Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 27 (2006) 1122–1126
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Image retrieval based on color distribution entropy

Junding Sun [a,*], Ximin Zhang [b], Jiangtao Cui [b], Lihua Zhou [b]

[a] *School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China*
[b] *Research Institute of Multimedia, Xidian University, Xi'an 710071, China*

Received 2 September 2004; received in revised form 8 November 2005
Available online 28 February 2006

Communicated by R. Manmatha

Abstract

Color histogram is an important technique for color image database indexing and retrieving. However, the main problem with color histogram indexing is that it does not take the color spatial distribution into consideration. Previous researches have proved that the effectiveness of image retrieval increases when spatial feature of colors is included in image retrieval. In this paper, two new descriptors, color distribution entropy (CDE) and improved CDE (I-CDE), which introduce entropy to describe the spatial information of colors, are presented. In comparison with the spatial chromatic histogram (SCH) and geostat which also measure the global spatial relationship of colors, the experiment results show that CDE and I-CDE give better performance than SCH and geostat.

Keywords: Annular color histogram; Color distribution entropy; Improved color distribution entropy

1. Introduction

In the past decade, content-based image retrieval (CBIR) has attracted extensive research interest. The color histogram method introduced in (Swain and Ballard, 1991) has shown to be very effective and simple to implement. However, the main disadvantage of the color histogram method is that it is not robust to significant appearance changes because it does not include any spatial information. Recently, several schemes including spatial information have been proposed. Pass et al. (1996) suggested classifying each pixel as coherent or noncoherent based on whether the pixel and its neighbors have similar colors. Then, a split histogram called *color coherence vector* (CCV) is used to represent this classification for each color in an image. Huang (1997) proposed a color *correlograms* method, which collects statistics of the co-occurrence of two colors some distances apart. A simplification of this feature is the *autocorrelogram*, which only captures the spatial correlation between identical colors. Rao et al. (1999), Cinque et al. (1999) and Lim and Lu (2003) respectively introduced annular color histogram, spatial-chromatic histogram (SCH) and geostat to describe how pixels of identical color are distributed in the image.

In this paper we propose a *color distribution entropy* (CDE) method, which takes account of the correlation of the color spatial distribution in an image. The main difference between CDE, geostat and SCH is that CDE describes how pixel patches of identical color are distributed in an image. Furthermore, three methods are presented to improve the retrieval performance of CDE. In the end, we will analyze and compare the performance of the proposed method with SCH and geostat.

The organization of the rest of paper is as follows. In Section 2, we briefly review related literature on how to represent the color spatial feature. Section 3 outlines the proposed techniques. Section 4 provides analysis on how to improve the performance of CDE. Section 5 describes the experimental results. Section 6 gives a conclusion.

* Corresponding author. Tel.: +86 29 88202244.
E-mail address: sunharry@hotmail.com (J. Sun).

0167-8655/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.12.014

2. Related work

This section describes three related work which take the global spatial relationship of the colors into consideration.

2.1. Annular color histogram

Rao et al. (1999) introduced annular color histograms to represent color spatial feature. Suppose $\boldsymbol{A}_i$ be the set of pixels with color bin i of an image and $|\boldsymbol{A}_i|$ be the number of elements in $\boldsymbol{A}_i$. Let $\boldsymbol{C}_i$ be the centroid and r_i be the radius of color bin i which are defined in (Rao et al., 1999).

With $\boldsymbol{C}_i$ as the center and with jr_i/N as the radius for each $1 \leqslant j \leqslant N$, we can draw N concentric circles. Let $|\boldsymbol{A}_{ij}|$ be the count of the pixels of color bin i inside circle j. Then the annular color histogram can be written as $(|\boldsymbol{A}_{i1}|, |\boldsymbol{A}_{i2}|, \hbar, |\boldsymbol{A}_{iN}|)$. This is illustrated in Fig. 1.

2.2. Spatial-chromatic histogram

For the spatial-chromatic histogram (SCH), Cinque et al. (1999) introduced parameter σ_i to describe how pixels of identical color are distributed in an image, which is defined as follows:

$$\sigma_i = \sqrt{\frac{\sum_{p \in \boldsymbol{A}_i} d(p, \boldsymbol{C}_i)^2}{|\boldsymbol{A}_i|}} \quad (1)$$

$$d(p, \boldsymbol{C}_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2} \quad (2)$$

where p is the position of a pixel at (x, y) and (x_i, y_i) is the coordinate of centroid $\boldsymbol{C}_i$. According to the analysis in (Lim and Lu, 2003), for a color bin i, the SCH index can be described by its histogram feature h_i and distribution feature σ_i. Therefore, the SCH index for an image can be written as $(h_1, \sigma_1, \hbar, h_i, \sigma_i, \hbar, h_n, \sigma_n)$, where n is the number of color bins.

2.3. Geostat

Lim and Lu (2003) proposed using the Looseness Parameter from geostat, a branch of statistics which deals with geographical data, to describe the global spatial relationship of colors. The Looseness Parameter defined in (Lim and Lu, 2003) is as follows:

$$L_i = \frac{1}{|\boldsymbol{A}_i|} 2\pi V_i \quad (3)$$

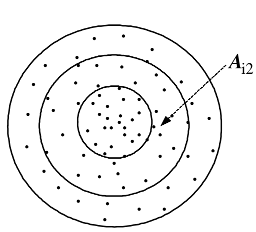

Fig. 1. Annular color histogram.

$$V_i = \frac{1}{|\boldsymbol{A}_i|} \sum_{p \in \boldsymbol{A}_i} d(p, \boldsymbol{C}_i)^2 \quad (4)$$

where $d(p, \boldsymbol{C}_i)$ was defined in formula (2). The geostat index of an image is $(h_1, L_1, \hbar, h_i, L_i, \hbar, h_n, L_n)$, where h_i is the histogram of color bin i, L_i is the looseness of color bin i and n is the number of bins.

3. Feature extraction

3.1. Color distribution entropy

Based on the annular color histogram, the NSDH (*the normalized spatial distribution histogram*) can be defined as $\boldsymbol{P}_i$ where $\boldsymbol{P}_i = (P_{i1}, P_{i2}, \hbar, P_{iN})$ and $P_{ij} = |\boldsymbol{A}_{ij}|/|\boldsymbol{A}_i|$.

John (2000) proposed using entropy which was developed by Shannon (1948) to represent color information of an image and retrieve images in CBIR. Based on the NSDH and the definition of entropy, we propose a new descriptor, CDE (*the Color Distribution Entropy*), describing the spatial information of an image. The CDE of color bin i can be defined as

$$E_i(\boldsymbol{P}_i) = -\sum_{j=1}^{N} P_{ij} \log_2(P_{ij}) \quad (5)$$

which gives the dispersive degree of the pixel patches of a color bin in an image. Large E_i means the distribution of the pixels is dispersed, otherwise the distribution is compact. Then the CDE index for an image can be written as $(h_1, E_1, \hbar, h_i, E_i, \hbar, h_n, E_n)$, where h_i is the histogram of color bin i, E_i is the CDE of color bin i and n is the number of bins.

Because of the lower dimension indexes, SCH, geostat and CDE are more efficient than annular color histograms. Additionally, annular color histograms mentioned in (Rao et al., 1999) are size variant because it is relative to the number of pixels of a color bin in the annular circles. The parameter σ described in (Cinque et al., 1999) is also size variant because it is relative to the number of pixels and the density of these pixels of a color bin. Normalized by the number of pixels in the bin, L mentioned in (Lim and Lu, 2003) is size invariant. In this paper, NSDH is size invariant because annular color histograms are normalized by the number of pixels of the color bins, therefore, the CDE is also size invariant.

3.2. Similarity measurement

With reference to the definition of similarity metric in (Lim and Lu, 2003, formula (5)), the distance of images $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ can be noted as

$$d(\boldsymbol{q}_1, \boldsymbol{q}_2) = \sum_{i=1}^{n} \min(h_i^{\boldsymbol{q}_1}, h_i^{\boldsymbol{q}_2}) \times \frac{\min(E_i^{\boldsymbol{q}_1}, E_i^{\boldsymbol{q}_2})}{\max(E_i^{\boldsymbol{q}_1}, E_i^{\boldsymbol{q}_2})} \quad (6)$$

With the same meaning as formula (5) defined in (Lim and Lu, 2003), the similarity metric is made up of two parts; the first part, $\min(h_i^{q_1}, h_i^{q_2})$, is histogram intersection which measures the similarity between $h_i^{q_1}$ and $h_i^{q_2}$, while the second part $\frac{\min(E_i^{q_1}, E_i^{q_2})}{\max(E_i^{q_1}, E_i^{q_2})}$ measures the similarity of the spatial distribution of a bin. In addition, $\frac{\min(E_i^{q_1}, E_i^{q_2})}{\max(E_i^{q_1}, E_i^{q_2})}$ is more meaningful than $\min(E_i^{q_1}, E_i^{q_2})$ or $\text{abs}(E_i^{q_1} - E_i^{q_2})$ according to the explanation of the formula (5) defined in (Lim and Lu, 2003).

4. Improvement

4.1. Removing the influence of small isolated color patches

Let p_i be the normalized color histogram and E_i the CDE of bin i. Because humans are more sensitive to color patches than to isolated pixels, color bin i is defined as an insensitive color when $p_i < \alpha_1$ and $E_i > \alpha_2$. Here, α_1 and α_2 are two thresholds. Such bins will not be considered when calculating CDE. That is to say, if the pixels of bin i are few and widely separated around the centroid in an image, bin i will be discarded.

4.2. The weighted CDE

Correspondingly, the pixels of bin i are more dispersed in the far annular circles than those in the near annular circles to the centroid. A weight function $f(j)$, which denotes the different contribution of each annular circle to the CDE, is introduced. Here j represents the different annular circle. Because the more dispersed the color bin, the higher the entropy, the weight function $f(j)$ should satisfy $f(j_1) > f(j_2)$ for $j_1 > j_2$. Here $j_1 > j_2$ means that annular circle j_1 is out of j_2. By $f(j)$, the effect of the near circles to the centroid on human vision is strengthened. On the other hand, the effect of the farther circles is weakened.

We introduce the weight function as follows:

$$f(j) = 1 + \frac{j}{N} \tag{7}$$

Accordingly, the above formula (5) can be further written as follows:

$$E_i(\boldsymbol{P}_i) = -\sum_{j=1}^{N} f(j) P_{ij} \log_2(P_{ij}) \tag{8}$$

4.3. Removing the influence of symmetrical property of entropy

Because of the symmetrical property of entropy, perceptually dissimilar color histograms may have the same entropy. That is to say, the entropy is invariant, while the element's order in the color histogram vector is changeable. For example, by using the entropy computed by Formula (5), we cannot distinguish the histograms $\boldsymbol{H}_1$, $\boldsymbol{H}_2$ and $\boldsymbol{H}_3$ in Fig. 2 because they have the same entropy.

In order to solve such influence, a new method, called histogram area, is proposed in this paper. Let $\boldsymbol{H} = \{p_1, p_2, \hbar, p_n\}$. The area of histogram $\boldsymbol{H}$ is defined as follows:

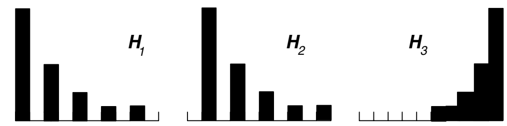

Fig. 2. Three examples of histogram.

$$A(\boldsymbol{H}) = \sum_{i=1}^{n} (p_i \times i) \tag{9}$$

It is clear that the histogram area, $A(\boldsymbol{H})$, gets maximum n when $p_n = 1$. Though the histograms $\boldsymbol{H}_1$, $\boldsymbol{H}_2$ and $\boldsymbol{H}_3$ have the same entropy, they have different histogram areas because the order of the element in different histograms is always different. Therefore, we introduce a weight function

$$g(\boldsymbol{H}) = 1 + \frac{A(\boldsymbol{H})}{n} \tag{10}$$

Using $g(\boldsymbol{H})$, the entropy of a histogram can be written as follows:

$$E(\boldsymbol{H}) = -g(\boldsymbol{H}) \sum_{i=1}^{n} p_i \log_2(p_i) \tag{11}$$

Let $\boldsymbol{H}_1 = \{1/2, 0, 1/4, 0, 1/8, 0, 1/16, 0, 1/16, 0\}$, $\boldsymbol{H}_2 = \{0, 1/2, 0, 1/4, 0, 1/8, 0, 1/16, 0, 1/16\}$, $\boldsymbol{H}_3 = \{0,0,0,0,0,1/16, 1/16, 1/8, 1/4, 1/2\}$ in Fig. 2. We can get $E(\boldsymbol{H}_1) = 2.41$, $E(\boldsymbol{H}_2) = 2.60$ and $E(\boldsymbol{H}_3) = 3.57$. It is clear that the weighted entropy computed by Formula (11) can effectively remove the influence of symmetrical property of entropy.

Combining the weight function $f(j)$ in Section 4.2 with the weight function $g(\boldsymbol{H})$ in Section 4.3, the improved CDE (I-CDE) can be defined as

$$E_i(\boldsymbol{P}_i) = -g(\boldsymbol{P}_i) \sum_{j=1}^{N} f(j) P_{ij} \log_2(P_{ij}) \tag{12}$$

Suppose $\boldsymbol{H}_1$, $\boldsymbol{H}_2$ and $\boldsymbol{H}_3$ in Fig. 2 be the NSDH of bin i of the three images respectively. We can get $E_i(\boldsymbol{H}_1) = 3.43$, $E_i(\boldsymbol{H}_2) = 3.95$ and $E_i(\boldsymbol{H}_3) = 6.58$. Then, the distances between them are $d_{L_1}(\boldsymbol{H}_1, \boldsymbol{H}_2) = 0.52$, $d_{L_1}(\boldsymbol{H}_1, \boldsymbol{H}_3) = 3.15$ and $d_{L_1}(\boldsymbol{H}_2, \boldsymbol{H}_3) = 2.63$ by L_1 distance metric. Consequently, the three histograms can be distinguished effectively. The results also show that $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ is similar and they are dissimilar to $\boldsymbol{H}_3$.

5. Experimental results

Experiments were carried out by using a database of about 8000 images which included natural scenes, indoor images, plants, animals, landscapes, people, vehicles, etc. All the images are down-loaded from various web sites so that no pre-restrictions such as camera type, brightness, rotation, zoom are specified for the testing. We indexed the image database by using CDE and I-CDE in HSV color space. The color space is uniformly quantized into nine lev-

Table 1
Range of α_1 and α_2

Testing set	α_1	α_2
First	[0.39%, 0.76%]	[3.7, 4.8]
Second	[0.43%, 0.86%]	[3.3, 4.6]
Third	[0.45%, 0.91%]	[3.3, 4.2]
Fourth	[0.40%, 0.89%]	[3.5, 4.4]

els of hue, three levels of saturation and value giving a total of 81 bins.

To choose the appropriate threshold values α_1 and α_2 described in Section 4.1, we collected four sets of testing data from the database. The first set consisted of 650 landscape images, the second consisted of 1000 animal images, the third consisted of 1000 vehicles images, and the forth set consisted of 2600 images including natural scenes, indoor images, plants, animals, landscapes, people, vehicles, paintings, etc. so that the variety of images in the test set prevented any bias toward particular type of images. Table 1 gives the range of α_1 and α_2. If α_1 and α_2 were located in the range shown in Table 1, good performance could be achieved according to different test sets. The thresholds used in the results below were $\alpha_1 = 0.7\%$ and $\alpha_2 = 3.8\%$.

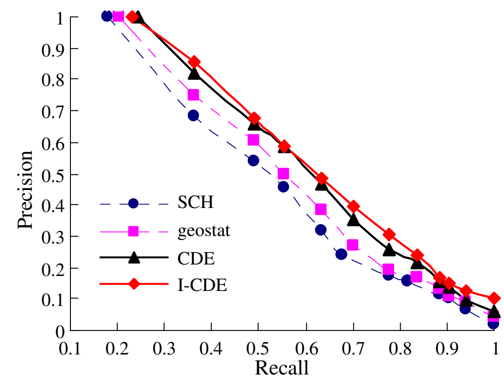

Fig. 3. Average recall-precision graph averaged over 50 query images.

Table 2
Results of ANMRR for the four methods

Methods	ANMRR
I-CDE	0.2803
CDE	0.2935
Geostat	0.3156
SCH	0.3408

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

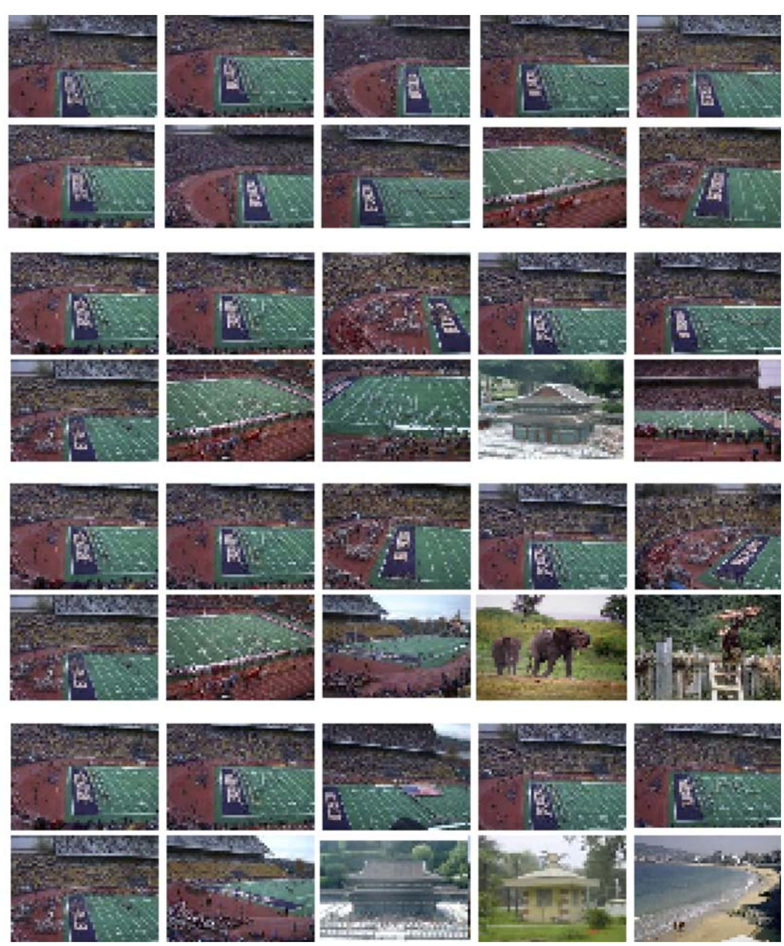

Fig. 4. Query images: (a) I-CDE, (b) CDE, (c) geostat and (d) SCH.

The retrieval accuracy was measured in terms of the recall, precision and ANMRR. The precision rate and recall rate are defined as follows (Choi et al., 2003):

$$\text{precision} = \frac{\text{number of relevant images selected}}{\text{total number of retrieved images}} \tag{13}$$

$$\text{recall} = \frac{\text{number of relevant images selected}}{\text{total number of similar images in the database}} \tag{14}$$

To measure the effectiveness of the methods, we also use the MPEG-7 retrieval metric the Average Normalized Modified Retrieval Rank (ANMRR). The ANMRR indicates not only how many of the correct items are retrieved, but also how highly they are ranked among the retrieved items. It is defined in (Manjunath et al., 2001) as

$$\text{ANMRR} = \frac{1}{\text{NQ}} \sum_{q=1}^{\text{NQ}} \frac{\frac{1}{\text{NG}(q)} \sum_{k=1}^{\text{NG}(q)} \text{Rank}^*(k) - 0.5 \times [1 + \text{NG}(q)]}{1.25K(q) - 0.5 \times [1 + \text{NG}(q)]} \tag{15}$$

where

$$\text{Rank}^*(k) = \begin{cases} \text{Rank}(k), & \text{if } \text{Rank}(k) \leqslant K(q) \\ 1.25K, & \text{if } \text{Rank}(k) > K(q) \end{cases}$$

$\text{NG}(q)$ is the size of the ground truth set for a query image q, $\text{Rank}(k)$ is the ranking of the ground truth images by the retrieval algorithm, NQ is number of query images, and $K(q)$ specifies the "relevant ranks" for each query. As the size of the ground truth set is normally unequal, a suitable $K(q)$ is determined by

$$K(q) = \min(4^*\text{NG}(q), 2^*\text{GTM}) \tag{16}$$

where GTM is the maximum of $\text{NG}(q)$ for all queries.

In order to evaluate the performance of the proposed method, CDE and I-CDE were compared with SCH and geostat in HSV color space. At the same time, fifty independent images belonging to ten types were chosen from the database as queries. Fig. 3 presents the precision-recall graph of the results averaged over 50 queries. Table 2 gives a comparison of the four methods in ANMRR. Fig. 4 shows the top ten retrieval results of a query (the top left image is the query image and the retrieved image). The retrieval results show that the proposed methods CDE and I-CDE give better performance in image retrieval than SCH and geostat. Meanwhile, I-CDE produces a better average recall-precision and ANMRR than CDE.

6. Conclusion

We introduced CDE as a method of measuring the global spatial relationship of colors in an image. Moreover, according to the properties of entropy, the improved CDE, I-CDE, is presented to strengthen the retrieval performance. Because CDE and I-CDE describe the distribution of pixels patches of a color bin in an image, they are more consistent with human vision than geostat and SCH which describe the distribution of isolated pixels of a color bin in an image. The experimental results show that CDE and I-CDE give better performance than geostat and SCH.

Acknowledgment

This research was supported by the doctor fund of Henan Polytechnic University (B050901).

References

Choi, W.P., Lam, K.M., Siu, W.C., 2003. Maximal disk based histogram for shape retrieval. ICASSP, Hong Kong 3, 9–12.

Cinque, L., Levialdi, S., Olsen, K., et al., 1999. Color-based image retrieval using spatial-chromatic histogram. In: Multimedia Computing and System, IEEE International Conference on Multimedia Computing and Systems, vol. 2, pp. 969–973.

Huang, J., 1997. Image indexing using color correlograms. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Juan., pp. 762–768.

John, Z.M., 2000. An Information Theoretic Approach to Content Based Image Retrieval, Louisiana State University and Agricultural and Mechanical College, Ph.D. thesis, pp. 45–62.

Lim, S., Lu, G.J., 2003. Spatial statistics for content based image retrieval. In: Proceedings of the International Conference on Information Technology: Computers and Communications, Las Vegas, Nevada, pp. 28–30.

Manjunath, B.S., Ohm, J.R., Vasudvan, V., et al., 2001. Color and texture descriptors. IEEE Trans. Circ. Syst. Video Tech. 11 (6), 703–715.

Pass, G., Zabih, R., Miller, J., 1996. Comparing images using color coherence vectors. In: ACM 4th International Conference on Multimedia, Boston, Massachusetts, United States, pp. 65–73.

Rao, A.B., Srihari, R.K., Zhang, Z.F., 1999. Spatial color histogram for content-based retrieval. Tools with artificial intelligence. In: Proceedings of 11th IEEE International Conference. pp. 183–186.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423.

Swain, M.J., Ballard, D.H., 1991. Color indexing. Int. J. Comput. Vision 7 (1), 11–32.