the classification results of our method. The combined features of edge-size histogram (ESH) and autocorrelation (AC) give us better performance than one feature alone. The final overall classification accuracy of 94.6098% is achieved. The authors also compare SVM classifier with decision tree and Bayesian method, the features used here are combined edge-size histogram and autocorrelation. Results showed in table 2 validate that SVM has comparable or better performance among classification methods. Figure 6 shows some correctly classified image, and Figure 7 are some misclassified paintings.

(a) Correctly classified Gongbi paintings

(b) Correctly classified Xieyi paintings

Fig. 6. Correctly classified traditional Chinese paintings

(a) False classified Gongbi paintings

(b) False classified Xieyi paintings

Fig. 7. False classified traditional Chinese paintings

5 Conclusion

Although the two schools of Guohua try to achieve the same end: the creation of beauty, Gongbi and Xieyi images employ different drawing techniques. This paper gives a method to differentiate these two important categories of Chinese paintings using low-level image features. Encouraging results was obtained on a medium-sized dataset. Future works include establishing algorithms to identify other important semantic information for digitized Guohua images and other types of art images.

Acknowledgments. This work was partly supported by China-American Digital Academic Library (CADAL) project, and partly sup-ported by NEC-JDL video analysis and retrieval project.

References

1. Smeulders A.W.M., Worring M., Gupta A., Jain R.,: Content-Based Image Retrieval at the End of the Early Years. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, 2000
2. Szummer M., Picard R.: Indoor-Outdoor Image Classification. IEEE International Workshop on Content-based Access of Image and Video Databases, in conjunction with ICCV'98. Bombay, India, 1998
3. Serrano N., Savakis A., Luo J.: A Computationally Efficient Approach to Indoor/Outdoor Scene Classification. IEEE ICPR'02, Québec City, Canada, Aug. 2002
4. Vailaya A., Jain A., Zhang H.J.: On Image Classification:City vs. Landscape. IEEE Workshop on Content-Based Access of Image and Video Libraries, Santa Barbara, California, June21, 1998
5. Prabhakar S., Cheng H., Handley John C., Fan Z.G., Lin Y.W.: Picture-Graphics Color Image Classification. IEEE ICIP 2002
6. Wang J.Z., Li J., Wiederhold G., Firschein O.: Systems for Screening Objectionable Images. Computer Comm., vol.21, no.15, 1998
7. Fleck M., Forsyth D.A., Bregler C.: Finding Naked People. Proc. European Conf. Computer Vision, vol.2, 1996
8. Chen C.C., Bimbo A., Amato G., Boujemaa N.: Report of the DELOS-NSF Working Group on Digital Imagery for Significant Cultural and Historical Materials. DELOS-NSF Reports, December, 2002
9. Li J., Wang J.Z,: Studying Digital Imagery of Ancient Paintings by Mixtures of Stochastic Models. IEEE Transactions on Image Processing, vol. 12, no. 2, 15 pp., 2004
10. Leykin A., Cutzu F., Hammoud H.: Visual Properties Differentiating Art from Real Scenes. Technical Report No. 565, Computer Science Department, Indiana University
11. Jiang Shuqiang, Gao Wen, Wang Weiqiang: Classifying Traditional Chinese Painting Images. The 4th International Conference on Information, Communications and Signal Proc-essing - 4th IEEE Pacific-Rim Conference On Multimedia (ICICS-PCM2003)
12. Colombo C., Bimbo A.D., Pala P.: Semantics in Visual Information Retrieval. IEEE Multimedia, July-September 1999
13. International Organization for Standardization Organization International De Normalization ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio, ISO/IEC JTC1/SC29/WG11 N4031, Singapore, March 2001
14. Shim Seong-O, Choi Tae-Sun: Edge Color Histogram for Image Retrieval. IEEE ICIP 2002
15. Tuceryan M., Jain A.K.: Texture Analysis. In The Handbook of Pattern Recognition and Computer Vision (2nd Edition), by C. H. Chen, L. F. Pau, P. S. P. Wang (eds.), pp. 207-248, World Scientific Publishing Co., 1998. (Book Chapter)
16. Burges B.: Tutorial on Support Vector Machine for Pattern Recognition. Data Mining and Knowledge Discovery, vol.2, no.2, 1998