Optimized Method of Multi-Feature for Content-based Image Retrieval

Zhengyan Dai[1,a*], Sujuan Qin[2,b]

[1,2]State Key Laboratory of Networking and Switching Technology

Beijing University of Posts and Telecommunications

Beijing, 100876, China

[a]zhengyandai1004@gmail.com, [b]qsujuan@bupt.edu.cn

Keywords: multi-feature, hierarchy clustering, image-retrieval

Abstract. Content-based image retrieval has been focused on attention during recent years. Traditional methods of CBIR most relied on single feature extraction of images, which could only reflect single characters of the image. We proposed an optimized structure of multi-feature for content-based image retrieval. We extracted the color feature using HSV bin to reflect the holistic feature of the image, and then extract the sift descriptor to produce Bag-of-Words bin to reflect the local feature of the image. Using those two features to get fusion vectors represents the image synthetically. Finally, we use hierarchy clustering to cluster images in the database to get efficient retrieval results. In the experiment, we test the precision of our method comparing with the same type of feature extraction combining method to validate our promotion.

Introduction

As the development of mobile Internet, the number of approaches of getting image information is gaining, and image information stored in the Internet becomes more informative, the accuracy and efficiency of handling image information give rise to image retrieval system. Initially, the main approach for image retrieval was Text-based image retrieval (TBIR), supposing images could be represented by the labels put by people or text information surrounding the picture [1]. The traditional image retrieval was mainly based on this theory. However, labeling picture could not satisfy people's demand for the low efficiency confronting with more variety of images in our life, for instance, this old approach could not solve label information in different language. Thus, image retrieval based on image's own feature, such as color, texture, shape, etc. has been focused on more attention recently. The earlier content-based image retrieval (CBIR)'s main process is explained in figure 1. [2] It is the basic structure for most of the image searching process. Combining image's feature and text information, Barrios J M etc. offered a new approach finding the similar image using both image information and text meaning. [3] The main process is showed in figure 2. This approach, however, also could cost much of human's efforts to label the pictures that would reduce the efficiency of the system, even though the accuracy could be improved to some extent. Hence, people have been looking for more efficient method for image retrieval, and the extraction of origin features, namely low-level features, has been the main issue for recent years.
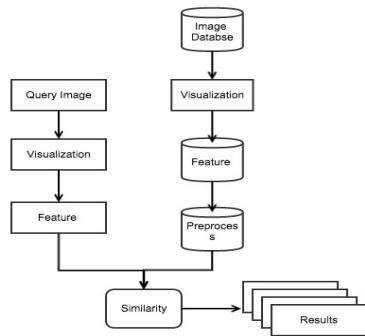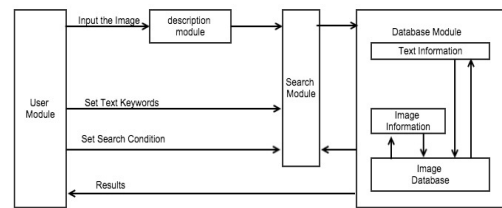
Fig.1, earlier CBIR process

Fig.2, image retrieval process combining text information and image feature

The features of image have great significance on efficiency and accuracy of the image retrieval system, including local features and holistic features. Fusion of different features for image retrieval could improve the quality of the CBIR. Recently, Roshi Choudhary etc. proposed an integrated approach to CBIR, [4] which combines the color and texture information of the image. Lacking appropriate preprocess for image database, the system's efficiency had some flaws. Timothee Cour etc. proposed a method based on graph-fusion, [5] extracting features like HSV, GIST to test the experiment, gaining good results. However, the frequency of inserting and deleting the nodes of graph cost much time.

We proposed an optimized structure of multiple feature fusion for CBIR, testing on the main test image databases. Comparing with the original structure for retrieval, the efficiency and accuracy of the system both improved. The remaining work is described as mentioned: In section 2, related work for feature extraction for picture preprocessing and some clustering method are explained. The proposed architecture is explained in section3, and the experiment and analysis is displayed in section4. In section 5, the work is concluded and some potential implementation is explained.

Related works

Holistic Features. Holistic features are extracted for representing the whole picture's primary information. It could compress the total data that we need match, and keep stable when the picture rotates, resisting the influence by noise, spectator and scale. While it could lose some important information related to local position, which could make pictures more instinctive in some condition. To solve the problem, some novel vectors based on histogram have been proposed, such as coherence vector, correlograms, etc. [6] which all took the information about space into consideration.

Local Features. Local Features are extracted for representing local textures of the image. The local feature could be used to represent the local objects in the image which could be ignored by the holistic features and resist the influence from the light condition, posture and perspective of the sample image. Although local features give a fantastic approach for representing the objects in the image, the number of the feature is large, and the reference and learning about space consumes lots of time. To solve it, experts proposed the method Bag-of-feature, [7] the main process is explained as follow: First, we should detect the key points in images, using specific descriptor to describing those local features. Then clustering those local features for calculation efficiently gets quantified features. At last, using bag-of-features summarizes and unifies those features to a histogram. This approach could stand the little nuance among key points, and separate mostly different key points simultaneously.

Clustering. The clustering aggregates objects sharing similar features, and separating objects with distinctive features. [8] The main approaches of clustering include K-means, hierarchy clustering, SOM etc. and some optimized methods related to those basic approaches. K-means is normally used in clustering. [9] The process of K-means clustering is efficient, and its computational complexity is merely $O(tkn)$, t means how many of iteration, k is the number of clustering, n is the number of the data. While the exact number of classification, the K, should be given before the process but the appropriate number of clustering could not be deemed easily. Hierarchy clustering is more direct, namely clustering one hierarchy by one hierarchy. [10] We could beneficially get the whole clustering process just by one time. According to the clustering tree we could calculate how many cluster we need, it solves the issue we mention above, while we sacrifice more time and space to calculate the clustering tree.

Proposed Method

Considering the limits of using single feature of image to employ the CBIR and low efficiency of traditional matching process, we proposed a structure that combines holistic feature and local feature of the images, then clustering the images with similar features into distinct categories in which could enhance the accuracy and efficiency of the CBIR system.

Feature Extraction. Here we choose HSV histogram to represent the color feature of the images, reflecting the holistic feature of the image, converting the image to HSV mode to extract feature vectors. The range of HSV's 3 elements are $H \in [0,360]$, $S \in [0,255]$, $V \in [0,1]$. As the number of color image's pixels is too large, we quantify the color space of HSV to improve the efficiency. We separate H to 8, S to 12 and V to 3 fractions respectively, to get a 288 dimensional color vector.

Given an image I, m*n, quantifying the color to C_1, C_2, … C_{72}. Supposing the set N(I) = {n_1, n_2, …n_{72}}, in which n_i represents the number of pixels, equals to C_i . The color histogram reflects the number of pixels, proportions to the total number of pixels, namely: $P(pixel \in C_i) = n_i/m*n$.

Sift is commonly used in local feature extraction, which could identify objects from different perspectives and different scenes based on its resistance to scale transition and rotation of image. Firstly, we extract the sift descriptor from each image in the database, clustering those features using K-Means to get visual feature. Then we calculate the feature histogram using sift for each image, storing the feature vectors to feature database. [11] The main steps in extracting SIFT displayed in figure 3.
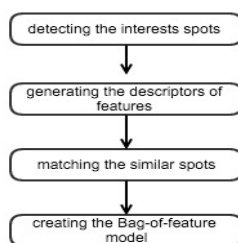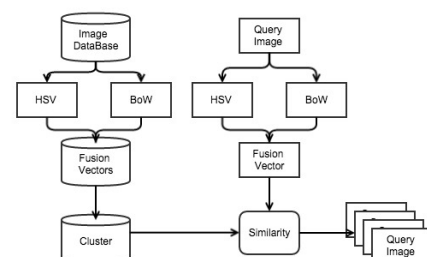

Fig.3, extracting SIFT features

Fig.4, main process of my algorithm

Integrating the features of images. Here, we take hierarchy clustering as an example, which we use in our experiment, to introduce the process of image clustering. [12]

When we deal with the vectors from holistic and local extraction, we use flatten(), a function in OpenCV, to get a fusion vector from the image, and use PCA() to reduce the dimensions of the vectors. Like the process in the figure 4, we loaded all the vectors to one matrix, viewing each vector as a sole cluster. And then we calculate the distance between each two vectors using Euclidean distance, until get anticipating number of clustering. For clustering, the key step is evaluating the similarity between two data, and different approach demands different methods to evaluate the similarity.

Supposing we use the image in the processed image database as the query image, so we could use the distances between the query and other images which clustered the same group with the query image to get the most similar images, without comparing with all other images which could make the retrieval much more efficient.

Experiment and Analysis

We use Corel1000 dataset [13] as our test images, which includes 1000 pictures, containing flowers, mountains, Africans, beach, buildings, food, automobiles, horses, dinosaurs, and elephants.

Here, we use precision to evaluate the quality of our novel method. We set S as image set similar to the Query image in the image dataset, R as the total retrieval image set, s as similar pictures in one retrieval result, u as dissimilar pictures in one retrieval result, the precision could be defined as follow: [14]

$$precision = P(S \mid R) = \frac{P(S \cap R)}{P(R)} = \frac{s}{s+u}$$

To start the experiment, we randomly select several pictures in each category of the Corel dataset, using the retrieval result to calculate the precision. Firstly, we use only color feature, HSV, to test each category pictures and quantify the precision. Secondly, we use local features, Bag-of-Feature, using sift, to calculate the precision. At last, we use the novel method to calculate the precision.

Table 1 displays the mAP of the three method, and the related statistic results shows in figure 2. We could see the novel method get higher precision than method using only one feature.

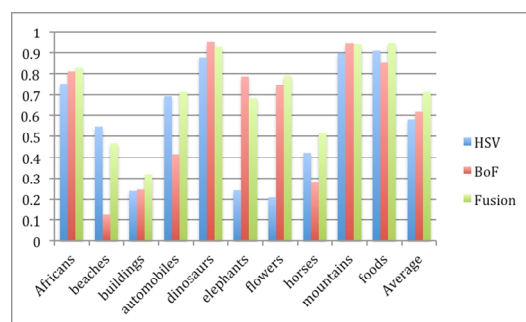|	HSV	BoF	Fusion
Africans	0.753	0.81	0.825
beaches	0.546	0.124	0.467
buildings	0.24	0.247	0.32
automobiles	0.694	0.415	0.712
dinosaurs	0.877	0.95	0.926
elephants	0.242	0.787	0.684
flowers	0.21	0.748	0.791
horses	0.424	0.28	0.513
mountains	0.897	0.943	0.94
foods	0.91	0.85	0.945
Average	0.5798	0.6156	0.7123

Table1.

Fig.5, results compared with single-feature retrieval

We also used our results to compare with the results provided by Roshi Choudhary [4] and

Kommineni Jenni [15]. The result displayed in figure 6 proves our method had improved the average precision rate.
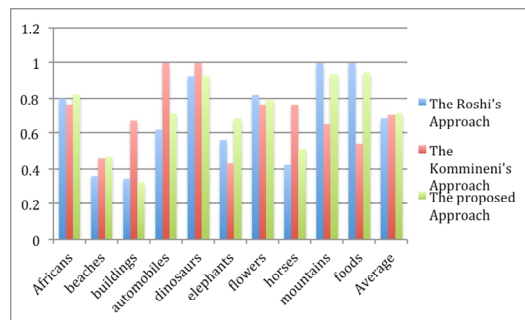

Fig.6, results compared with other methods

Conclusion

In this paper, we proposed a new architecture to do content-based image retrieval, improving the accuracy of the system by considering holistic feature and local feature of the image. Through hierarchy clustering the images in the database, we could save some time spent on getting the results. When we produced the bag-of-feature, there could be some improvement to avoid memory overflow, as images in the database could produce a large number of parameters. So, in the future, we could do more work to deal with the process and also try other feature extraction methods to retrieval images.

Acknowledgements

References

[1] Li, Wen, et al. "Text-based image retrieval using progressive multi-instance learning." Computer Vision (ICCV), 2011 IEEE International Conference on IEEE, 2011:2049-2055.

[2] Santini A G, Jain R. Content based image retrieval at the end of the early years[J]. Pattern Analyis & Machine Intelligence IEEE Transactions on, 2000, 22(12):1349-1380.

[3] Barrios J M, Diaz-Espinoza D, Bustos B. Text-Based and Content-Based Image Retrieval on Flickr: DEMO[J]. Proceedings of International Workshop on Similarity Search & Applications, 2009:156-157.

[4] Choudhary R, Raina N, Chaudhary N, et al. An integrated approach to Content Based Image Retrieval[C]// Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on. IEEE, 2014:2404-2410.

[5] Zhang S, Yang M, Cour T, et al. Query Specific Fusion for Image Retrieval[J]. Lecture Notes in Computer Science, 2012, 37(4):660-673.

[6] Zhang Y S, Zheng F U, Zhu Z Y. Image retrieval based on gradient texton coherence vector[J]. Application Research of Computers, 2012, 29(3):1119-1115.

[7] Liu L, Wang L. Exploring latent class information for image retrieval using the bag-of-feature model.[J]. Schools & Disciplines, 2011:1405-1408.

[8] Chen X T, Wang A Y, Xie W P, et al. Survey of clustering algorithm based on content-based

image retrieval[J]. Application Research of Computers, 2008, 25(12):3546-3549.

[9] Shi X, Xue A. Application study of improved k-means clustering algorithm in image retrieval[J]. Computer Engineering & Applications, 2011.

[10] Thies C, Malik A, Keysers D, et al. Hierarchical feature clustering for content-based retrieval in medical image databases[C]// Medical Imaging 2003. International Society for Optics and Photonics, 2003.

[11] Wu J, Zhao J, Zhu X, et al. Image Retrieval Based on Improved SIFT Features[J]. Microcomputer Applications, 2011.

[12] Yang Y, Xu D, Nie F, et al. Image clustering using local discriminant models and global integration[J]. Image Processing IEEE Transactions on, 2010, 19(10):2761 - 2773.

[13] The Corel 1000 database. http://wang.ist.psu.edu/docs/related/

[14] Singha M, Hemachandran K. Content Based Image Retrieval using Color and Texture[J]. Signal & Image Processing An International Journal, 2012, 3(1):271-273.

[15] Jenni K, Mandala S. Pre-processing image database for efficient Content Based Image Retrieval[C]// Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference onIEEE, 2014:968-972.