DOI: 10.2298/CSIS120122047C

A Novel Content Based Image Retrieval System using K-means/KNN with Feature Extraction

Ray-I Chang [1], Shu-Yu Lin [1,2], Jan-Ming Ho [2], Chi-Wen Fann [2], and Yu-Chun Wang [2]

[1] Dept. of Engineering Science and Ocean Engineering
National Taiwan University
Taipei, Taiwan (R.O.C)
{rayichang, d96525009}@ntu.edu.tw
[2] Research Center for Information Technology Innovation
Academia Sinica
Taipei, Taiwan (R.O.C)
{boymike, hoho, fann, zxaustin}@iis.sinica.edu.tw

Abstract. Image retrieval has been popular for several years. There are different system designs for content based image retrieval (CBIR) system. This paper propose a novel system architecture for CBIR system which combines techniques include content-based image and color analysis, as well as data mining techniques. To our best knowledge, this is the first time to propose segmentation and grid module, feature extraction module, K-means and k-nearest neighbor clustering algorithms and bring in the neighborhood module to build the CBIR system. Concept of neighborhood color analysis module which also recognizes the side of every grids of image is first contributed in this paper. The results show the CBIR systems performs well in the training and it also indicates there contains many interested issue to be optimized in the query stage of image retrieval.

Keywords: content based images retrieval; K-means clustering; feature extraction; image retrieval

1. Introduction

Image retrieval is the processing of searching and retrieving images from a huge dataset. As the images grow complex and diverse, retrieval the right images becomes a difficult challenge. For centuries, most of the images retrieval is text-based which means searching is based on those keyword and text generated by human’s creation.[1] The text-based image retrieval systems only concern about the text described by humans, instead of looking into the content of images. Images become a mere replica of what human has seen since birth, and this limits the images retrieval. This may leads to many drawbacks which will be state in related works.

To overcome those drawbacks of text-based image retrieval, content-based images retrieval (CBIR) was introduced [2][3]. With extracting the images features, CBIR perform well than other methods in searching, browsing and content mining etc. The need to extract useful information from the raw data becomes important and widely discussed. Furthermore, clustering technique is usually introduced into CBIR to perform well and easy retrieval. Although many research improve and discuss about those issues, still many difficulties hasn't been solved. The rapid growing images information and complex diversity has build up the bottle neck.

To overcome this dilemma, in this paper, we propose a novel CBIR system with an optimized solution combined to K-means and k-nearest neighbor algorithm (KNN). A creative system flow model, image division and neighborhood color topology, is introduced and designed to increase the clustering accuracy. The rest of this paper is organized as follows. Section 2 briefly describes the concepts of images retrieval, feature extracting, K-means and the structure of CBIR system. In Section 3, we provide a description of the Optimized Content Based Image Retrieval model with K-means and KNN combing with those module proposed in this paper. Experimental results and discussions are presented in Section 4. Finally, Sections 5 discuses our conclusion and future works.

2. Related Work

This chapter introduces some important literatures review in this chapter. First come to content based image retrieval (CBIR). Before CBIR, the traditional image retrieval is usually based on text. Text based image retrieval has been discussed over those years. The text-based retrieval is easy to find out some disadvantages such as:

1. Manually annotation is always involved by human's feeling, situation, etc. which directly results in what is in the images and what is it about.
2. Annotation is never complete.
3. Language and culture difference always cause problems, the same image is usually text out by many different ways.
4. Mistakes such as spelling error or spell difference leads to totally different results.

In order to overcome these drawbacks, content based images retrieval (CBIR) was first introduced by Kato in 1992. The term, CBIR, is widely used for retrieving desired images from a large collection, which is based on extracting the features (such as color, texture and shapes) from images themselves. Content-Based Image Retrieval concerns about the visual properties of image objects rather than textual annotation. And the most popular and directly features is the color feature, which is also applied in this paper. In this work, color is selected as a primary feature in images clustering.

2.1. Data Clustering

Data Clustering is often took as a step for speeding-up image retrieval and improving accuracy especially in large database. In general, data clustering algorithms can be divided into two types: Hierarchical Clustering Algorithms and Non-hierarchical Clustering Algorithms. However, Hierarchical Clustering is not suitable for clustering large quantities of data. Non- hierarchical is suggested to cluster large quantities of data. The most adapted method for non-hierarchical clustering is the K-Means clustering algorithm proposed by James MacQueen in 1967.

K-means clustering algorithm first defined the size of K clusters. Based on the features extracted from the images themselves, K-means allocates those into the nearest cluster. The algorithm calculates and allocates until there is little variation in the movement of feature points in each cluster. This paper applies k-means clustering algorithm for color clustering module based on this concept.

2.2. Feature Extraction

In the image retrieval, feature extraction is the most import issue of the first step. The features extracted from the images directly lead to the results. Some preprocessing is needed to avoid retrieval noise. Steps such as removing the background, highlights the objects. All the preprocessing steps help in feature extraction [18]. Different topic images usually contain different features [19]. Deciding the features needed to be extracted is always popular issues, and then it still comes back to basic features of the images such as:

- Color histograms are basic and fundamental feature of the images. It is the most commonly used and usually gets obvious effect result. Vary literatures apply color histograms as basic images comparison [20] [21].
- Tamura Features. Tamura proposed six images features based on human visualization [22], coarseness, contrast, directionality, line-likeness, regularity, and roughness. Those are usually considered in many related research [3] [2].
- Shape Feature. Besides color, shape is the most commonly used features. Some image retrieval applications require the shape representation to be invariant to rotation, translation, and scaling. Shape representation can be divided into two o categories, boundary-based and region-based [3].
- MPEG-7 Features. The Moving Picture Experts Group (MPEG) defines some description of visual named MPEG-7. It includes Color Layout, Color Structure, Dominant Color, Scalable Color, Edge Histogram, Homogeneous Texture, Texture Browsing, Region-based Shape, Contour-based Shape, Camera Motion, Parametric Motion and Motion Activity features [23].

2.3. Contrast Context Histogram

Contrast Context Histogram (CCH) is first proposed for image descriptor [15].CCH first uses histograms to measure the differences in intensity between various salient points around an object.

Through comparing the positive and negative histograms of the same object, it derives efficient and discriminative descriptors. The CCH method effectively resists image variability between photos of the same object. The CCH extract features in our system as fellow:

First, we assume that each photo has already been analyzed to derive several salient corners. For each center of a salient corner Pc, in the N*N pixel region R, we calculate the center-based contrast C(p) of point p in the region R as follows:

$$C(p) = I(p) - I(p_c) \tag{1}$$

Then, for each central point p in every region Ri, the positive contrast histogram and negative contrast histogram as shown in (2) and (3) respectively:

$$H_{R_i} + (p_c) = \frac{\sum \{C(p) \mid p \in R_i \quad and \quad C(p) \geq 0\}}{\#_{R_i +}} \tag{2}$$

$$H_{R_i} - (p_c) = \frac{\sum \{C(p) \mid p \in R_i \quad and \quad C(p) < 0\}}{\#_{R_i -}} \tag{3}$$

To combine the values of the positive and negative contrast histograms of all regions in Ri, the CCH descriptor of salient corners is defined as:

$$CCH\left(p_c\right) = (H_{R_1 +}, H_{R_1 -}, H_{R_2 +}, H_{R_2 -}, ..., H_{R_t +}, H_{R_t -}) \tag{4}$$

3. Proposed Method

In our CBIR system, it is divided into two parts: learning and querying. The learning step tells about the training process which a huge mount sample images are input in the first step, then the images' features are extracted for the clustering. K-means algorithm is selected to cluster the training data because of it is easy to implement, efficient and well developed in the recent 50 years. Finally the training output the clustering result as a learning code book. The query part describes the images searching process. Inputting the query images and matches to the training result. The output shows the most similar images for user's query. Figure 1 shows the overview of the CBIR system.

The main system architecture contains four modules both in learning stage and query stage: Segmentation and grid module, the K-Means clustering module, the feature extraction module and the neighborhood concept module. The flow chart of the system architecture is shown in Figure 2. Each module is described as follow.

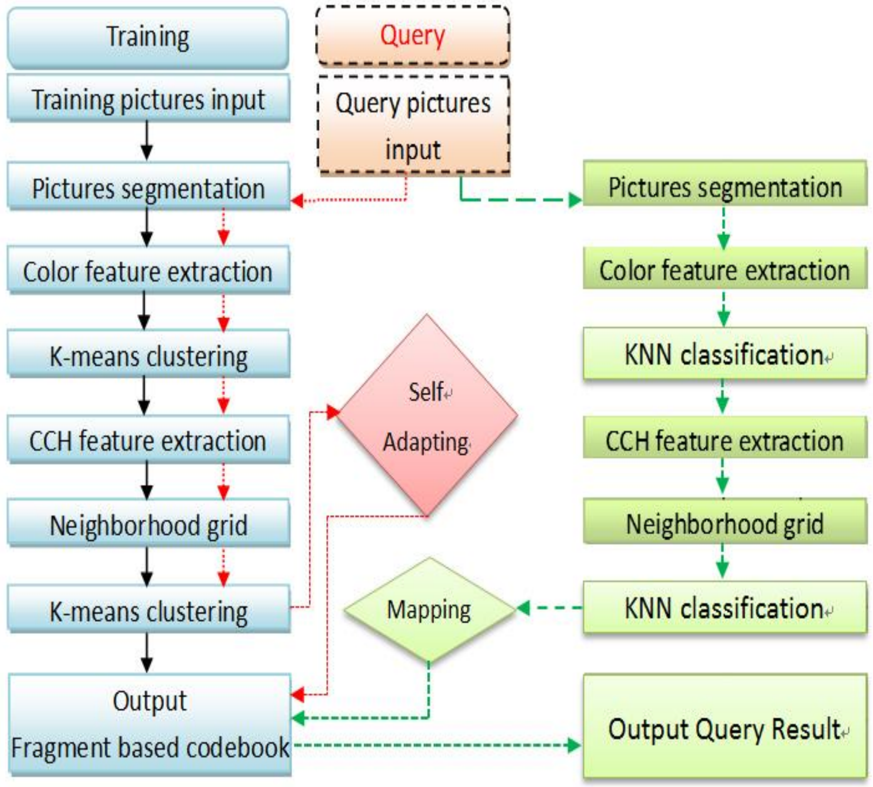


Fig. 1. Overview of the CBIR system

3.1. Image segmentation and grid module

The data images input into the system will be first processed in this module. In the images retrieval, larger images usually decrease the retrieval accuracy. Small images grids help in feature extraction and images processing. Therefore, this module first divides the images into F*F grid and every grid will divided again into S*S sub-grids while during the feature extraction module. In this stage, an input images will finally divided into (F*S)2 grids which the F*F grids uses both in feature extraction module and neighborhood module. The inside S*S grids are only used in color feature extraction. Figure 3 shows the sample of images segmentation.

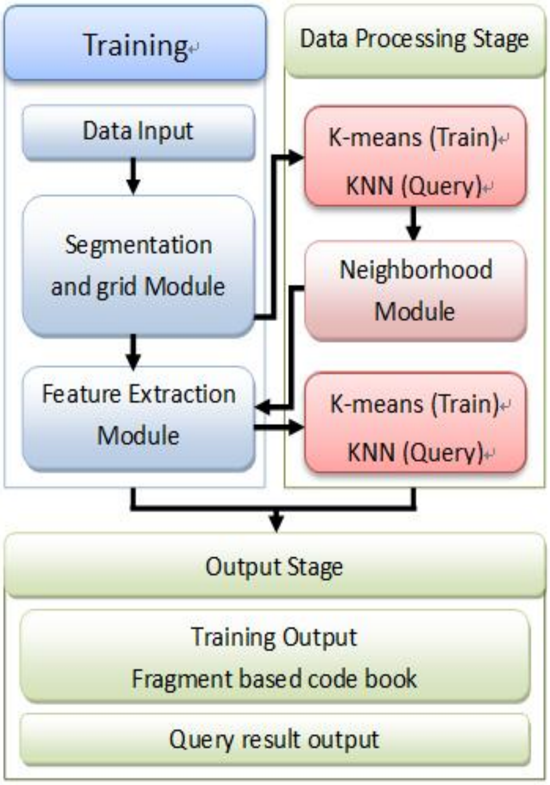


Fig. 2. Flow chart of the system architecture

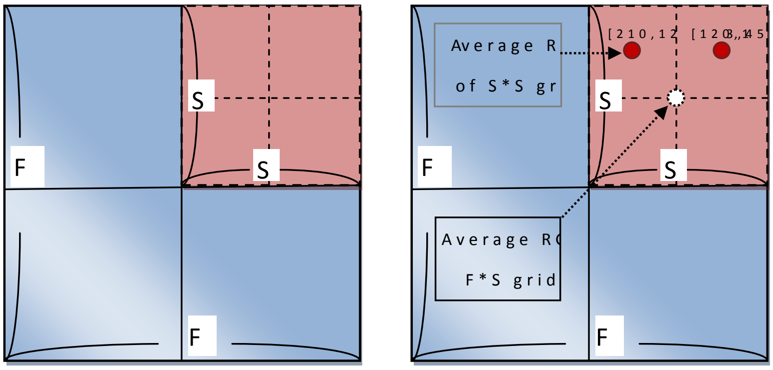


Fig. 3. Sample of Segmentation and grid module and Color feature extraction

3.2. Feature Extraction Module

The input images, including the training and query stage, are all processed in this module. It is also the most important in image retrieval. Since color is the most popular and intuitive feature based on human visualization, it is applied in the system. In order to get more powerful features, the CCH method is also applied for extracting the important feature point. The two feature extractions are described as below:

1. Color Feature Extraction

Input images will be divided into F*S grids before this stage. All grids are input to extract the color feature. First, the module compute the average RGB value of the F*F grids. Second, the inside S*S grids in every F*S grids will also be input to calculate the average RGB value. The S*S grids' detail RGB information is append after the F*S grids' color feature information. All those are prepared for first K-means clustering. Figure 3 illustrates the color feature extractions of this stage.

2. CCH feature extraction

The system utilizes CCH to find out the important feature points. All the points are detected for preparing the input data of the neighborhood module and K-means clustering or KNN classfying. The information of CCH feature points, including the 64 dimensions data, combines with the neighborhood module result. Taking it as the input for the second round K-means clustering, the K-means clustering results in a fragment-based database, call the Code book. As the same implementation in query step, K-means is replaced by KNN algorithm. Query data inputed will be classfied to improve the training code book, also correct classfied result helps for quicly retrieval. Figure 5 shows the sample for CCH feature extraction.

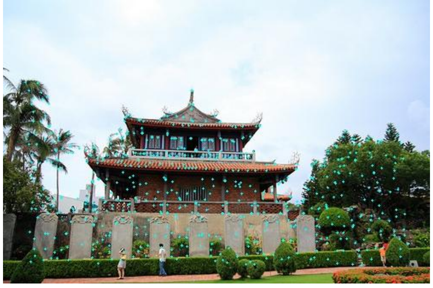


Fig. 4. Data presentation and processing of the neighborhood module

3.3. The neighborhood module

In this module, the input data is from the CCH feature points. Feature points of every image will consider as an index to build up the neighborhood table. Also the first K-means clustering result of every F*F grids are imported to denote the value of the neighborhood table. The steps of every detail are described as follow:

1. Input the CCH feature point Y of the X picture, represented as PICXY.
2. Get the first K-means clustering result based on the CCH feature point's coordinate.
3. Get the neighborhoods' first K-means clustering results.
4. Appending the results from step3 according with the order left to right then top to bottom. If there is no neighborhood, then the value will be "0", which stands for the side of the pictures.
5. Appending the CCH information into the neighborhood table.
6. K-means clustering based on the neighborhood table to generate the code book.

Figure 5 shows the data presentation and processing of the neighborhood module. Taking the same format, the query step apply KNN algorithm instead of K-means.

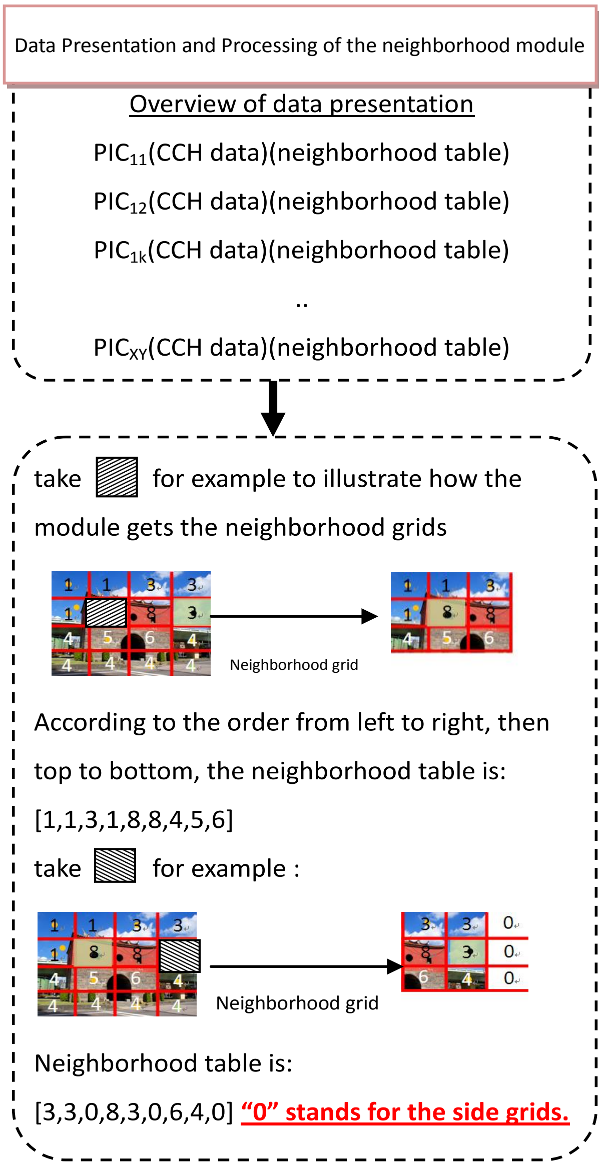
3.4. The K-means/KNN module

K-means clustering is applied twice in our system. The K-means clustering helps generate the code book. In order to keep those input (include the training and query stage) being clustered in the same standard, our architecture keeps the cluster central points in the training stage. The central points are imported into the K-means clustering in the query stage. Finally, the code book can be mapped in the same comparing standard.

K-nearest neighbor algorithm (KNN) is also involved into our CBIR system. Based on the training result, KNN is applied for the query data images. KNN helps to classify the input data; also it fixes the code book which means the training result can be self-adapted.

3.5. The image retrieval query module

This paper provides a query method based on modules mentioned before. Considering the grid fragment, color feature, and CCH feature points, all images input for query will be divided into pieces. Then KNN is applied to classify those images which maps to the training result (code book). The query grid images are compared with those picture grids in the same cluster, the system then calculates the difference based on the color feature which is introduced in 3.2. All fragments are tagged and linked with one grid in the code book. By calculate the most amount of the grids, the CBIR finally output the query and retrieval result.



Data Presentation and Processing of the neighborhood module

Overview of data presentation

PIC_{11}(CCH data)(neighborhood table)

PIC_{12}(CCH data)(neighborhood table)

PIC_{1k}(CCH data)(neighborhood table)

..

PIC_{XY}(CCH data)(neighborhood table)

take ▨ for example to illustrate how the module gets the neighborhood grids

According to the order from left to right, then top to bottom, the neighborhood table is: [1,1,3,1,8,8,4,5,6]

take ▧ for example :

Neighborhood table is:

[3,3,0,8,3,0,6,4,0] **"0" stands for the side grids.**

Fig. 5. Data presentation and processing of the neighborhood module

4. Implementation and Experimental Results

Based on the method provided in this paper, the experiments are designed to verify the architecture of the CBIR system. Also the experiments shows the modules proposed in this article perform good and well organized with the CBIR system architecture.

First the CBIR system splits and calculates the average RGB, we have implemented and figure 6 shows samples of the color feature and split result.

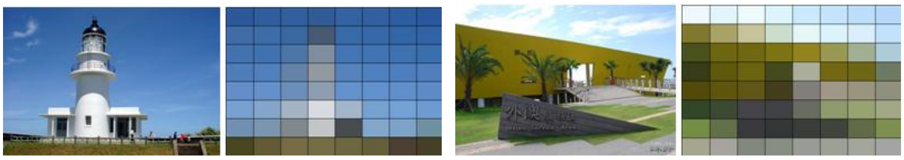


Fig. 6. Samples of the color feature result.

The color features extracted from X pictures divided into Y fragments is denoted as table 1 and pictures training clustered result is visualized as shown in figure 7 and figure 8.

The results show that the same color features of fragments are clustered together. Then the neighborhood module, CCH feature points are combined to be clustered to generate the training result, called the code book. Figure 9 and figure 10 show the visualization of the training results. The features extracted through CCH has been grouped together based on the color features.

In the code book, the color features and the CCH features are included. Based on the fragments, image can be retrieval in detail and the clustering helps decrease the computing cost.

Table 1. Representation of color feature

Data Representation			
Picture Number	Fragment Number	Average RGB	Fragment RGB
1	1	[123,213,21]	[12,23,23]….[13,43,234]
1	2	[23,3,23]	[1,43,123]…[5,10,87]
…	…	…	…
X	Y	[8,132,24]	[99,1,24]…[89,255,1]

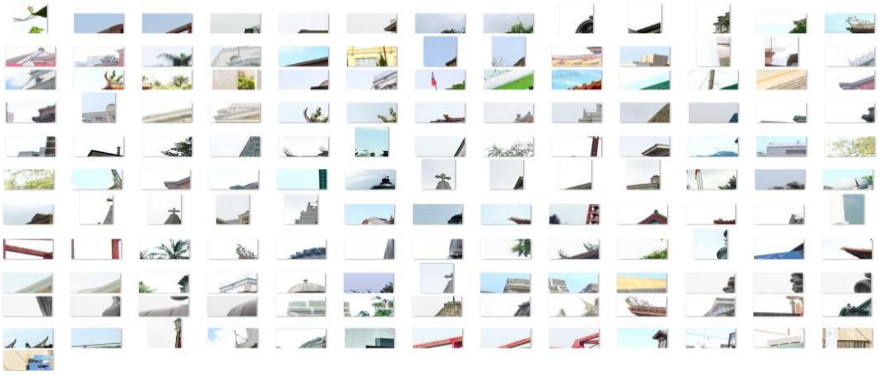


Fig. 7. Visulization of Cluster NO.2 derived by K-Means Color Clustering.

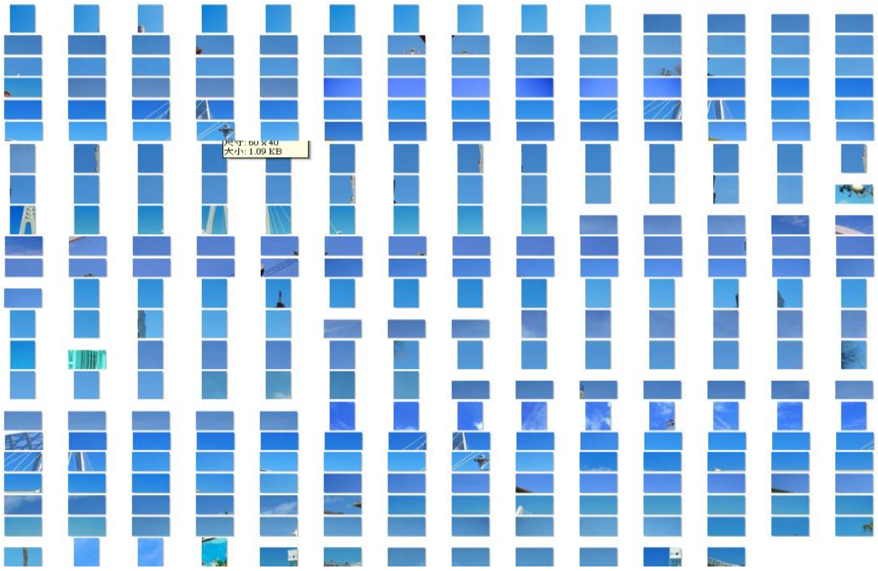


Fig. 8. Cluster NO.6 derived by K-Means Color Clustering.

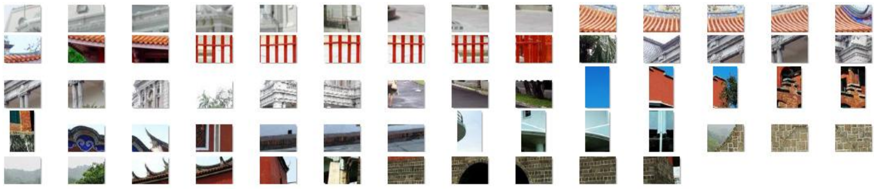


Fig. 9. Cluster NO.15 derived by K-Means in the code book.

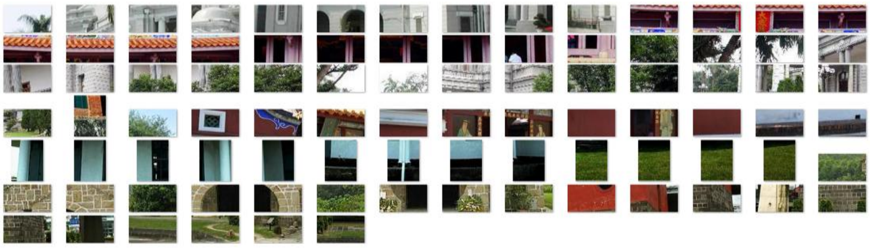
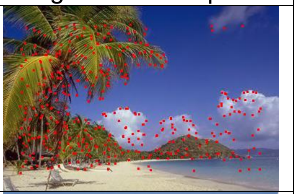
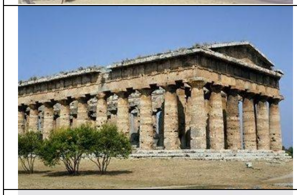
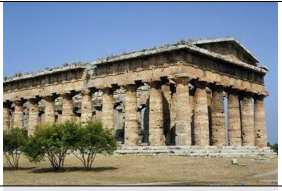
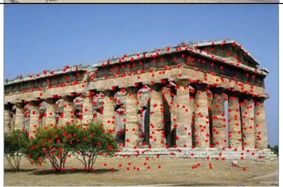
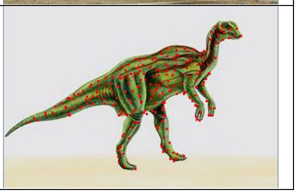


Fig. 10. Cluster NO.13 derived by K-Means in the code book.

In the code book, the color features and the CCH features are included. Based on the fragments, image can be retrieval in detail and the clustering helps decrease the computing cost. This paper has verified several settings for proposed CBIR system architecture. The results indicates the system perform well for training and querying.

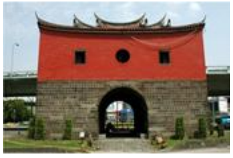
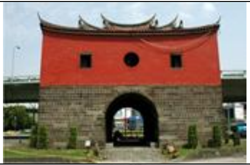
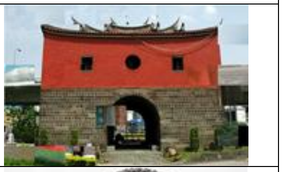
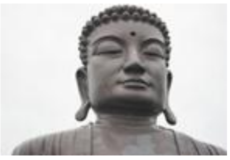
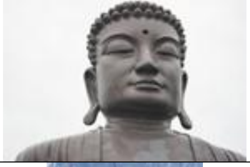
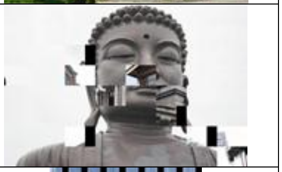
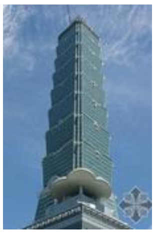
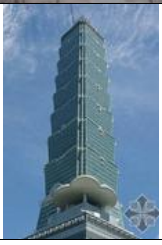
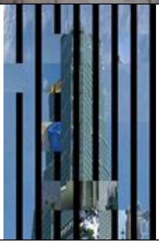
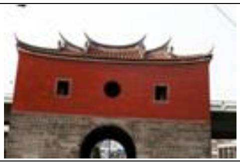
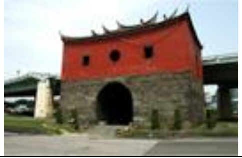
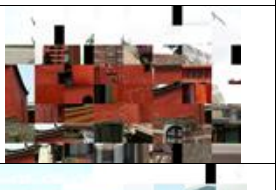
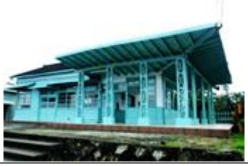
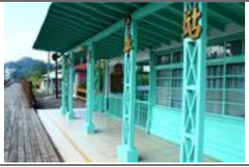
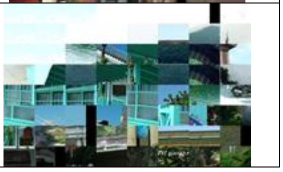
In order to verify the CBIR system, the Wang's dataset which includes 386*254 pixel images is applied as the testing and training data [25]. The 1,000 images from the Wang's dataset are applied into the codebook. 50 images are randomly selected as query images. With the CCH feature points, the CBIR system proposed in this paper successfully retrievals the correct images for those query images. Furthermore, landscaped images are selected from 10,000 images as the training and querying images [25]. The results indicates the CBIR system proposed retrievals correctly when it is well trained. Table 2 shows the query sample of the Wang's dataset.

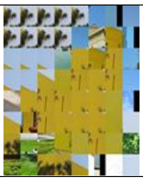
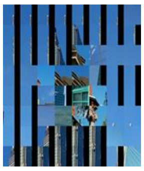
Table 2. Retrieval results for the Wang's dataset.

Images Input	Retrieval Results	Image with CCH points
		
		
		

Several images are input for images retrieval. To query for similar images, the system operates as section 3 described. Table 3 shows the image retrieval results. The CBIR system proposed in this paper successfully find out the pictures which are included in the code book. For those which are not in the code book, the similar pictures are also figured out as table 3 shown. Also the retrieval processes which are generated by the grid module is shown in table 3, the grid puzzle images. The experiment indicates that the CBIR system retrievals the easy-to-tell similar image, even though the image inputted for query is never trained in the codebook.

Table 3. Image retrieval results

Images Input	Retrieval Results	Grid puzzle images
Pictures included in the Code book		
		
		
		
Pictures not included in the Code book		
		
		

Images Input	Retrieval Results	Grid puzzle images
Pictures included in the Code book		
		
		

5. Conclusion and Future Work

The proposed system is designed to operate the content based image retrieval system. It has been verified with the photos of places of interest in Taiwan and the Wang's dataset [24][25]. Our experimental results demonstrate that our CBIR system architecture not only works well for image retrieval, but also improves its precision.

In our knowledge, this paper first combines segmentation and grid module, feature extraction module, K-means clustering and neighborhood module to build the CBIR system. Furthermore, the concept of neighborhood module which recognizes the side of every grids of image is first contributed in this paper. Applying the concept of fragment based code book into the content based image retrieval system also contributes in our system architecture. The experimental results confirm that the proposed CBIR system architecture attains better solution for image retrieval. Our model represents the first time in which combine new modules and techniques proposed in the paper have been integrated with CBIR system.

Images can be retrieval correctly through the proposed CBIR system. For those images which are contained in the code book, all of them can be searched as the most similar result. Also for general images selected randomly, the query results are similar to the input data. Since the CBIR system is based on the color feature, the retrieval results are directly and easy to tell the performances. In the future work, we hope to build a generalized query method which increase the system searching ability and provide more accurate content descriptions of places of interest places by performing color feature analysis and CCH image extraction simultaneously. As a result, the CBIR system will be able to suggest more relevant annotations and descriptions.

Furthermore, we hope to optimize the system architecture and modules proposed in this paper. There exists some detail setting can be discussed and optimized with the images retrieval issues.

Acknowledgment. This research was supported in part by Taiwan E-learning and Digital Archives Programs (TELDAP) sponsored by the National Science Council of Taiwan under NSC Grants: NSC101-2631-H-001-005, NSC101-2631-H-001-006.

References

1. Kato,T.:Database architecture for content-based image retrieval. Image Storage and Retrieval Systems, 112–123. (1999)
2. Datta, R., Joshi, D., Li, J., Wang, J. Z.: Image Retrieval: Ideas, Influences, and Trends of the New Age. ACM Computing Surveys, Vol. 40, No. 2, Article 5, April. (2008)
3. Rui, Y., Huang, T.S., Chang, S. F.: Image Retrieval: Current Techniques, Promising Directions and Open Issues. Journal of Visual Communication and Image Representation, Transaction on Systems, Man, and Cybernetics, vol. 8, 460–472. (1999)
4. Rui, Y., She, A. C., Huang, T. S.: Modified Fourier Descriptors for Shape Representation - A Practical Approach. Proc. of First International Workshop on Image Databases and Multi Media Search. (1996)
5. Gupta, A., Jain, R.: Visual Information Retrieval. Communications of the ACM, vol. 40, 70–79. (1997)
6. Gross, M. H., Koch, R., Lippert, L.: A. Dreger, Multiscale image texture analysis in wavelet spaces. Proc. IEEE Int. Conf. on Image Proc. (1994)
7. Chua, T. S., Tan, K.-L., Ooi, B. C.: Fast signiture-based color-spatial image retrieval. Proc. IEEE Conf. on Multimedia Computing and Systems. (1997)
8. Chuang, G. C.-H., Kuo, C. -C. J.:Wavelet descriptor of planar curves: Theory and applications. IEEE Trans. Image Proc., vol. 5, 56–70.(1996)
9. Faloutsos, C., Flickner, M., Niblack, W., Petkovic, D., Equitz, W., Barber , R.: Efficient and Effective Querying by Image Content. (1993)
10. Rui, Y., Huang, T.S., Mehrotra, S.: Relevance feedback techniques in interactive content-based image retrieval. In Storage and Retrieval for Image and Video Databases VI, 25–36. (1996)
11. Stricker, M., Orengo, M.: Similarity of color images. Proc. SPIE Storage and Retrieval for Image and Video Databases. (1995)
12. Smith, J. R., Chang, S.-F.: Automated binary texture feature sets for image retrieval. Proc. IEEE Int. Conf. Acoust, Speech, and Signal Proc, May. (1996)
13. Mirmehdi, M., Palmer, P. L., Josef K.: Optimising the Complete Image Feature Extraction Chain. Proceedings of the Third Asian Conference on Computer Vision, Vol. 2, 307–314, January. (1998)
14. Kwak, N., Choi, C.-H., Choi, C.-Y.: Feature extraction using ICA. Proceedings of the International Conference on Artificial Neural Networks, 568–573. (2001)
15. Huang, C. R., Chen, C. S., Chung, P. C.: Contrast context histogram - A discriminating local descriptor for image matching. IEEE International Conference on Pattern Recognition, 53–56. (2006)
16. Yu, H. H., Wolf, W.: Hierarchical, multi-resolution algorithms for dictionary-driven content-based image retrieval. Proc. IEEE Int. Conf. on Image Proc. (1997)
17. Ono, A., Amano, M., Hakaridani, M.: A flexible content-based image retrieval system with combined scene description keyword. Proc. IEEE Conf. on Multimedia Computing and Systems. (1996)

18. Prasad, B. G., Biswas, K. K., Gupta, S. K.: Region-based image retrieval using integrated color, shape, and location index. Comput. Vis. Image Underst. Vol. 94, 193–233. (2004)
19. Deslaers, T., Keysers, D., Ney, H.: Features for image retrieval: an experimental comparison. Information Retrieval, vol. 11, 77–107. (2008)
20. Puzicha, J., Rubner, Y., Tomasi, C., Buhmann, J.: Empirical Evaluation of Dissimilarity Measures for Color and Texture. Proc. Int"l Conf. Computer Vision. (1999)
21. Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1349–1380. (2000)
22. Tamura, H., Mori, S., Yamawaki, T.: Textural features corresponding to visual perception. IEEE Transactions on Systems, Man, and Cybernetics, 460–473. (1978)
23. Eidenberger, H.: How good are the visual MPEG-7 features?. In Proceedings SPIE Visual Communications and Image Processing Conference, Vol. 5150, 476–488. (2003)
24. Jia, L., Wang, J. Z.: Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, 1075–1088. (2003)
25. Wang, J. Z., Jia, L., Gio, W.: SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol 23, 947–963. (2001)

Ray-I Chang joined the Department of Engineering Science, National Taiwan University, in 2003. He has published over 150 original papers, including papers published in journals such as IEEE Transactions on Multimedia, IEEE Transactions on Broadcasting, and IEEE Transactions Neural Networks. His current research interests include multimedia networking and data mining. Dr. Chang is a member of IEEE.

Shu-Yu Lin is a Ph.D candidate in Department of Engineering Science, National Taiwan University, Taiwan, R.O.C. His research interests include artificial intelligence, machine learning, data mining and multimedia.

Jan-Ming Ho received the B.S. degree in Electrical Engineering from National Cheng Kung University, Taiwan, in 1978 and the M.S. degree from Institute of Electronics at National Chiao Tung University, Taiwan, in 1980. He received the Ph.D. degree in Electrical Engineering and Computer Science from Northwestern University, USA, in 1989. He joined the Institute of Information Science, Academia Sinica, Taipei, Taiwan, as an associate research fellow in 1989 and was promoted to research fellow in 1994. He had served as Associate Editor of IEEE Transaction on Multimedia. He was Program Chair of Symposium on Real-time Media Systems, Taipei, 1994 - 1998, General Co-Chair of International Symposium on Multi-Technology Information

Processing, 1997 and will be General Co-Chair of IEEE RTAS 2001. He was also steering committee member of VLSI Design/CAD Symposium, and program committee member of several previous conferences including ICDCS 1999, and IEEE Workshop on Dependable and Real-Time E-Commerce Systems (DARE'98), etc.

Chi-Wen Fann is a research assistance in Institute of Information Science, Academia Sinica. He was an well-experienced project manager. His research interests includes multimemedia system design and data mining.

Yu-Chun Wang is a research assistance in Institute of Information Science, Academia Sinica. His research includs cloud computing, data mining and machine learning.

Received: January 22, 2012; Accepted: December 04, 2012.