An Orientation Independent Texture Descriptor for Image Retrieval

Chao-Bing Huang,Quan Liu
School of Information Engineering
Wuhan University of Technology
Wuhan, Hubei, 430070, China
E-mail: huang_cb@163.com, qliu@public.wh.hb.cn

***Abstract*—In this paper, a new orientation independent texture descriptor based on the relation characteristic of pixels in their local neighbor is proposed, which is termed Quantized Compound Change Histogram. For every pixel in the image, the compound rate of change of gray value of the pixel in four different directions in its neighbor is computed and is quantized into 40 bins to get the quantized compound change at the pixel. Calculating the distribution of the quantized compound change over the whole image, the Quantized Compound Change Histogram is obtained. It is invariant to image rotation and translation. Experimental results show that the texture descriptor has better descriptive power for natural color images.**

K
h
C

I. Introduction

Multimedia information is getting more abundant and the means to produce it are becoming a commodity, but finding and managing multimedia content is getting harder and harder. How to efficiently and effectively describe and retrieve multimedia information is a growing need in many application domains. In the early 1990s, Content-based image retrieval (CBIR) was proposed, in which images would be indexed by their visual content, such as color and texture. Since then, many techniques in this research direction have been developed and many image retrieval system, both research and commercial, have been built.

In CBIR, the accuracy of content-based image retrieval depends greatly on the description of low-level visual features. Because of perception subjectivity, there does not exist a single best presentation for a given feature. What features and representations should be used in image retrieval is application dependent. There is a need of developing an image content description model to organize the features. MPEG-7 [1-2] sets a standard for multimedia description tools, it is generic in the sense that it is not specially designed or optimized for a particular application domain. Researchers all over the world can design and develop multimedia feature descriptors to describe and retrieve multimedia information in their application domains.

In recent ten years, many methods [3-6] have been proposed to describe the color-spatial feature of image. In [3], Color Correlogram is used to express how the spatial correlation of pairs of colors changes with distance. In [4], Annular Color Histogram (ACH) is used to express the distribution of each identical color bin in concentric circles centered at the centroid of the bin with different radiuses. In [5], Spatial-Chromatic Histogram (SCH) describes how pixels of identical color are distributed in the image. The spatial distribution is expressed by the standard deviation of identical color bin from its centroid. In [6], Geographical Statistics method describes the spatial distribution of identical color with one parameter of “Looseness”, which is size invariant. These methods are efficient for natural color image.

Texture is an important visual feature of image. Texture is usually defined as a certain local feature, or it measures the relation between pixels in a local area. MPEG-7 [2] considered three texture descriptors, i.e. Texture Browsing Descriptor, Homogeneous Texture Descriptor, and Edge Histogram Descriptor. Homogeneous Texture Descriptor is computed by first filtering the image with a bank of orientation and scale sensitive filters, and computing the mean and standard deviation of the filtered outputs in the frequency domain. Manjunath et al. [7] proposed the use of Gabor wavelet for extracting the texture descriptor and showed that this descriptor is robust, effective. However, the computation for the extraction of Gabor texture features is intensive and the time consumption is long. In [8], a texture descriptor, texture spectrum, is proposed. The local texture information for a given pixel and its 3×3 neighbor is represented by the corresponding texture unit, and the global texture of an image is characterized by its texture spectrum. In our previous work [10-11], two texture descriptors are proposed. The local texture information is represented by the salient changes between a given pixel and its neighbor in several directions, and the global texture of an image is characterized by the statistics of the local texture information. The texture descriptors are effective for natural color images and the computation of extracting the descriptor is simple. However, the descriptor is orientation dependent. In this paper, a texture descriptor is proposed, which is orientation independent.

The remainder of the paper is organized as follows. Section 2 describes texture descriptor proposed in our paper. Section 3 describes image retrieval. Section 4 reports the experimental results and gives some discussion. Section 5 concludes the paper.

II. A NEW TEXTURE DESCRIPTOR

One of the most popular signal processing based approaches for texture extraction has been the use of Gabor filters. It has been proposed that Gabor filters can be used to model the responses of the human visual system. In [7], the feature is computed by first filtering the image with a bank of orientation and scale sensitive filters and then computing the mean and standard deviation of the output in the frequency domain.

In our previous work [10], a texture descriptor termed Gradient Unit Histogram is proposed, in which the gradient unit of a pixel is used to describe the characteristic of the salient changes of the pixel in its neighbor in four different directions. The descriptor is effective for natural color images, but it is orientation dependent. Here, we give another texture descriptor, Quantized Compound Change Histogram (QCCH), which is orientation independent.

In this paper, Let us define $N_r(i, j)$ as a square neighbor of radius r centered at (i, j). In Fig. 1, $N_2(i, j)$, a square neighbor of radius 2 centered at (i, j) is shown. The part with green is $N_1(i-1, j)$, a square neighbor of radius 1 centered at (i-1, j).

(i-2,j-2)	(i-1,j-2)	(i,j-2)	(i+1,j-2)	(i+2,j-2)
(i-2,j-1)	(i-1,j-1)	(i,j-1)	(i+1,j-1)	(i+2,j-1)
(i-2,j)	(i-1,j)	(i,j)	(i+1,j)	(i+2,j)
(i-2,j+1)	(i-1,j+1)	(i,j+1)	(i+1,j+1)	(i+2,j+1)
(i-2,j+2)	(i-1,j+2)	(i,j+2)	(i+1,j+2)	(i+2,j+2)

Figure 1. $N_2(i,j)$--a square neighbor of radius 2 centered at (i, j).

Let y(i, j) denote the gray value of a pixel at (i, j), and can be calculated according to the following formula: y=0.299R+0.587G+0.114B (R, G, B are red, green, blue component value of color image, respectively). The average gray value over $N_r(i, j)$ is denoted by $y_r(i, j)$. Then, the expression:

$$|y_r(i-r, j) - y_r(i+r, j)| \equiv H_r^y(i, j) \quad (1)$$

is a measure of the horizontal rate of change of average gray value at (i, j). Similarly,

$$|y_r(i, j-r) - y_r(i, j+r)| \equiv V_r^y(i, j) \quad (2)$$

is a measure of the vertical rate of change of average gray value at (i, j).

$$|y_r(i-r, j-r) - y_r(i+r, j+r)| \equiv D_r^y(i, j) \quad (3)$$

is a measure of the diagonal rate of change of average gray value at (i, j).

$$|y_r(i+r, j-r) - y_r(i-r, j+r)| \equiv A_r^y(i, j) \quad (4)$$

is a measure of the anti-diagonal rate of change of average gray value at (i, j).

Here, we set radius r=1, and, $H_1^y(i, j)$, $V_1^y(i, j)$, $D_1^y(i, j)$, $A_1^y(i, j)$ can be computed according to equation (1), (2), (3), (4), respectively. $H_1^y(i, j)$, $V_1^y(i, j)$, $D_1^y(i, j)$, $A_1^y(i, j)$ describes the rate of change of gray value of the pixel at (i, j) in four different directions in its 5×5 neighbor, respectively.

To describe the compound rate of change of gray value of the pixel at (i, j) in four different directions in its 5×5 neighbor, the average value of $H_1^y(i, j)$, $V_1^y(i, j)$, $D_1^y(i, j)$, $A_1^y(i, j)$ are computed. i.e.:

$$v(i, j) = (H_1^y(i, j) + V_1^y(i, j) + D_1^y(i, j) + A_1^y(i, j)) / 4 \quad (5)$$

It can be seen: $0 \leqslant H_1^y(i, j) \leqslant 255$.

$0 \leqslant V_1^y(i, j) \leqslant 255$.

$0 \leqslant D_1^y(i, j) \leqslant 255$.

$0 \leqslant A_1^y(i, j) \leqslant 255$.

So, it can be obtained: $0 \leqslant v(i, j) \leqslant 255$.

In our experiment, it is found that the value of v(i, j) is not large, and it almost concentrate on the range [0, 86]. To represent effectively the compound rate of change of gray value of the pixel at (i, j), a non-uniform quantization is chosen to quantize v(i, j) into 40 integers. The quantized result is denoted by t(i, j), which is the compound change of pixel at (i, j), and, $t(i, j) \in T$, T = { 0, 1, 2, ……, 39 }.

The quantization process is following:

If $v(i, j) \in [0, 15.5]$, then they are uniformly quantized into 16 bins, i.e., t(i, j)=0, 1, …, 15.

If $v(i, j) \in (15.5, 35.5]$, then they are uniformly quantized into 10 bins, i.e., t(i, j)=16, 17, …, 25.

If $v(i, j) \in (35.5, 85.5]$, then they are uniformly quantized into 10 bins, i.e., t(i, j)=26, 1, …, 35.

If $v(i, j) \in (85.5, 115.5]$, then they are uniformly quantized into 3 bins, i.e., t(i, j)=36, 37, 38.

If $v(i, j) \in (115.5, 255]$, then they are quantized into 1 bin, i.e., t(i, j)=39.

For every pixel at location (i, j) (i=0,1,⋯,m, j=0,1,⋯,n), its quantized compound change t(i, j) can be computed. The value range of t(i, j) is T. Thus calculating the distribution of the quantized compound change in the value range, the normalized quantized compound change histogram is obtained:

$$H_t(t) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\delta(t(i,j)-t)\ ,\ \forall t \in T \qquad (6)$$

δ (x) is the unitary impulse function, it has two values, 0 or 1; if x=0, then δ (x)=1, otherwise δ (x)=0.

III. IMAGE RETRIEVAL

The image retrieval problem is the following: let D be an image database and M be the query image. Obtain a permutation of the images in D based on M, i.e., assign rank (I) $\in\{1,2, \ldots, |D|\}$ for each I$\in$D, using some notion of similarity to M. This problem is usually solved by sorting the images I$\in$D, inversely according to |f (I)-f (M)|, where f (*) is a function computing features vectors of images and |*| is some distance measure defined on feature vectors.

A generic description of visual content should be based on a flexible combination of different feature descriptors. Even though there are some examples, where one single descriptor would be sufficient to characterize similarity of images, in most cases a more distinguishable specification of the query will be necessary, which implies that a combination of different descriptor types is necessary within a description scheme. Then, during search for a visual item, an appropriate weighting of the similarity resulting from different descriptors has to be performed according to the query requirements. In this paper, the texture feature is combined with color feature, and an appropriate weighting of the similarity resulting from them is assigned.

A. Color descriptor

The HSV (hue, saturation, value) color space is used. After the conversion from RGB to HSV, the obtained values *h* ,*s* ,and *v* are the hue, saturation, and bright component values, and $h\in[0,2\pi]$, $s\in[0,1]$, $v\in[0,1]$. A non-uniform quantization is chosen in HSV color space to divide 3-dimensional HSV color space into 166 color bins according to [9], which corresponds to human color perception. The quantized color set is denoted by Q, i.e. Q = { 0, 1, 2, …, 165 }.

Given a color image *f*, of size m by n pixels, characterized by the quantized color q at location (i, j), i.e. q = *f* (i, j), the value range of f(i, j) is Q={0, 1, 2, …, 165}. Thus, calculating the distribution of the quantized color in the value range over whole image, the normalized quantized color histogram can be gotten:

$$H_c(q) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\delta(f(i,j)-q)\ ,\ \forall q \in Q \quad (7)$$

B. Distance measure

L_1-distance metric is used. Let M be the query image, I be an image in the image database. For image M and I, their color histograms are $H_c^M(q)$, $H_c^I(q)$, the distance metric between them is d_c(M,I); their quantized compound change histograms are $H_t^M(q)$, $H_t^I(q)$, the distance metric between them is d_t(M,I).

$$d_c(M,I)=\sum_{q\in Q}|H_c^M(q)-H_c^I(q)|$$

$$d_t(M,I)=\sum_{t\in T}|H_t^M(t)-H_t^I(t)|$$

An overall distance measure between two image M and I can then be defined as a linearly combination of its two components:

$d(M, I) = w_1 d_c(M, I) + w_2 d_t(M, I)$

where w_i (i = 1, 2) are weights assigned to them.

The distance sequence is normalized using Gaussian normalization by assuming each distance sequence to be a Gaussian sequence.

For the distance metric of color histograms d_c, suppose there are a distance sequence of $d_c=\{d_{c1}, d_{c2}, \ldots\ldots, d_{ck}\}$, where k is the total number of images in the database. We first compute the mean and standard deviation of the sequence, then normalize the original sequence by the following equation:

$$d_{ci}=\left(\frac{(d_{ci}-m)/(3\sigma)+1}{2}\right)$$

where m is the mean and σ is the standard deviation.

The same normalization process is applied to the distance metric of quantized compound change histograms d_t.

After normalization, the probability of a distance value being in the range of [0, 1] is approximately 99%. Any value outside of the range could be mapped into value 0 or 1. The normalization process ensures the equal emphasis of the two distance measures in the computation of the overall distance.

The choice for the weights w_1, w_2 is dependent on the class of images we are working with, and could be assigned subjectively or determined experimentally using a set of training images. The weights allow us to assign different degree of importance to the two components..

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In our experiments, an image retrieval system termed NFIR is implemented with Visual C++ 6.0 programming language in Windows XP in our Personal Computer with Intel Pentium D CPU 2.80GH_Z and 512MB RAM.

The images are obtained from three image database [12-14], which are 24-bit JPEG color images. The size of images is 384×256 or 256×384. 800 images are selected as our image database. The images are divided into ten categories of different scenes (“bus”, “dragon”, “elephant”, “flower”, “horse”, “beach”, “flag”, “aircraft”, “fungi”, “persons”). There are 80 images in each category. The ten sample images got from ten image categories are showed in Fig. 2, they are “bus” with 384×256, “elephant” with 256×384, “persons” with 256×384, “flower” with 384×256, “horse” with 384×256, “dragon” with 384×256, “beach” with 384×256, “flag” with 384×256, “aircraft” with 384×256, “fungi” with 384×256, respectively.

Figure 2. The ten sample images got from ten image categories.

Three methods are used to compare retrieval performance.

The first one is SCH [5]. The distance measure is also the one used by [5].

The second one is combining color histogram $H_c(q)$ with Gabor wavelet texture descriptor [7], which is denoted by CGabor. In Gabor wavelet texture feature, the total number of orientations K=6, the number of scales in the multi-resolution decomposition S=5. The texture feature vector is:

$$t_{gabor}(I) = [\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \cdots, \mu_{29}, \sigma_{29}]$$

Where μ_{mn} is the mean of the magnitude of the transform coefficients, σ_{mn} is the standard deviation of the magnitude of the transform coefficients. In CGabor, $H_c(q)$ and t_{gabor} are combined. The distance measure is same as the one described in Section Ⅲ.B, and the weights are set as: w_1=0.4, w_2=0.6.

The third one is combining color histogram $H_c(q)$ with the Quantized Compound Change Histogram given in Section Ⅱ, which is denoted by QCCH. The distance measure is given in Section Ⅲ.B, and the weights are set as: w_1=0.4, w_2=0.6.

Five categories in ten categories are tested. The five categories are “bus”, “flower”, “horse”, “aircraft”, “fungi”.

Fig. 3 shows a retrieval result with “aircraft” as query example by using QCCH.

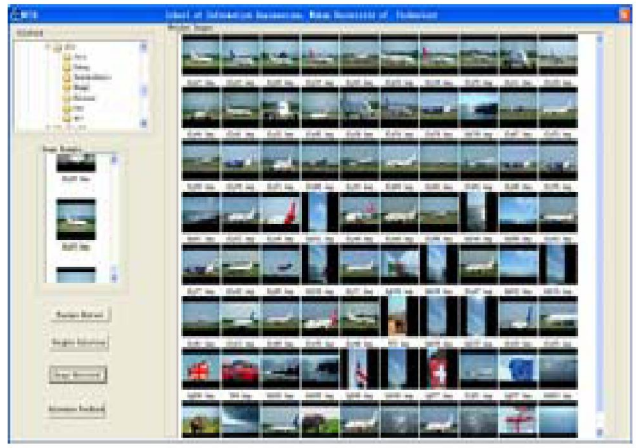

Figure 3. A retrieval result with “aircraft” as query example by using QCCH.

Precision is the most common evaluation measure used in image retrieval. It is the ratio of relevant images retrieved to the total number of images retrieved. Relevant images are referred to images in the same category.

Here, $P(N_R)$ proposed by [15] is used, $P(N_R)$ stands for the precision value after N_R images are retrieved. N_R =20, 50, 80, are set in our experiment. The average-P(20), average-P(50), average-P(80) of five image categories by three methods are showed in Table 1, Table 2, Table 3, respectively.

Table 1. The average-P(20) values of five image categories with three methods

	SCH	CGabor	QCCH
bus	0.8	0.87	0.91
flower	0.74	0.95	0.93
horse	0.89	0.80	0.9
aircraft	0.74	0.97	0.93
fungi	0.71	0.80	0.80

Table 2. The average-P(50) values of five image categories with three methods

	SCH	CGabor	QCCH
bus	0.78	0.804	0.836
flower	0.54	0.852	0.752
horse	0.708	0.688	0.728
aircraft	0.588	0.872	0.844
fungi	0.576	0.69	0.69

Table 3. The average-P(80) values of five image categories with three methods

	SCH	CGabor	QCCH
bus	0.683	0.675	0.76
flower	0.425	0.755	0.618
horse	0.55	0.567	0.6
aircraft	0.485	0.635	0.613
fungi	0.51	0.552	0.575

From Table 1 and Table 2, the retrieval performance of QCCH is almost same as that of CGabor, and outperforms SCH for all five categories.

From Table 3, CGabor can achieve better retrieval performance for image categories “flower”, “aircraft” than others, especially, for image category “flower”. QCCH can achieve better retrieval performance for image categories “bus”, “horse”, “fungi” than others. QCCH can achieve better retrieval performance than SCH for all five image categories.

The category of “flower” is a kind of natural color image with homogeneous texture regions. CGabor is more effective for the “flower” category than SCH, QCCH. The categories of “bus”, “horse”, “fungi” have complex spatial layout. QCCH is more effective for them than SCH, CGabor.

In feature extraction, the computing time of extracting Gabor wavelet texture descriptor of an image is 4.9 second, that of SCH is 0.2 second, and that of QCCH is 0.04 second. The computing time of extracting QCCH is much shorter than that of Gabor wavelet texture descriptor.

In image retrieval, the retrieval time depends on the number of feature indexes. The number of feature indexes of SCH is 166+166, that of CGabor is 166+60, and that of QCCH is 166+40. The retrieval time of QCCH is shorter than others.

For every pixel in an image, their quantized Compound Change defined in Section II is orientation independent, QCCH is a distribution of the quantized directional difference computed globally on the image, so, it is insensitive to image rotation and translation. Gabor wavelet texture descriptor is also insensitive to image rotation and translation. SCH is sensitive to image rotation and translation.

V. CONCLUSION

Texture feature is an important visual feature for image retrieval. Gabor wavelet texture descriptor has powerful descriptor power for natural color image with homogenous texture regions, it is non-efficient for others. Based on the relation of pixels in their neighbor, a texture descriptor, Quantized Compound Change Histogram, is proposed. It is insensitive to variations in image rotation and translation. Compared with other related descriptors, it has more powerful descriptor for natural color image with complex spatial layout, and has fewer feature indexes, the computing time of extracting of it is shorter. Experimental results have validated the effectiveness of the texture descriptor.

REFERENCES

[1] F. Pereira, R. Koenen, “MPEG-7: A Standard for Multimedia Content Description”, International Journal of Image and Graphics, Vol 1, pp. 527-546, 2001.

[2] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, A. Yamada, “Color and Texture Descriptors”, IEEE Transactions on Circuits and System for Video Technology, Vol 11, pp. 703-715, 2001.

[3] J. Huang, “Image indexing using color correlograms”, Proceedings of the IEEE Computer society Conference on Computer Vision and Pattern Recognition, San Juan, pp. 762-768, 1997.

[4] A. Rao, R. Srihari, and Z. Zhang, “spatial color histograms for content-based image retrieval”, Proceedings of 11th IEEE International Conference on Tools with Artificial Intelligence, pp. 554-559, 1999.

[5] L. Cinque, S. Levialdi, K. Olsen, etc., “Color-based image retrieval using spatial-chromatic histograms”, Proceedings of IEEE International Conference on Multimedia Computing and System, pp. 969-973, 1999.

[6] S. Lim, G. Lu, “Spatial statistics for content based image retrieval”, Proceedings of IEEE International Conference on Information Technology: Computers and Communications, pp. 155-159, 2003.

[7] B. S. Manjinath, W. Y. Ma, “Texture Feature for Browsing and Retrieval of Image Data”, IEEE Transactions Pattern Analysis and Machine Intelligence, Vol 18, pp. 837-841, 1996.

[8] D.-C. He, L. Wang, “Texture Features Based on Texture Spectrum”, Pattern Recognition, Vol 24, pp. 391-399, 1991.

[9] C.-B. Huang, S.-S. Yu, J.-l. Zhou, H.-W. Lu. “Image retrieval using both color and local spatial feature histograms”, Proceeding of IEEE second international conference on Communications, Circuits and Systems, Chengdu, China, pp.927-931, 2004.

[10] C.-B. Huang, S.-S. Yu, J.-L. Zhou, H.-W. Lu, “Combining Neighbor Fuzzy Entropy, Gradient Unit Feature with Color-Spatial Feature for Image Retrieval”, Proceedings of International Conference on Machine Learning and Cybernetics, Shanghai, China, pp. 4319-4324, 2004.

[11] C. B. Huang, S. S. Yu, J. L. Zhou, H. W. Lu, “Color and Neighbor Edge Directional Difference Feature for Image Retrieval”, Chinese Optics Letters, Vol 3, pp. 66-68, 2005.

[12] http: //wang.ist.puu.edu

[13] http: //elib.cs.berkeley.edu

[14] http: //www.freefoto.com

[15] H. Muller, W. Muller, D. M. Squire, “Performance evaluation in content-based image retrieval: overview and proposals”, Pattern Recognition Letters, Vol 22, pp. 593-601, 2001.