Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Pattern Recognition Letters 24 (2003) 1935–1941

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Image retrieval based on perceptive weighted color blocks

Xuelong Li

Information Processing Center, Department of Electron. Eng. & Information Science, University of Science and Technology of China, Hefei, Anhui 230027, PR China

Received 26 August 2002; received in revised form 16 January 2003

Abstract

A novel algorithm based on running sub-blocks with different similarity weights is proposed for object-based image retrieval. By splitting the entire image into certain sub-blocks, we use color region information and similarity matrix analysis to retrieval images under the query of special object.

Keywords: Image retrieval; Sub-block; Perceptive weighted

1. Introduction

With the tremendous growth of digital images, content-based image retrieval (CBIR) has gained much attention (Smeulders et al., 2000; Remco and Mirela, 2000; Li, 2002a) in recent years. In order to retrieve a group of content-related images from the image databases, various CBIR algorithms have been proposed, and some systems (Niblack et al., 1993; Pentland et al., 1994) have been developed according to these algorithms. In these algorithms, visual features (color, texture, shape and etc. (Tao et al., 2002a)) are extracted from the images, and used as a criterion to retrieve those content-related images from the databases. Though the retrieval processing was approached in three levels (pixel-level, low-level features and high-level concepts) originally, the semantic gaps between them are being decreased (Smeulders et al., 2000; Yuan et al., 2002).

Of all the proposed approaches based on low-level features, color histogram (Swain and Ballard, 1991; Hafner et al., 1995; Li et al., 2002) is employed extensively. From the mathematical viewpoint, color histogram can be regarded as a color distribution over a certain color space. Therefore, difference color spaces can be employed in CBIR algorithms to achieve an optimum performance, such as RGB, HSV, CIELuv (Plataniotis and Venetsanopoulos, 2000), etc.

On a certain color space, each color channel is quantized into some bins, and the similarity between the query image and the images in the databases can be calculated over these bins by distance metric calculation. To describe the similarity between query image and the images in databases, many metrics are proposed, for example,

E-mail address: xuelong_li@hotmail.com (X. Li).

0167-8655/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0167-8655(03)00032-1

histogram intersection (Swain and Ballard, 1991), Euclidean metric (Hafner et al., 1995), quadratic distance (Niblack et al., 1993), etc.

However, color histogram can only represent the coarse distribution of colors over a certain color space, and all spatial information is discarded. The histogram specifies the number of pixels in each color bin, but shows no information about the location of these pixels within the image. Therefore, simple color histogram matching may provide many false-alert retrieval results.

In order to distinguish different images with similar histograms, several methods (Gong et al., 1995; Chua et al., 1997; Lu et al., 1994; Pass et al., 1996) are proposed to incorporate spatial information into the color histogram to avoid the false-alert retrieved images. In these approaches based on spatial-color histogram and color layout (color feature and spatial relations), the whole image is divided into sub-blocks, and color features are extracted for each sub-block, and then the similarity between corresponding sub-blocks is calculated.

Actually, object is an important concept in Moving Picture Experts Group (MPEG) (MPEG Requirements Group, 1998a,b), and image semantic segmentation is one way for object-based image retrieval, for example, some region-based image retrieval systems have been built (Deng and Manjunath, 2001; Carson et al., 2002), and some methods have been proposed for matching the sub-regions (Wang et al., 2001), but SEMANTIC is still a hard mission (Tao et al., 2002b), because of not only the semantic problem but also the computing speed for the tremendous growth of digital visual information.

These approaches can work well under the condition that the image contents are fixed over a certain background. However, with the spatial-color-histogram-based CBIR technology, some content-related images with the query content set on different backgrounds can be missed due to the one-to-one sub-block histogram matching. Thus, another approach should be proposed to implement CBIR according to the movable image objects.

In this paper, a novel CBIR algorithm based on running sub-blocks with different similarity weights for the image retrieval by movable contents is presented. With the proposed algorithm, the CBIR system can be extended to retrieve the images with the same content located on different areas.

In the next section, the proposed CBIR algorithm will be presented with detailed analysis. In Section 3, experimental results from test bed will be provided. Finally, Section 4 concludes the proposal and forecasts some possible future work. Acknowledgements are also given in the following.

2. Algorithm

Like other CBIR approaches based on sub-block color histogram, the first step of our algorithm is to split the images into different sub-blocks. In order to retrieve the images with the query content located on different areas, the query sub-block is running over the sub-blocks of the compared image to process histogram matching. Moreover, different parts in an image contribute difference effect to human vision perception (Matlin and Foley, 1991). To take this effect into consideration, different weights are added to sub-blocks according to their different locations. Therefore, the sub-blocks are compared not only by their locations in an image but also the perceptions, and this algorithm can rapid improve the retrieval performance for certain kinds of objects with little influence of the entire color histograms.

Let I_1, I_2 be the query image and an image in the image database respectively, and SBS_1, SBS_2 are the sub-blocks-sets corresponding to I_1 and I_2. Define a kind of image blocks set:

$$\mathrm{SBS}_k = \{\mathrm{sbs}_{ij} | \mathrm{sbs}_{ij} \subset I_k\}$$

where sbs_{ij} is a sub-block in image I_k ($k = 1, 2$), and i, j are the sub-blocks' sequences in the entire image, what's more, we assume that there are $\|\mathrm{SBN}_1\| = m \times n$ sub-blocks in I_1, and $\|\mathrm{SBN}_2\| = l \times k$ sub-blocks in I_2.

Then, the similarity matrix between the query image and the sampled image can be given by (1).

$$M(\mathrm{sbs}_{i'j'}, \mathrm{sbs}_{i''j''}) \underset{\substack{i'<m, j'<n, \mathrm{sbs}_{i'j'} \subset I_1; \\ i''<l, j''<k, \mathrm{sbs}_{i''j''} \subset I_2}}{=} \begin{bmatrix} C_{\langle 0,0,0,0\rangle} & C_{\langle 0,0,0,1\rangle} & \cdots & C_{\langle 0,0,0,k\rangle} & C_{\langle 0,0,k,0\rangle} & \cdots & C_{\langle 0,0,1,k\rangle} & \cdots & C_{\langle 0,0,l,k\rangle} \\ C_{\langle 0,1,0,0\rangle} & C_{\langle 0,0,0,0\rangle} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{\langle 0,n,0,0\rangle} & \cdots & \cdots & C_{\langle 0,n,0,k\rangle} & C_{\langle 0,0,0,0\rangle} & \cdots & C_{\langle 0,n,1,k\rangle} & \cdots & C_{\langle 0,n,l,k\rangle} \\ C_{\langle 1,0,0,0\rangle} & \cdots & \cdots & C_{\langle 1,0,0,k\rangle} & C_{\langle 0,0,0,0\rangle} & \cdots & C_{\langle 1,0,1,k\rangle} & \cdots & C_{\langle 1,0,l,k\rangle} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{\langle 1,n,0,0\rangle} & \cdots & \cdots & C_{\langle 1,n,0,k\rangle} & C_{\langle 1,n,1,0\rangle} & \cdots & C_{\langle 1,n,1,k\rangle} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & C_{\langle m,n-1,l,k\rangle} \\ C_{\langle m,n,0,0\rangle} & \cdots & \cdots & C_{\langle m,n,0,k\rangle} & C_{\langle m,n,1,0\rangle} & \cdots & \cdots & C_{\langle m,n,l,k-1\rangle} & C_{\langle m,n,l,k\rangle} \end{bmatrix} \tag{1}$$

where $C(i', j', i'', j'') = C(\mathrm{sbs}_{i'j'}, \mathrm{sbs}_{i''j''})$ is the distance metric between two sub-blocks ($\mathrm{sbs}_{i'j'}$ and $\mathrm{sbs}_{i''j''}$). From this matrix, the similarity between any sub-blocks of the query image and the sampled image can be found by our algorithm. With the information, the images with the same content located on different parts can be retrieved for post-processing. For example, when a bird moving from sky to the forest (Fig. 1), it will be difficult to retrieve the related images from a large database by traditional methods, while we can find that one or more sub-blocks are really similar to some of our matching blocks from the query image I_1.

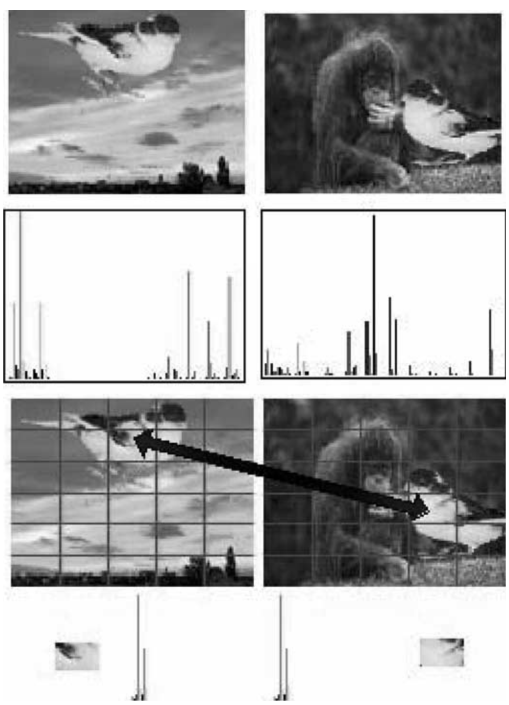

Fig. 1. The GCH and the sub-blocks.

The overall similarity between the images can be given by an experimental formula:

$$D(I_1, I_2) = \sum_{\substack{i'<m, j'<n, \mathrm{sbs}_{i'j'} \subset I_1; \\ i''<l, j''<k, \mathrm{sbs}_{i''j''} \subset I_2}} \omega_{i'j'} / [C(\mathrm{sbs}_{i'j'}, \mathrm{sbs}_{i''j''}) + \sigma] \tag{2}$$

$\omega_{i'j'}$ is the similarity weight of every sub-blocks which will be discussed later in this paper and σ is a fixed value for the case that $C(\mathrm{sbs}_{i'j'}, \mathrm{sbs}_{i''j''})$ is zero, and it is defined as 0.1 in the proposed scheme. This overall similarity provides flexibility for CBIR criterion.

In the following, the proposed CBIR algorithm is described with two detailed steps: image sub-block splitting and similarity measurement.

2.1. Image sub-blocks splitting

For our CBIR algorithm, the number of split image sub-blocks mainly determines the algorithm complexity and performance. In general, there are three ways to split the query and sampled image into sub-blocks: static splitting, semi-dynamic splitting and dynamic splitting.

For static sub-block splitting (Gong et al., 1995; Chua et al., 1997), the image is split into sub-blocks according to a fixed pattern. This way is simple to implement, but the retrieval may have no significant meaning since the size of the query content varies in different image.

For semi-dynamic sub-blocks splitting (Lu et al., 1994), the number of the sub-blocks is increased by the rule of $1/4^q$, and the user should predefine a threshold for the steps. By this way, the color spatial detail will be more and more precise, and it is also an important character of our proposed CBIR algorithm.

In fact, the size of the query object in the image is a very important factor to decide how to split the image into sub-blocks. Thus, we propose a dynamic sub-block splitting method based on the size of the query object. In this splitting method, the image is split into sub-blocks by the size of object in the query image. Therefore, balance between the algorithm complexity and retrieval precision can be achieved with dynamic sub-block splitting based on the size of object in the query image.

2.2. Similarity measurement

After splitting the image into small sub-blocks, jobs are to extract color features of the sub-blocks, and to calculate the similarity between the query image and the sampled image in the database.

For algorithm optimization, we set a size threshold τ. When a sub-block size is less than the pre-determined threshold τ, and τ can be selected as 16×16. We regard the average value of all the pixels in this sub-block as its feature value; otherwise, if the sub-block size is larger than the threshold τ, color histogram of this sub-block is used to describe its color attribute.

Various kinds of color histogram can been presented in different color spaces, such as RGB, HSV, CIELuv and etc. In our algorithm, the CIELuv color space is preferred for its precision. A pixel with a RGB value can be changed into CIELuv color space (Plataniotis and Venetsanopoulos, 2000).

$$L^* = \begin{cases} 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16 \rightarrow & \text{if } \frac{Y}{Y_n} > 0.008856 \\ 903.3\left(\frac{Y}{Y_n}\right) \rightarrow & \text{otherwise} \end{cases}$$

$$u^* = 13L^*(u' - u'_n)$$

$$v^* = 13L^*(v' - v'_n)$$

where

$$\begin{cases} u' = 4X/(X+15Y+3Z) \\ v' = 9Y/(X+15Y+3Z) \end{cases} \begin{cases} u'_n = 4X_n/(X_n+15Y_n+3Z_n) \\ v'_n = 9Y_n/(X_n+15Y_n+3Z_n) \end{cases}$$

X_n, Y_n and Z_n are the X, Y and Z of the chosen reference white.

In CIELuv color space, the distance metric between two colors is given as follow:

$$d_{ij} = \sqrt{(L_i - L_j)^2 + (u_i - u_j)^2 + |(v_i - v_j)^2}$$

Every of the color channels, L, U or V, is quantized into several bins, and the numbers of the bins are N_L, N_u and N_v.

Thus, the whole color space is fragmentized to N_{total} areas, and these areas can build a global color histogram (GCH) where

$$N_{\text{total}} = N_L * N_u * N_v$$

and the distance metric between two color histograms can be calculated by:

$$D = \sum_{i}^{N_{\text{total}}} \sum_{j}^{N_{\text{total}}} (1 - \mathrm{d}_{ij}/\mathrm{d}_{\max})(x_i - y_i)(x_j - y_j)$$

where x and y are the values of every corresponding bin and $\mathrm{d}_{\max}$ is the maximum value of d_{ij}.

Moreover, it was found that human eyes are naturally drawn towards the center of an image (Matlin and Foley, 1991). If an object is located near the center, it is more notable.

Thus, we arrange the sub-blocks as Fig. 2, and assign a similarity weight to every sub-block as $1/\log R$, where R is the Euclidean distance between a certain sub-block and the central sub-block.

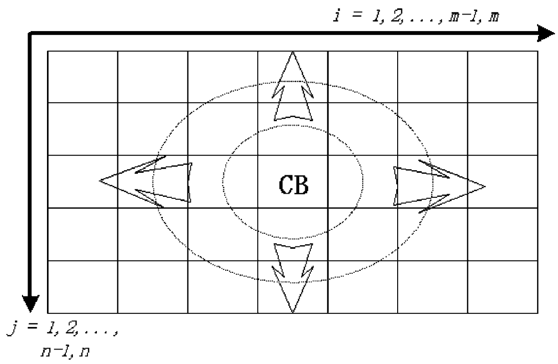

Fig. 2. Similarity weight by $1/\log R$, where CB stands for the central block.

The central sub-block can be located as follows. Assume the sub-block-sets of an image has $m * n$ sub-blocks, the central one of all the blocks can be found by:

$$\text{CentralBlock} = \text{sbs}_{([\frac{m}{2}]+1),([\frac{n}{2}]+1)}$$

$$\text{SBS}_k = \left\{ \text{sbs}_{ij} \middle| \begin{array}{l} \text{sbs}_{ij} \subset I_k \\ i < m, j < n \end{array} \right\}$$

where $[\cdot]$ is the Gauss function to get the maximum integer small than the input parameter.

The similarity matrix and the overall similarity between the query image and the sampled image can be calculated using (1) and (2).

3. Experimental results

According to our proposed CBIR algorithm of running sub-blocks with different similarity weight, a testbed has been realized with Delphi programmable language and WIN98.

With the test bench, some retrieval experiments (Fig. 3) are performed as shown in two steps. The first step is to calculate the overall similarity value. Before the experiments, we assume a pre-set threshold value for the overall similarity value. If the overall similarity value between a sampled image and the query image is larger than the threshold value, the sampled image is selected from the database for the second step. The second step is to calculate the similarity matrix. If there is any sub-block in the matrix whose values are small than a certain value, it is claimed that the image has the object that is related to the query image. Otherwise, the image temporally retrieved at the first step is omitted.

In Fig. 3, the query image is an image with a bird as the main object. First, the overall similarity is calculated. Then case (a), (b) and (c) are retrieved from the database since the overall similarity value is over the pre-set threshold τ'. However, image (d) has no the query object. Thus, in the second step, the similarity matrix M described in (2) are analyzed as shown in Fig. 4. The values stand for similarity between sub-blocks. If two sub-blocks own the same histograms, the value should be zero, while two quite different sub-blocks will return a high value.

The areas of small values are located nearly in Fig. 4(a), while they are scattered in Fig. 4(b), this shows: though two sampled image may give the approximate similarities to a query image containing a certain object, their difference will be reflected in our approached matrix. Thus, sampled image (a) can be retrieved finally, and sampled image (b) should be discarded.

We got some preliminary results on the small image database, and test hundreds of queries. E.g., Fig. 5 shows the retrieval results by traditional GCH and our method; our results are much relevant to the query object. Because the entire image is partitioned into many sub-parts, we can identify certain object without much concern with the

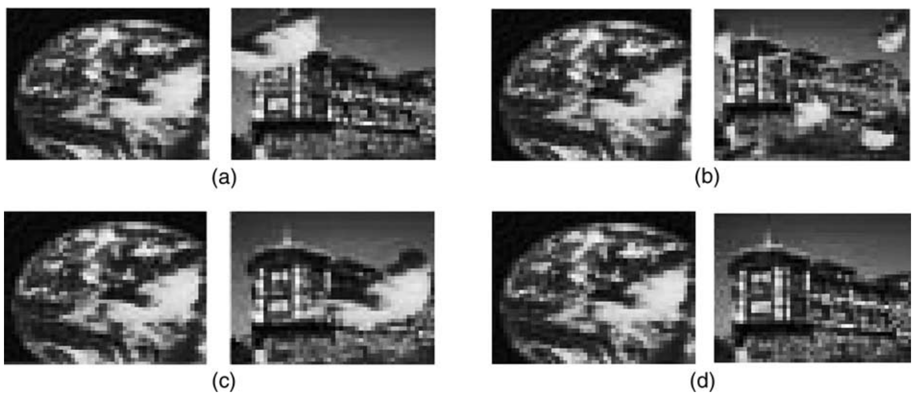

Fig. 3. Query image & sampled images overall similarity: (a) 0.759, (b) 0.690, (c) 0.898, (d) 0.182.

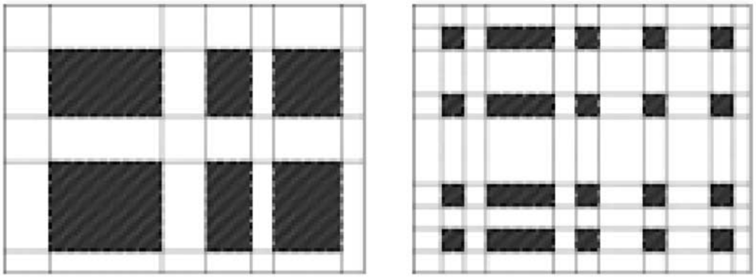

Fig. 4. The similarity matrix analysis (SMA). ( ) stand for areas of small values; the left (a) and the right (b) figures stand for the M matrix between images in Fig. 3(a) and (b) respectively.

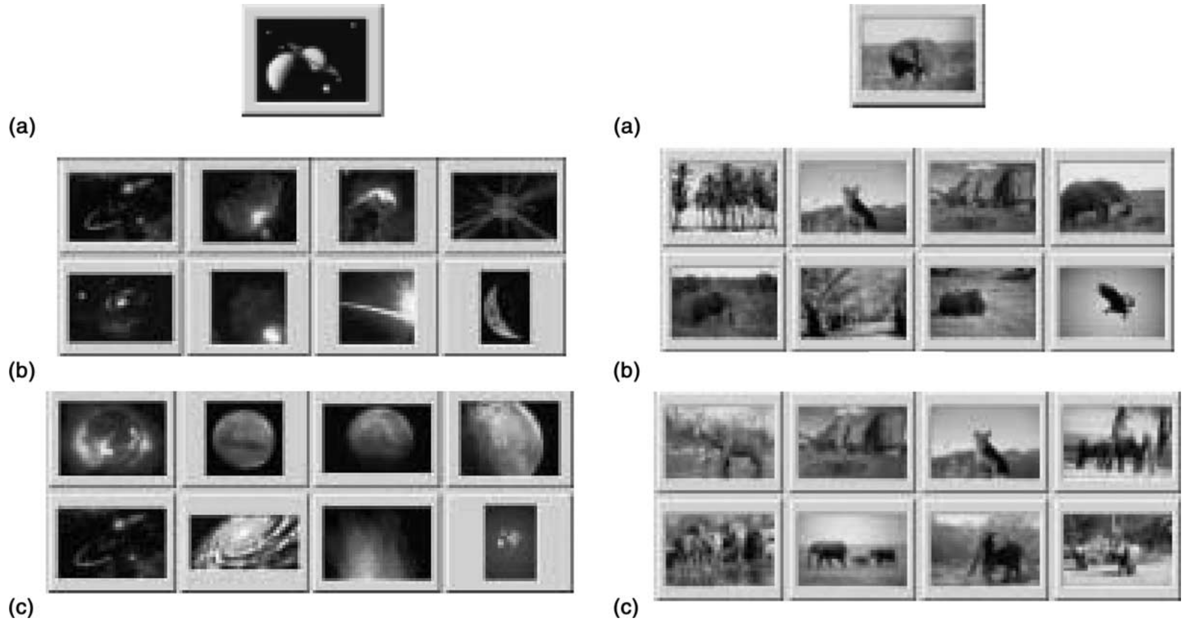

Fig. 5. (a) Query image, (b) GCH, (c) SMA.

background; in another words, our algorithm is region-based. In the future, we will use some more image segmentation algorithms to divide images into sub-regions. By this way, we can also build an indexing structure, which is organized by sub-regions and spatial relationships to accelerate the indexing/retrieval speed.

4. Conclusions

CBIR has become an important research field in multimedia information processing. In this paper, we present a novel CBIR algorithm based on running sub-blocks with different similarity weights. It is a two-step retrieval processing: (1) overall similarity calculation and threshold decision; (2) similarity matrix calculation and sub-block recognition. By splitting the entire image into sub-blocks, we use color-layout information to retrieval images under the query of special object. With the two-step image retrieval processing, our algorithm can retrieve content-related images from a large image database and discard those images without the query object. Further more, after the CBIR system retrieve the images with the same content located on different parts of the sampled images from the image database, some post-processing, such as relevance feedback (Li, 2002b) can be incorporated to enhance the CBIR systems for other applications.

Acknowledgements

Dr. Xuelong Li would like to thank Prof. Nenghai Yu and Prof. Zhengkai Liu (with Dept. of Elec. Engi. & Info. Sci., Univ. of Sci. & Tech. of China), Prof. Xiaoou Tang (with Dept. Info. Engi., Chinese Univ. of Hong Kong), Dr. Yong-gang Wen (with Dept. E.E., Massachusetts Institute of Technology) and Dr. Yuan Yuan (with Dept. E.E., University of Bath), for their kind help during the present work.

The research is supported by the Nature Science Fund of Anhui Province, PR China, under grant no. 03042307.

References

Carson, C., Belongie, S., Greenspan, H., Malik, J., 2002. BlobWorld: image segmentation using Expectation-Maximization and its application to image query. IEEE Trans. Pattern Anal. Machine Intell. 24 (8), 1026–1038.

Chua, T.S., Tan, K., Ooi, B.C., 1997. Fast signature-based color-spatial image retrieval. In: IEEE Int. Conf. on CVPR, pp. 362–369.

Deng, Y., Manjunath, B.S., 2001. Unsupervised segmentation of color-texture regions in images and video. IEEE Trans. Pattern Anal. Machine Intell. 23 (8), 800–810.

Gong, Y., Chua, H., Guo, X., 1995. Image indexing and retrieval based on color histogram. In: Proc. 2nd Int. Conf. on Multimedia Modeling, pp. 115–126.

Hafner, J., Sawhney, H.S., Wquitz, W., 1995. Efficient color histogram indexing for quadratic form distance functions. IEEE Trans. Pattern Anal. Machine Intell. 17 (7), 729–736.

Li, X.L., 2002a. A security-based visual information retrieval scheme. In: IEEE Internat. Conf. Machine Learning and Cybernetics, pp. 1119–1121.

Li, X.L., 2002b. Content based image retrieval system with new low-level features, new similarity metric, and novel feedback learning. In: IEEE Internat. Conf. Machine Learning and Cybernetics, pp. 1126–1132.

Li, X.L., Tao, D.C., Yuan, Y., Yu, N.H., Liu, Z.K., Tang, X.O., 2002. Illumination-invariant color histogram spectrum feature for content-based image retrieval. In: IEEE Internat. Conf. Methods & Models in Automation & Robotics, pp. 555–559.

Lu, H., Ooi, B.C., Tan, K., 1994. Efficient image retrieval by color contents. In: Internat. Conf. on Applications of DB.

Matlin, M.W., Foley, H.J., 1991. Sensation and Perception, third ed. Allyn and Bacon.

MPEG Requirements Group, 1998a. MPEG-7: Context & Objectives. ISO/IEC JTC1/SC29/WG11/N2 207.

MPEG Requirements Group, 1998b. MPEG-7: Requirements. ISO/IEC JTC1/C29/ WG11/N2 208.

Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Pektovic, D., Yanker, P., Faloutsos, C., Taubin, G., 1993. The QBIC project: querying images by content using color, texture, and shape. SPIE Storage and Retrieval for Image and Video Databases, pp. 173–181.

Pass, G., Zabih, R., Miller, J., 1996. Comparing images using color coherence vectors. In: ACM 4th Internat. Conf. Multimedia, pp. 65–73.

Pentland, A., Picard, R., Sclaroff, S., 1994. Photobook: Content-based manipulation of image databases. SPIE Storage and Retrieval for Image and Video Databases II, pp. 34–47.

Plataniotis, K.N., Venetsanopoulos, A.N., 2000. Color image processing and applications. Berlin, pp. 25–35.

Remco, C.V., Mirela, T., 2000. Content-based image retrieval systems: a survey. Tech. Report, Department of Computing Science, Utrecht University.

Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R., 2000. Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Machine Intell. 22 (12), 1349–1380.

Swain, M.J., Ballard, D.H., 1991. Color indexing. Internat. J. CV 7 (1), 11–32.

Tao, D.C., Li, X.L., Yuan, Y., Yu, N.H., Liu, Z.K., Tang, X.O., 2002a. A set of novel textural features based on 3D co-occurrence matrix for content-based image retrieval. In: IEE Internat. Conf. Information Fusion, pp. 1403–1407.

Tao, D.C., Li, X.L., Yuan, Y., Yu, N.H., Liu, Z.K., Tang, X.O., 2002b. Color image semantic segmentation using modified JSEG algorithm. In: IEEE Internat. Conf. Methods & Models in Automation & Robotics, pp. 535–539.

Wang, J.Z., Li, J., Wiederhold, G., 2001. SIMPLIcity: semantics-sensitive integrated matching for picture libraries. IEEE Trans. Pattern Anal. Machine Intell. 23 (8), 947–963.

Yuan, Y., Yu, N.H., Li, X.L., Tao, D.C., Liu, Z.K., 2002. Supervised clustering algorithm based visual information features classification. SPIE Internat. Conf. Image Graphics 4875 (1), 614–618.