Color and neighbor edge directional difference feature for image retrieval

Chaobing Huang (黄朝兵), Shengsheng Yu (余胜生), Jingli Zhou (周敬利), and Hongwei Lu (鲁宏伟)

National Storage System Laboratory, College of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074

Received April 29, 2004

A novel image feature termed neighbor edge directional difference unit histogram is proposed, in which the neighbor edge directional difference unit is defined and computed for every pixel in the image, and is used to generate the neighbor edge directional difference unit histogram. This histogram and color histogram are used as feature indexes to retrieve color image. The feature is invariant to image scaling and translation and has more powerful descriptive for the natural color images. Experimental results show that the feature can achieve better retrieval performance than other color-spatial features.

OCIS codes: 100.2000, 100.5010.

Content-based image retrieval (CBIR) has been a hot topic of research for the last few years[1−3]. The color histogram proposed in Ref. [4] and the cumulative color histogram proposed in Ref. [5] were simple, effective, and invariant to image translation and rotation. However, a color histogram provides no spatial information. In order to incorporate spatial information in color histogram, researchers have proposed many methods[6−8]. The spatial-chromatic histogram (SCH)[7] described how pixels of identical color distributed in the image. The spatial distribution was expressed by the standard deviation of identical color bin from its centroid. The geographical statistics (Geostat)[8] described the spatial distribution of identical color with one parameter of "Looseness", which was size invariant. The texture spectrum[9] was a new image texture descriptor and described image texture in a simple way, which incorporated spatial information in an image. Another texture spectrum[10] has fewer number of feature indexes than that given in Ref. [9].

In this paper, a novel image feature termed neighbor edge directional difference unit histogram is proposed. For every pixel in the image, a neighbor edge directional difference unit is defined and computed in its 5×5 neighbor. Calculating the distribution of the neighbor edge directional difference unit over the image, the neighbor edge directional difference unit histogram is obtained, which is used as image feature. Compared with the features given by Refs. [6 − 8], our feature is invariant to scaling and translation in the image. Go a step further, our feature has more powerful descriptive for the natural color images than the features obtained by Refs. [6 − 8] and [10].

First, in order to get color feature, the HSV (hue, saturation, value) color space is used. The hue, saturation, and bright component value are denoted as h, s, v, respectively, 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. The colors with $v < 0.25$ are coded as $q = 0$. The colors with $s < 0.2$ and $v \geq 0.25$ are uniformly classified into 3 gray levels according to the values of v, they are coded as $q = 1, 2, 3$, respectively. All remaining colors ($s \geq 0.2$ and $v \geq 0.25$) are regarded as chromatic region, they are quantized into 162 color bins, the codes of them are: $q = 9 \times h + 3 \times s + v + 4$, where $h \in[0,\ 2\pi]$ is uniformly quantized into 18 bins with $h = 0, 1, \cdots, 17$, $s \in[0.2,\ 1]$ is uniformly quantized into 3 bins with value $s = 0, 1, 2$, $v \in[0.25,\ 1]$ is uniformly quantized into 3 bins with value $v = 0, 1, 2$. The quantized color set is denoted by Q, i.e. $Q = \{0, 1, 2, \cdots, 165\}$. Given a color image f, of size $m \times n$ pixels, characterized by the quantized color q at (i, j), i.e. $q = f(i, j)$. Thus, calculating the distribution of the quantized color in the value range, we can get the normalized quantized color histogram

$$H_{\mathrm{c}}(q) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\delta[f(i,j) - q], \quad \forall q \in Q, \tag{1}$$

where $\delta()$ is the unitary impulse function.

Second, we get spatial feature. A pixel and its 5×5 neighbors are taken into consideration, as shown in Fig. 1, where $y(i, j)$ is 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).

We record the salient gray value changes of the neighboring pixels in eight different directions using a binary sequence R_k (k=0, 1, 2, 3, 4, 5, 6, 7). They are defined as follows:

1) Horizontal direction: if $|y(i+2, j) - y(i-2, j)| >$BTH, $R_0 = 1$; otherwise $R_0 = 0$.

2) 22.5° angle direction: if $|y(i+2, j-1) - y(i-2, j+1)| >$BTH, $R_1 = 1$; otherwise $R_1 = 0$.

3) 45° angle direction: if $|y(i+2, j-2) - y(i-2, j+2)| >$BTH, $R_2 = 1$; otherwise $R_2 = 0$.

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

Fig. 1. 5×5 pixel neighbor edge centered at location (i, j).

1671-7694/2005/020066-03 http://www.col.org.cn

4) 67.5° angle direction: if $|y(i+1,j-2)-y(i-1,j+2)|$ >BTH, $R_3=1$; otherwise $R_3=0$.

5) Vertical direction: if $|y(i,j-2)-y(i,j+2)|$ >BTH, $R_4=1$; otherwise $R_4=0$.

6) 112.5° angle direction: if $|y(i-1,j-2)-y(i+1,j+2)|$ >BTH, $R_5=1$; otherwise $R_5=0$.

7) 135° angle direction: if $|y(i-2,j-2)-y(i+2,j+2)|$ >BTH, $R_6=1$; otherwise $R_6=0$.

8) 157.5° angle direction: if $|y(i-2,j-1)-y(i+2,j+1)|$ >BTH, $R_7=1$; otherwise $R_7=0$. Where BTH is a predefined positive threshold of gray value. As each R_k (k=0, 1, 2, 3, 4, 5, 6, 7) has two possible values, the combination of all the eight elements results in $2^8=256$ possible values in total. We use the 256 possible values to define the neighbor edge directional difference unit at pixel (i,j) as

$$r(i,j)=\sum_{k=0}^{7}2^k\times R_k. \qquad (2)$$

Thus calculating the distribution of the neighbor edge directional difference unit in the value range $R=\{0,1,2,\cdots,255\}$, we can get the normalized neighbor edge directional difference unit histogram

$$H_{\mathrm{nr}}(r)=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\delta[r(i,j)-r],\quad \forall r\in R. \qquad (3)$$

The normalized quantized color histogram ($H_{\mathrm{c}}(q)$) and the normalized neighbor edge directional difference unit histogram ($H_{\mathrm{nr}}(r)$) are used as image features. We denote the image features as the neighbor edge directional difference feature (NEDDF). A form of "relative" L_1-distance metric[6] is adopted to measure the distance of feature vectors. Let M be the query image, I be an image in the image database D. For images M and I, the distance metric between their color histograms is denoted by $d_1(M,I)$; the distance metric between their neighbor edge directional difference unit histograms is denoted by $d_2(M,I)$. The distance between two images M and I is $d(M,I)=w_1d_1(M,I)+w_2d_2(M,I)$, where w_i are weights assigned to $d_i(M,I)$ $(i=1,2)$. 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 distance components. Finally, for the query image M and every image $I\in D$, the results of content-based image retrieval can be gotten, they are ordered inversely according to the distance, $d(M,I)$.

In the experiment, 1000 24-bit JPEG color images of size 384×256 or 256×384 pixels are downloaded from a web site[2] as our image database. The images are divided into 10 categories of different scenes (bus, flower, horse, elephant, beach, building, dragon, mountain, food, person). There are 100 images in each category. To compare performance, we use four features, SCH[7], Geostat[8], TS[10], and NEDDF (we set $w_1=w_2=1$, BTH = 5). Precision is the most common evaluation measure used in image retrieval: Precision=(number of relevant images retrieved)/(total number of images retrieved).

Table 1. The Average-$P(100)$ Values of Ten Image Categories with Four Features

	SCH	Geostat	TS	NEDDF
Bus	0.545	0.46	0.55	0.67
Flower	0.475	0.415	0.655	0.665
Horse	0.62	0.58	0.63	0.675
Elephant	0.30	0.215	0.265	0.305
Beach	0.35	0.31	0.30	0.4
Building	0.35	0.305	0.405	0.43
Dragon	0.94	0.87	0.935	0.94
Mountain	0.23	0.18	0.23	0.27
Food	0.365	0.31	0.365	0.38
Person	0.475	0.385	0.405	0.52

Relevant images are referred to images in the same class. Here, $P(N_R)$[11] is used, $P(N_R)$ stands for the precision value after N_R images are retrieved, where N_R is the total number of relevant images in the image database, i.e. $N_R=100$. The average-$P(100)$ value of ten image categories with four features are displayed in Table 1.

From Table 1, it can be found, there are three image categories in SCH, two image categories in Geostat, four image categories in TS, five image categories in NEDDF, whose average-$P(100)$ values $>50\%$, the number of the image categories in NEDDF is the most. Furthermore, the average-$P(100)$ values in NEDDF are the largest ones for those in all four features in all ten categories. So, the retrieval performance of NEDDF is the best among the four features.

SCH and Geostat roughly describe the spatial arrangement of pixels using their centroid, standard deviation, and Looseness, respectively. These three features provide an acceptable approximation for a substantially homogeneous region's spatial arrangement, but loose descriptive power when pixels having the same color are arranged in more complex spatial layout. However, NEDDF has powerful descriptive for color images with more complex spatial layout. For example, there are three 24-bit JPEG color images of size 384×256 in image database displayed in Fig. 2, the content of image 1 is "a flower", where the standard deviation $\sigma(4)=131.57$, $\sigma(5)=129.18$, the Looseness $L(4)=1035.92$, $L(5)=1081.03$, and the normalized neighbor edge directional difference unit histogram $H(255)=0.192$; the content of image 2 is "persons", where $\sigma(4)=121.96$, $\sigma(5)=98.37$, $L(4)=7.18$, $L(5)=6.32$, $H(255)=0.399$; the content of image 3 is "a flower" too, where $\sigma(4)=0$, $\sigma(5)=0$, $L(4)=0$, $L(5)=0$, $H(255)=0.162$. The retrieval results of SCH using image 1 as the query image are: image 1, image 2, image 3. The retrieval results of Geostat using image 1 as the query image are: image 1, image 2, image 3, too. However, the retrieval results of NEDDF using image 1 as the query image are: image 1, image 3, image 2, which are the results we expect. TS[10] takes 3×3 neighbor into consideration, NEDDF uses 5×5 neighbor, which is a larger window and incorporates more information. TU[10] takes the change in gray value between two neighbor pixels in horizontal direction into consideration, NEDDF uses the change in gray value between

Fig. 2. Three images in our database.

two neighbor edge pixels in which one pixel is opposite to the other. NEDDF is more efficient than TU[10].

In conclusion, we have proposed a new image feature, the neighbor edge directional difference feature (NEDDF). Compared with other related features, NEDDF is insensitive to variations in image scaling and translation, and can capture more color-spatial information in an image. Experimental results have validated the effectiveness of our method. For future work, we shall integrate the low-level visual feature with high-level semantic information to retrieve and classify color image.

C. Huang's e-mail address is cbhuang@wtwh.com.cn or huang_cb@163.com.

References

1. Y. Rui, T. S. Huang, and S.-F. Chang, J. Visual Communication and Image Representation **10,** 39 (1999).
2. J. Z. Wang, J. Li, and G. Wiederhold, IEEE Trans. Pattern Analysis and Machine Intelligence **23,** 947 (2001).
3. G. Mu, H. Zhai, and S. Zhang, Chin. Opt. Lett. **1,** 583 (2003).
4. M. J. Swain and D. H. Ballard, Intern. J. Comput. Vision **7,** 11 (1991).
5. M. Strick and M. Orengo, Proc. SPIE **2420,** 381 (1995).
6. J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, in *Proceedings of the IEEE Computer society Conference on Computer Vision and Pattern Recognition* 762 (1997).
7. L. Cinque, S. Levialdi, K. A. Olsen, and A. Pellicanò, in *IEEE International Conference on Multimedia Computing and System* 969 (1999).
8. S. Lim and G. Lu, in *IEEE International Conference on Information Technology: Computers and Communications* 155 (2003).
9. D.-C. He and L. Wang, Pattern Recognition **24,** 391 (1991).
10. H. L. Wan, U. C. Morshed, H. Hu, and Z. Z. Shi, J. Computer-Aided Design & Computer Graphics (in Chinese) **15,** 195 (2003).
11. H. Müller, W. Müller, D. M. Squire, S. Marchand-Maillet, and T. Pun, Pattern Recognition Lett. **22,** 593 (2001).