Content-Based Image Retrieval Using Texture Structure Histogram

Gang Hou , Qinghe Feng, Xiaoxue Zhang, Jun Kong and Ming Zhang

Abstract. In this paper, we proposed a novel image feature descriptor, namely texture structure histogram (TSH) for content-based image retrieval. This method using the color and edge orientation information to describe the image texture structure information. Considering the HSV color space conforms to humans' visual perception mechanism, the feature extraction is conducted in the HSV color space. This paper puts forward the non-equal interval quantization scheme that makes the expression of the image information to be more reasonable. In feature representation phase, we use the feature fusion mechanism that makes the color and shape information merge together and get a better results. The experiment results demonstrate that the proposed method more efficient and have a high retrieval performance.

Keywords: image retrieval · texture structure histogram · feature representation

Gang Hou

College of Communication Engineering of Jilin University,

College of Humanities & Sciences of Northeast Normal University.

Qinghe Feng, Xiaoxue Zhang, Jun Kong

School of Computer Science and Information Technology, Northeast Normal University.

Ming Zhang (⬚): corresponding author.

School of Computer Science and Information Technology, Northeast Normal University,

Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University.

e-mail: zhangm545@nenu.edu.cn

1 Introduction

In recent years, with the rapid development of digital image processing technology, helping the user to find the multimedia information what they need quickly and effectively becomes a hot research topic at present. Image retrieval is a major component of multimedia information retrieval technology, and also one of the basic theory of video information retrieval, it play a significant role in the field of information retrieval. Image retrieval is based on users' query requests, extract an image or image set that related to the query image from the image dataset. Generally, three categories of methods for image retrieval are used: text-based, content-based and semantic-based [1]. The content-based image retrieval (CBIR) has been proposed in the early 1990's [2]. This approach is to retrieve images using low-level features like color, texture and shape that can represent an image.

Color is a very important visual cue for image retrieval and object recognition. Color histograms are invariant to orientation and scale, and this feature makes it more powerful in image classification. Color histogram-based image retrieval is easy to implement and widely used in CBIR systems. Some of the commonly used color descriptors including compact color moments, the color coherence vector, and color correlo-gram[3].

Texture is one of the most important characteristics of an image. Texture features are also widely used in CBIR systems. Various algorithms have been designed for texture analysis, such as gray level co-occurrence matrices [4], the Tamura texture feature [5], the Markov random field model [6], Gabor filtering [7], and local binary patterns [8]. Tamura et al., based on human visual psychology research put forward some different methods to describe the texture feature, give a description of several different terms: coarseness contrast and directionality, linelikeness, regularity, roughness, etc.

In addition to color and texture features, shape feature is the most essential characteristics of depicting objects. But it is also most difficult to describe. The major difficulty is the segmentation of interested target. Currently commonly used to describe the shape of image retrieval methods mainly include two categories: based on edge and based on the shape of the area. The former using the edge information of images, while the latter using the area of gray level distribution information. Classical methods of describing shape features include the use of moment invariants, Fourier transforms coefficients, edge curvature and arc length [9,10].

This paper uses the color and edge orientation feature that describe the texture information correctly. The rest of this paper is organized as follows: In section 2, we describe the HSV color space and the color and edge orientation quantization scheme in the HSV color space. In section 3, we describe the feature extract and feature fusion. In section 4 the experiments and results were introduced. Finally is the conclusion of this paper.

2 HSV Color Space and Quantization

Color information is the bottom and intuitive physical characteristics. Because color is robust to the effects of noise, size and orientation of image, so color feature is most commonly used in content-based image retrieval. Color quantization is closely related to the color space. A lot kinds of color spaces have been proposed and used for image retrieval. However, different color space has different application, we usually hard to decide which kind of color space is most suitable for our image retrieval algorithm. The HSV color space could mimic human color perception well. The HSV color space can be modeled as a cylinder as shown in [2].

2.1 Color Non-Equal Interval Quantization in HSV Color Space

In order to cut down the computing complexity and extract the color features in efficient way, we use HSV color space and quantize it into non-equal interval 72 bins, thus we get the color index image *C(x, y)*. As is known to all, quantizing the *H, S* and *V* channels uniformly is not suitable for human's visual perception and recognition. We give our non-equal interval quantization scheme as follows. Fig.1 shows that our quantization scheme is better than the equal interval scheme.

$$H = \begin{cases} 0, & H \in [0,25] \, or \, (335,360] \\ 1, & H \in (25,37] \\ 2, & H \in (37,70] \\ 3, & H \in (70,160] \\ 4, & H \in (160,200] \\ 5, & H \in (200,275] \\ 6, & H \in (275,295] \\ 7, & H \in (295,335] \end{cases}$$

$$S = \begin{cases} 0, & S \in [0,0.37] \\ 1, & S \in (0.37,0.76] \\ 2, & S \in (0.76,1] \end{cases}$$

$$V = \begin{cases} 0, & V \in [0,0.37] \\ 1, & V \in (0.37,0.76] \\ 2, & V \in (0.76,1] \end{cases}$$
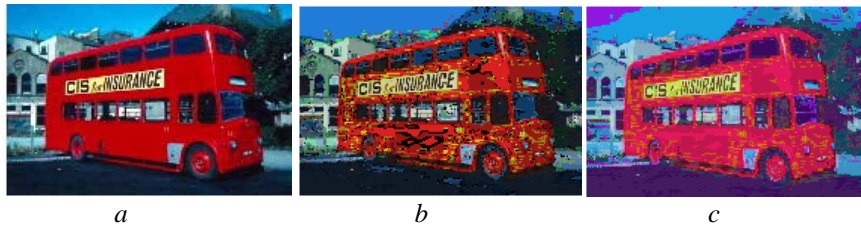

| *a* | *b* | *c* |

Fig. 1 HSV quantization scheme
a. RGB image   *b*. equal interval quantization   *c*. non-equal interval quantization

2.2 Edge Orientation Quantization in HSV Color Space

Edge orientation also plays an important role in content-based image retrieval. Edge orientation information can represent the object's shape and boundaries perfectly. In this paper, we use a simple and effective method for edge orientation detection.

Let $a = (H_x, S_x, V_x)$ and $b = (H_y, S_y, V_y)$, where H_x, S_x and V_x denotes the gradient in H, S and V channel along horizontal direction respectively, and H_y, S_y and V_y denotes the gradient in H, S and V channel along vertical direction respectively. Here we use the Sobel operator introduced in [2].

After the edge orientation $\theta(x, y)$ of each pixel has been computed, the orientations are uniformly quantized into 18 bins with each corresponding to angle of 20°. It is most reasonable according to the experiments. Thus we get the edge orientation index image $\theta(x, y)$.

3 Feature Extraction

3.1 Structure Map Construction and Feature Representation

According to the color index image, we build the color structure map. Similarly, we can build the edge orientation structure map and we omit it here. The details of the process are illustrated in Fig.2. For the color index image, a 3×3 block filter throughout the image from left to right and top to bottom with 3-step length. Then, we consider the value of central point of the 3×3 block and the values of its four neighborhoods. If the value of central point equals to that of its four neighborhoods, then the values of the 3×3 block will be remained, otherwise, the values will be set to zero. Finally, we get the color structure map $T(x, y)$. Similarly, the edge orientation structure map $O(x,y)$. Fig.2 shows the processing of the color information detection.

After previous operation, we can obtain the color structure map $T(x, y)$ and its index value is denoted as $p \in \{0, 1, ..., P-1\}$. For each point (x_0, y_0) in $T(x, y)$, we denote its index value $T(x_0, y_0) = p_0$, and the index values of are $T(x_i, y_i) = p_i, (i = 1,2,3,4)$. The values of edge orientation structure image $O(x,y)$ are denoted as $q \in \{0, 1, ..., Q-1\}$. Similarly, the angle at (x_0, y_0) and its four neighbourhoods (x_i, y_i) are denoted as $O(x_0, y_0) = q_0$ and $O(x_i, y_i) = q_i, (i = 1,2,3,4)$, respectively. In our experiments, P=71, Q=17. Finally we define the color feature and edge orientation feature as follows:

$$H_{color}(T(x,y)) = \sum_{i=1}^{N(T(x,y))} N_i(T(x,y)) \qquad (3.1)$$

$$H_{ori}(O(x,y)) = \sum_{i=1}^{N(O(x,y))} N_i(O(x,y)) \qquad (3.2)$$

where $N(\cdot)$ denotes the total number of the index persisted in color (edge orientation) structure map for each variable p and q, $Ni(\cdot)$ denotes the number of the points whose value is the same as center point among 4-neighborhood for i-th point in color (edge orientation) structure maps. Here $H_{color}(T(x,y))$ is a 72 dimensional vector and $H_{ori}(O(x,y))$ is a 18 dimensional vector. Finally our methods use a 90 dimensional feature vector for image retrieval.
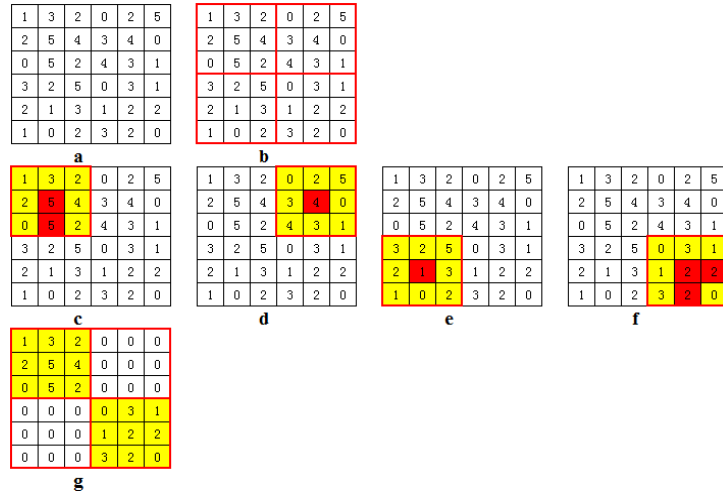

Fig. 2 The process of color structure map constructing

3. 2 Feature Fusion

For some images, its texture information is very rich. However, for some kinds of images, its edge orientation information seems to be much richer. Thus we give different weights to the two features according to the attribute of images. According to the experiment, we train the parameters α to get the most suitable one whose precision is the highest. We combine $H_{color}(T(x,y))$ and $H_{ori}(O(x,y))$ as the final feature vector H, namely, texture structure histogram (TSH, for short).

$$H = [\alpha\, H_{color}(T(x,y)),\ (1-\alpha)\, H_{ori}(O(x,y))] \qquad (3.3)$$

4 Experiments and Results

In our experiments, we use the Corel_5000 dataset and Corel_ 10000 dataset. The Corel_5000 dataset is a subset of the Corel_10000 dataset. We randomly choose 50 categories in Corel_10000 dataset and each category contains 100 images with the size of 192×128 or 128×192 in JPEG format. The second dataset is the Corel_10000 dataset, it contains 100 categories.

In order to match the images, the similarity distance metric between T and Q is calculated using

$$D(T,Q) = \sum_{i=1}^{K} \left| \frac{T_i - Q_i}{1 + T_i + Q_i} \right| \quad (4.1)$$

For the proposed method, K=90 for all the images. A smaller $D(T, Q)$ means more similar to the query image.

In our experiments, we use the precision and recall curves that commonly used in content based image retrieval. The precision and recall is defined as follows:

$$P(N) = I_N / N \qquad R(N) = I_N / M \quad (4.2)$$

The precision of P is defined as ratio between the number of the retrieved relevant images I_N and the total number of the retrieved images N; it measures the accuracy of the retrieval. Recall R is defined as ratio between the number of the retrieved relevant images I_N and the total number of the relevant images M of the whole dataset; it measures the robustness of the retrieval. In our image retrieval method, N=12 and M=100.

Table 1 displays the experiments results compared with the MTH[11] and MSD[12],the precision and recall shows that our method is better than this two methods. In the Corel_5000 dataset, our method performs about 8 percent higher than that of the MTH, and about 4 percent higher than that of the MSD.

Fig. 3 shows an example of image retrieval results with the Corel_10000 datasets; Fig. 4 shows the average precision ratio and recall ratio curves on two datasets. The horizontal axis corresponds to recall, whereas the vertical axis corresponds to the precision. If the curves go far from the original coordinate, it demonstrates the curves has a higher precision and recall. In Fig. 4, it shows that our method performs better than the MTH and MSD, respectively.

Table 1 The average retrieval precision and recall results on two datasets

Dataset	Performance	Method		
		MTH	MSD	TSH
Corel_5000	Precision(%)	52.12	55.92	60.47
	Recall(%)	6.25	6.71	7.26
Corel_10000	Precision(%)	39.1	45.62	47.24
	Recall(%)	4.69	5.48	5.67

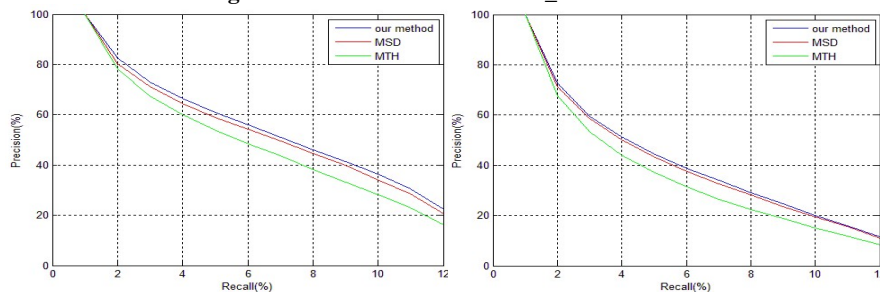**Fig.3** The retrieval results on Corel_10000 dataset

Fig. 4 The precision and recall ratio curves of three methods on two datasets, respectively

5 Conclusion

In this paper, we propose a novel image feature representation method, namely, texture structure histogram (TSH). It is an effective image feature for image retrieval. First, the HSV color space is used, because HSV color space conforms to people's subjective judgment of color similarity, that is to say, HSV color space conforms to human visual perception. Then, we propose a novel non-equal interval quantization scheme that according to the different attribute and distribution of the *H, S* and *V* channels. Building texture structure map is also a key step, which provides a good feature representation for the texture information of image. TSH integrates the advantages of both color and texture features, and shows a good retrieval performance. The experimental results demonstrate that the proposed algorithm have a better performance for image retrieval.

References

1. Liu G H, Yang J Y. Content-based image retrieval using color difference histogram[J]. Pattern Recognition, 2013, 46(1): 188-198.

2. R.C. Gonzalez, R.E. Woods, Digital Image Processing, third ed. , Prentice Hall, 2007.

3. J. Huang, S.R. Kumar, M. Mitra, et al., Image indexing using color correlograms, in: IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 762–768.

4. R.M. Haralick, K. Shangmugam, I. Dinstein, Textural feature for image classification, IEEE Transactions on Systems, Man and Cybernetics SMC-3 (6) (1973) 610–621.

5. H. Tamura, S. Mori, T. Yamawaki, Texture features corresponding to visual perception, IEEE Transactions on Systems, Man, and Cybernetics 8 (6) (1978) 460–473.

6. G. Cross, A. Jain, Markov random field texture models, IEEE Transactions on Pattern Analysis and Machine Intelligence 5 (1) (1983) 25–39.

7. B.S. Manjunathi, W.Y. Ma, Texture features for browsing and retrieval of image data, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8) (1996) 837–842.

8. T. Ojala, M. Pietikanen, T. Maenpaa, Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) (2002) 971–987

9. W. Burger, M.J. Burge, Principles of Digital image processing: Core Algorithms, Springer, 2009.

10. R.C. Gonzalez, R.E. Woods, Digital Image Processing, 3rd edition, Prentice-Hall, 2007.

11. Liu G H, Zhang L, Hou Y K, et al. Image retrieval based on multi-texton histogram[J]. Pattern Recognition, 2010, 43(7): 2380-2389.

12. Liu G H, Li Z Y, Zhang L, et al. Image retrieval based on micro-structure descriptor[J]. Pattern Recognition, 2011, 44(9): 2123-2133.