http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 11 (5): 626-631, 2012
ISSN 1812-5638 / DOI: 10.3923/itj.2012.626.631

Image Retrieval Based on Topological Features of Gray-level Co-occurrence Networks

[1]Zheng-Liang Huang, [2]Shi-Ze Guo and [1]Zhe-Ming Lu
[1]School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China
[2]North Electronic Systems Engineering Corporation, Beijing 100083, China

Abstract: This study presents to construct special networks from images and then apply their topological properties to image retrieval. Each input color image is divided into three separate gray-level images in the RGB space. For each gray-level image, we view the 256 gray-levels as nodes and construct the Horizontal Gray-level Co-occurrence Network (HGCN) and Vertical Gray-level Co-occurrence Network (VGCN) by counting the number of horizontal and vertical occurrences for each possible gray-level pair. Based on the obtained six directed weighted networks HGCN_R/G/B and VGCN_R/G/B, we extract their topological features including in-degrees, out-degrees, in-strengths and out-strengths for image retrieval. Simulation results demonstrate the superiority of our features to some existing features in terms of P-R curve.

Key words: Image retrieval, complex network, horizontal gray-level co-occurrence network, vertical gray-level co-occurrence network, topological features

INTRODUCTION

Content-based Image Retrieval (CBIR) is any technology that in principle helps us to organize digital image archives based on their visual content. Thus, the research ranging from image similarities to robust image annotation techniques falls under the field of CBIR. People from all kinds of fields, such as computer vision, machine learning, information retrieval, human-computer interaction, database systems, data mining, information theory, statistics and psychology are contributing and becoming part of the CBIR community. In the image retrieval field, the important parameter to evaluate the user-system interaction level is the complexity of queries provided by the system. From the users' perspective, various querying modalities, such as keywords, free-text, images, graphics or their composition can be used to query a system. These query modalities need different processing methods or support for user interaction. Query processing methods can be classified into text-based, content-based, composite processing based, interactive-simple based, interactive-composite based methods. Processing text-based queries involves keyword matching using simple set-theoretic operations and therefore a response can be generated very fast. However, in huge systems working with millions of images and keywords, efficient indexing methods are indispensable. Composite querying methods provide the users with more flexibility for expressing themselves. Interaction-based querying paradigms statistically model the user's interest or help the user refine his queries by providing cues and hints. This study focuses on content-based methods. CBIR methods generally index an image based on its actual contents such as color (Konstantinidis *et al*., 2005), texture (Hejazi and Ho, 2007) and shape (Zhang and Lu, 2002), rather than the metadata such as keywords, tags and descriptions associated with it (Huang and Lee, 2003; Zhang *et al*., 2008; Shanthi and Nadarajan, 2009).

Recently, several image coding schemes such as Vector Quantization (VQ) (Lv and Lu, 2011a, b) and Block Truncation Coding (BTC) (Li *et al*., 2011) have been applied to image retrieval for the compressed code of an image is essentially its compact description. VQ is a classical quantization technique for signal processing that models probability density functions by the distribution of prototype vectors. It was originally applied for data compression. Its main idea is to divide a large set of points (vectors) into groups owning approximately the same number of points closest to them. Each group is represented by its centroid as in k-means and some other clustering algorithms. BTC is a type of lossy image compression technique for grayscale images. It divides the original image into blocks and then uses a quantizer to reduce the number of gray levels in each block while maintaining the same mean and standard deviation. Reference (Panchanathan and Huang, 1999) extracted features from the index table of the spatial-domain VQ compressed image. Uchiyama *et al.*

Corresponding Author: Zhe-Ming Lu, Teaching Building 5, Yuquan Campus, Zhejiang University, 38 ZheDa Road, Hangzhou 310027, Pepol's Republic of China

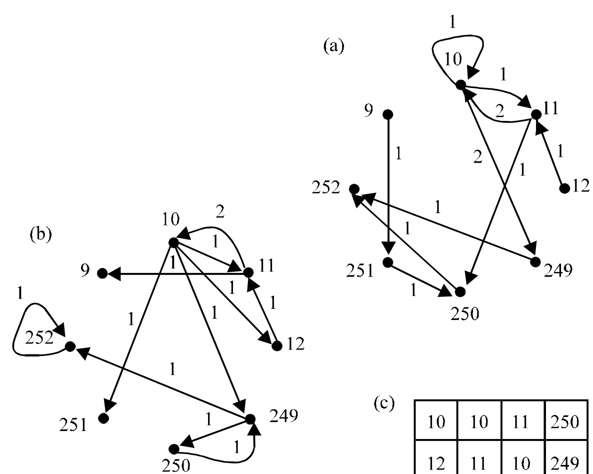
(2001) extracted features from the individual VQ codebook generated from the image. Idris and Panchanathan (1995) uses the index histogram of VQ-compressed index sequence to describe the features which are denoted as VQ Index Histograms (VQIH) in this paper. Lu and Burkhardt (2005) extracted DCT-domain VQ index histograms from the image in the YCbCr color space. Zheng *et al.* (2006) proposed two new compressed-domain features for color image retrieval based on the YCbCr color space. They are named Multi-Stage Vector Quantization Index Histograms and Mean-Removed Vector Quantization Index Histograms. Qiu (2003) extracted a block color co-occurrence matrix and a block pattern histogram from the single bitplane BTC compressed image in the RGB color space. Gahroudi and Sarshar (2007) extracted three block pattern histograms from the three-bitplane BTC compressed image in the RGB color space. Yu *et al.* (2011) extracted two pattern co-occurrence matrices from the BTC compressed Y image and VQ compressed Cb and Cr images, respectively.

This study has proposed a new train of thought to extract image features from networks or graphs that are constructed from the input image based on a simple mechanism. Our idea comes from the popular complex network research area. Complex network is a young and booming research area inspired mainly by the empirical study of computer networks and social networks. Most social, biological and technological networks display substantial non-trivial topological features, with patterns of connection between their elements that are neither purely regular nor purely random. Two most well-known and intensively studied types of complex networks are scale-free networks (Barabasi and Albert, 1999) and small-world networks (Watts and Strogatz, 1998). The former type is characterized by power-law degree distributions and the latter one is in possession of short path lengths and high clustering coefficients. The study of complex networks has attracted researchers from various areas. Since the topological features can be used to describe a network, if we can construct a special network from each image, then we can use its topological features to describe the image and thus we can adopt this feature to search similar images in the image database.

GRAY-LEVEL CO-OCCURRENCE NETWORK CONSTRUCTION

The first key step in our image retrieval system is to construct six gray-level co-occurrence networks (GCNs) from each input color image in the RGB space. Before describing the construction process, we give some knowledge of graph (or network). A network G is composed of a set of nodes $V = \{v_1, v_2, \ldots, v_N\}$ joined by edges $E = \{e_{ij}\}$, where, e_{ij} denotes the edge from node v_i to node v_j. In many real-world networks, edges are often associated with weights $W = \{w_{ij}\}$ that differentiate them in terms of their strength, intensity or capacity. We call these networks weighted networks. On the other hand, many real-life networks are directed networks composed of directed edges (i.e., arcs). If a network is both weighted and directed, we call it a directed weighted network. Next we turn to our directed weighted network construction problem.

Without loss of generality, we take the R image for example to describe the construction process of horizontal and vertical GCNs. In our network, we take the gray-levels as the network nodes. Since the R image is a 256 gray-level image, thus the number of nodes in each GCN is 256 and each node is labeled with a gray-level $a \in (0,255)$. Next problem is how to associate these nodes, i.e., which mechanism is adopted to connect each pair of nodes with a directed weighted link. In this study, for each possible gray-level pair (a, b), $0 \le a \le 255$, $0 \le b \le 255$, we scan the whole R image row by row to count the number of occurrences of (a, b) as a horizontal neighboring pixel pair and we denote the obtained number as w_{ab}. If $w_{ab}>0$, we construct a directed link (called arc) from Node a to Node b with w_{ab} being its weight. Otherwise, there is no link between Node a and Node b. Based on the above simple mechanism, we can obtain the horizontal GCN (HGCN). Similarly, by scanning the whole R image column by column to count the number of occurrences of (a, b) as a vertical neighboring pixel pair, we can obtain the vertical GCN (VGCN). Figure 1 shows an example of HGCN and



(c)

10	10	11	250
12	11	10	249
11	10	249	252
9	259	250	252

Fig. 1: Example of HGCN and VGCN constructed from a small image. (a) A small example image (b) The corresponding HGCN (c) The corresponding VGCN

VGCN that are constructed from a small image. For example, since there are two pairs of (11,10) and one pair of (10,11) in the horizontal direction, thus the incoming weight to Node "11" from Node "10" is 1 and the outgoing weight from Node "11" to Node "10" is 2 in the HGCN. Similarly, since there are one pair of (249, 250) and one pair of (250, 249) in the vertical direction, thus the incoming weight to Node "249" from Node "250" is 1 and the outgoing weight from Node "249" to Node "250" is 1 in the VGCN.

THE PROPOSED SCHEME

The second key step in our image retrieval system is to compute the topological features from each GCN. For undirected unweighted networks, three measures can be used to characterize their features. Degree distribution is the first important topological feature. The degree of a node is defined as the number of edges incident from it. Degree distribution P (k) is defined as the probability that a randomly selected node is with degree k. The degree distribution is very important in studying both real networks, such as the Internet and social networks and theoretical networks. The second important feature is the clustering coefficient which is a measure of degree to which nodes in a graph tend to cluster together. The global clustering coefficient is based on triplets of nodes. A triplet is three nodes that are connected by either two (open triplet) or three (closed triplet) undirected ties. A triangle consists of three closed triplets, one centred on each of the nodes. The global clustering coefficient is the number of closed triplets over the total number of triplets. The clustering coefficient can be also obtained by averaging the local clustering coefficients over all the nodes. The local clustering coefficient C_i for Node i with degree k_i is defined as the number of links L_i that actually exist between its nearest neighbors divided by the number of links that could possibly exist between them, i.e., $C_i = 2L_i/[k_i (k_i\text{-}1)]$. The third important feature is called Average Path Length (APL) that is defined as the average distance over all possible pairs of nodes in a connected network.

In fact, GCN is a kind of directed weighted complex network which can be characterized by the distributions of node degrees, node strengths and local clustering coefficients. Given a directed weighted graph G on N nodes, its weight matrix $W = \{w_{ij}\}_{N\times N}$ and its adjacency matrix $A = \{a_{ij}\}_{N\times N}$, where, the entry a_{ij} satisfies: If there exists an arc from the head node i to the tail node j, then $a_{ij} = 1$; otherwise, $a_{ij} = 0$. The number of head nodes directly adjacent to Node i is called the in-degree of Node i, while the number of tail nodes directly adjacent to it is its out-degree. The in-degree is denoted as k_i^{in} and the out-degree as k_i^{out}. Furthermore, the sum of arc weights inbound to Node i is called the in-strength of Node i and the sum of arc weights outbound from it is its out-strength. The in-strength is denoted as s_i^{in} and the out-strength as s_i^{out}. Based on above definitions, we have:

$$k_i^{in} = \sum_{j=0}^{N-1} a_{ji}, \qquad k_i^{out} = \sum_{j=0}^{N-1} a_{ij} \tag{1}$$

and:

$$s_i^{in} = \sum_{j=0}^{N-1} a_{ji} w_{ji}, \; s_i^{out} = \sum_{j=0}^{N-1} a_{ij} w_{ij} \tag{2}$$

Thus, for each GCN, we can obtain four sequences:

$$K^{in} = \{k_0^{in}, k_1^{in}, \ldots, k_{255}^{in}\}, \;\; K^{out} = \{k_0^{out}, k_1^{out}, \ldots, k_{255}^{out}\} \tag{3}$$

and:

$$S^{in} = \{s_0^{in}, s_1^{in}, \ldots, s_{255}^{in}\}, \; S^{out} = \{s_0^{out}, s_1^{out}, \ldots, s_{255}^{out}\} \tag{4}$$

Based on these four sequences, we can obtain four 32 dimensional normalized vectors as follows:

$$\begin{aligned} F_1 &= \{\overline{k}_{0\to 7}^{in}, \overline{k}_{8\to 15}^{in}, \ldots, \overline{k}_{248\to 255}^{in}\} \\ F_2 &= \{\overline{k}_{0\to 7}^{out}, \overline{k}_{8\to 15}^{out}, \ldots, \overline{k}_{248\to 255}^{out}\} \\ F_3 &= \{\overline{s}_{0\to 7}^{in}, \overline{s}_{8\to 15}^{in}, \ldots, \overline{s}_{248\to 255}^{in}\} \\ F_4 &= \{\overline{s}_{0\to 7}^{out}, \overline{s}_{8\to 15}^{out}, \ldots, \overline{s}_{248\to 255}^{out}\} \end{aligned} \tag{5}$$

Where:

$$\overline{k}_{n\to(n+7)}^{in(out)} = \left(\sum\nolimits_{i=n}^{n+7} k_i^{in(out)}\right)/(8\cdot k_{max}) \tag{6}$$

and:

$$\overline{s}_{n\to(n+7)}^{in(out)} = \left(\sum\nolimits_{i=n}^{n+7} s_i^{in(out)}\right)/(8\cdot s_{max}) \tag{7}$$

Here, $k_{max} = 256$ and s_{max} equals the image size. Since there are six GCNs generated from each input color image, we can obtain 24 normalized vectors and join them together into the final 768 dimensional feature vector of the input image.

EXPERIMENTAL RESULTS AND DISCUSSION

To demonstrate the effectiveness of the proposed feature, we compare our feature with the traditional color-histogram-based (CH) feature and VQ index histogram

(Idris and Panchanathan, 1995). We adopt the standard database in the experiment that is carried out on a Pentium IV computer with the 2.80 GHz CPU. This database consists of 1000 JPEG images of size 384×256 or 256×384 which are categorized into ten categories including people, beach, building, bus, dinosaur, elephant, flower, horse, mountain and food, each containing 100 images. Figure 2 shows 10 typical images from this database:

$$\begin{aligned} \text{Precision} &= \frac{\text{No. of relevant images}}{\text{No. of images returned}} \\ \text{Recall} &= \frac{\text{No. of relevant images}}{100} \end{aligned} \quad (8)$$

For the traditional CH feature, we extract three 256-bin color histograms (i.e., the feature vector is also 768-dimensional) from each image in the RGB color space. For each image, the VQIH feature extraction process for each color component can be illustrated as follows: (1) Divide the image into blocks with the same size of each codeword, typically 4×4. (2) For each block, find its nearest neighbor codeword in the codebook and assign the codeword's index to this block. (3) Calculate the index histogram from the obtained codeword index sequence, with the number of bins being the number of codewords in the codebook. Here the number of bins is selected to be 256. Because there are three color components, we can in total get a 768-dimensional feature vector for each image. To compare the performance more reasonably, we randomly select 5 images from each class and thus in total 50 images, as the test query images. For each test query image, we perform the retrieval process based on the proposed feature, the traditional CH feature and the VQ index histogram respectively. For each number of returned images (from 1 to 1000), we average the recall and precision value over 50 test query images. Here, the precision and recall are defined as follows.

Figure 3 compares the average P-R curves among the proposed feature, the conventional CH feature and VQ index histogram (Idris and Panchanathan, 1995). From Fig. 3, we can see that our feature is much better than the traditional CH feature, because our feature can reflect the relationships between neighboring pixels. We can also see that our proposed feature can achieve a bit better performance than the VQIH feature.

To have an intuitive comparison of the retrieval performance between the proposed feature and the CH feature, we also give a query example as shown in Fig. 4,

Fig. 2: Ten typical images from the test database

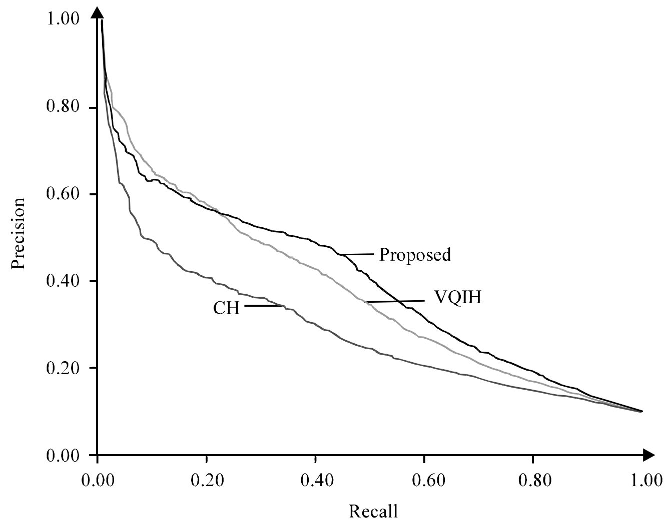


Fig. 3: Comparisons of P-R curves among various features

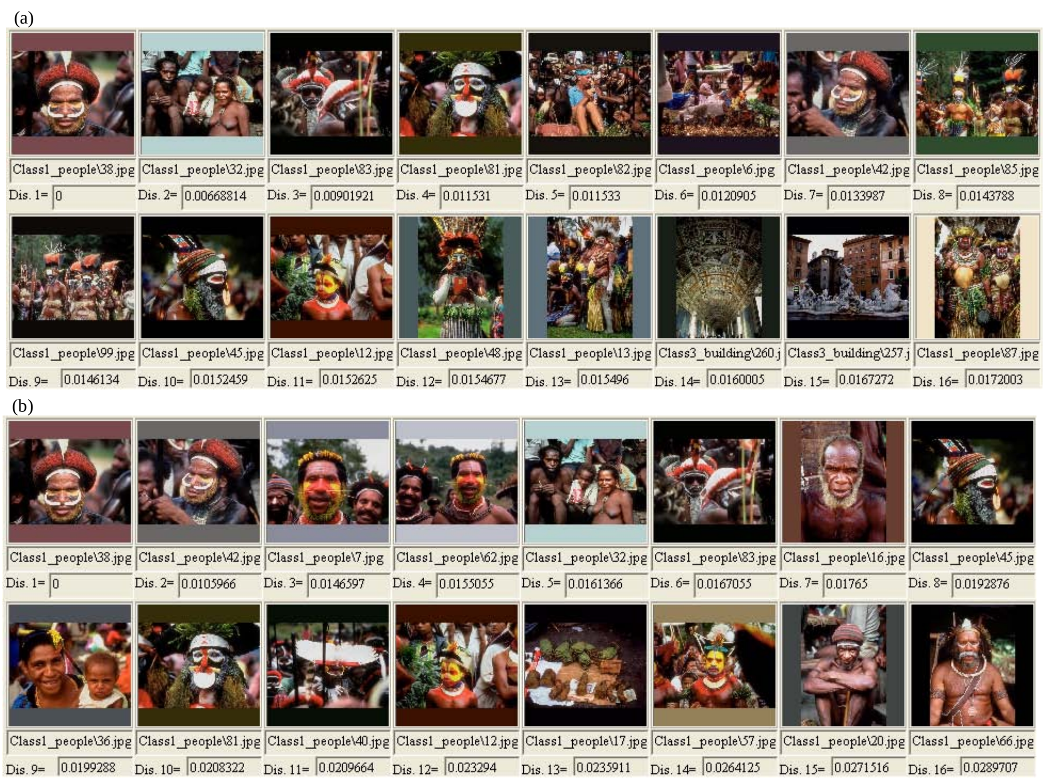


Fig. 4 (a-b): Real query examples based on different kinds of features. (a) Retrieval results based on the CH feature and (b) Retrieval results based on the proposed feature

where the query image is the first image in the returned results. Figure 4a shows the first 16 retrieved images based on the CH feature, while Fig. 4b shows the first 16 retrieved images based on the proposed features. We can see that, two images in Fig. 4a are not from the same class, while all images in Fig. 4b are from the same class. Therefore, we can know that the proposed feature can better represents the image content.

CONCLUSIONS

In this sudy, we present a new type of feature for color image retrieval. This new kind of feature is extracted from two networks (HGCN and VGCN) that are constructed from each image. The network construction process is very simple by scanning all possible pixel pairs in horizontal (row by row) and vertical (column by column) directions in each image respectively. We extract the in-degree, out-degree, in-strength and out-strength histograms from HGCN and VGCN as the final feature for retrieval. This feature can reflect the relationship between neighboring pixels, such that it can obtain much better retrieval performance than the traditional color histogram. Experimental results also show that the proposed feature is better than the VQIH feature. Future work will concentrate on other topological features or spectrum properties of GCNs to further improve the retrieval performance.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the National Natural Scientific Foundation of China under Grant No. 61003255 and No. 61070208. These research projects were conducted in China from January 2011 to December 2013.The authors also gratefully thank Prof. Jia Li for providing the image database for research.

REFERENCES

Barabasi, A.L. and R. Albert, 1999. Emergence of scaling in random networks. J. Sci., 286: 509-512.

Gahroudi, M.R. and M.R. Sarshar, 2007. Image retrieval based on texture and color method in BTC-VQ compressed domain. Proceedings of The 9th International Symposium on Signal Processing and its Applications, February 12-15, 2007, Sharjah, pp: 1-4.

Hejazi, M.R. and Y.S. Ho, 2007. An efficient approach to texture-based image retrieval. Int. J. Imaging Sys. Technol., 17: 295-302.

Huang, P.W. and C.H. Lee, 2003. An efficient method of organizing bit-string signatures for searching symbolic images. Inform. Technol. J., 2: 159-172.

Idris, F. and S. Panchanathan, 1995. Storage and retrieval of compressed images. IEEE Trans. Consum. Electron., 41: 937-941.

Konstantinidis, K., A. Gasteratos and I. Andreadis, 2005. Image retrieval based on fuzzy color histogram processing. Opt. Commun., 248: 375-386.

Li, C.H. Z.M. Lu and Y.X. Su, 2011. Reversible data hiding for Btc-compressed Images based on bitplane flipping and histogram shifting of mean tables. Inform. Technol. J., 10: 1421-1426.

Lu, Z.M. and H. Burkhardt, 2005. Colour image retrieval based on DCT-domain vector quantisation index histograms. Electron. Lett., 41: 956-957.

Lv, M.L. and Z.M. Lu, 2011a. An image hashing scheme based on mean-removed vector quantization for multiple purposes. Inform. Technol. J., 10: 120-126.

Lv, M.L. and Z.M. Lu, 2011b. Multipurpose perceptual image hashing based on block truncation coding. Inform. Technol. J., 10: 207-212.

Panchanathan, S. and C. Huang, 1999. Indexing and retrieval of color images using vector quantization. SPIE Proc. Appl. Digital Image Process., 3808: 558-568.

Qiu, G., 2003. Color image indexing using BTC. IEEE Trans. Image Process., 12: 93-101.

Shanthi, I.E. and R. Nadarajan, 2009. An index structure for fast query retrieval in object oriented databases using signature weight declustering. Inform. Technol. J., 8: 275-283.

Uchiyama, T., M. Yamaguchi, N. Ohyama, N. Mukawa and H. Kaneko, 2001. Multispectral image retrieval using vector quantization. Proc. IEEE Int. Conf. Image Process., 1: 30-33.

Watts, D.J. and S.H. Strogatz, 1998. Collective dynamics of' small-world' networks. Nature, 393: 440-442.

Yu, F.X., H. Luo and Z.M. Lu, 2011. Colour image retrieval using pattern co-occurrence matrices based on BTC and VQ. IET Elect. Lett., 47: 100-101.

Zhang, D. and G. Lu, 2002. Shape-based image retrieval using generic Fourier descriptor. Signal Process. Image Commun., 17: 825-848.

Zhang, M., J. Shen and Z. Shang, 2008. A novel unsupervised color image fuzzy classification algorithm with salient regions. Inform. Technol. J., 7: 560-569.

Zheng, W.M., Z.M. Lu and H. Burkhardt, 2006. Color image retrieval schemes using index histograms based on various spatial-domain vector quantizers. Int. J. Innovative Comp. Inform. Control, 2: 1317-1326.