IJCSN International Journal of Computer Science and Network, Volume 5, Issue 1, February 2016
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 1.02

Content Based Effective Searching of Images from Large Data Repository

[1] Shine Raj G, [2] Anver Muhammed K M

[1] Assistant Professor, Baselios Mathews II College of Engineering
Sasthamcotta, Kollam, Kerala

[2] System Administrator, Baselios Mathews II College of Engineering
Sasthamcotta, Kollam, Kerala

Abstract - Image searching is one of the most interesting topics in the coming future. Comparing two images and linear search will take a lot of time. Rather than comparing the images, its better to look for similarity search. Feature extraction and bag of words are the most effective way of searching for similarity of images. We have propsed these two approaches in this paper.

Keywords - *CBIR, Feature Extraction, Bag of Words.*

1. Introduction

Over the past few years Content-Based Image Retrieval has become an exciting and in-depth area of research. The relevance of visual information retrieval in many areas such as fashion and design, crime prevention, medicine, law, and science makes this research field one of the important and fastest growing in information technology. Image retrieval has come a long way where it started off with text-based retrieval. However, there are many problems associated with retrieving images based on text such as manual annotation of keywords, differences in perceptions and interpretations, and a few others. Due to this, researchers came up with CBIR where images are retrieved based on automatically derived low-level features (human vision related), middle-level features (objects related), or high-level features (semantic related). Among these features, the low-level features are the most popular due to its simplicity. One of the important low-level features is color as it plays an important role in CBIR due to its robustness to complex background and independent of image size and orientation. Several color descriptors have been approved including number of histogram descriptors and a dominant color descriptor (DCD). MPEG-7 specifies seven color descriptors. It includes dominant colors, scalable color histogram, color structure, color layout, and GoF/GoP color. The early perception in human visual system performs dominant color identification, and eliminates the fine details and colors in small areas. Whatever be the exact color distribution human beings perceive images as a combination of the dominant colors. Dominant color descriptor (DCD) described as in MPEG-7 lacks certain semantic information but provides compact, intuitive semantic information and effectively describes the color distribution in the interested region of an image.

The feature descriptors are defined using two main components: (1) representative colors and (2) the percentage of each color. This method lacks clarity, because dissimilar images which have large percentage of the same dominant color we are searching for will lead to erroneous results. Usually, the dissimilarity occurs when the background color of an image has the largest percentage of dominant color. Due to this so many enhancements are done to the MPEG-7 DCD. A semantic feature is added to the DCD to improve its accuracy in an object-based image retrieval application and it is considered as feature level-based solution to the background dominance problem. In this paper, we are proposing an effective color extraction scheme for the image retrieval from large database.

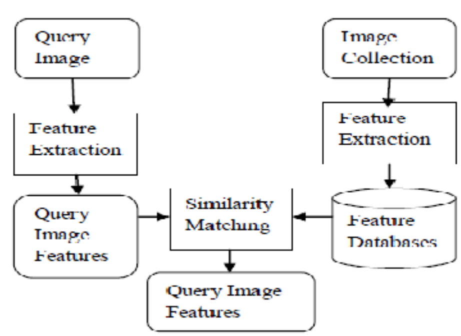

Fig 1: Architecture of the Content Based Retrieval System

The paper is organized in the following way. Section 2 is concerned with explicating content based image retrieval and the similarity measures. Section 3 is mainly concerned with the feature extraction, the newly proposed modification that helps improves the similarity measure and introduces new clustered and indexed database. Section 4 discussed the bag of words approach. Section 5 illustrates the extensive experiment that contains visual results. Finally comes the conclusion in Section 6.

2. Content Based Image Retrieval

Content-based Image Retrieval has witnessed great interest in the last decade. Many techniques have been done with respect to content-based image retrieval (CBIR). But most of the techniques proposed extract low-level features (e.g. color, texture, shapes and layout of objects) to measure the similarities among images, by comparing the feature differences. Color, texture and shape features have been used for describing image content. Color is one of the most widely used low-level visual features and is invariant to image size and orientation. As conventional color features used in CBIR, there are color histogram, color correlogram, and dominant color descriptor (DCD). Color histogram is the most commonly used color representation, but it does not include any spatial information. Color correlogram describes the probability of finding color pairs at a fixed pixel distance and provides spatial information. Therefore color correlogram yields better retrieval accuracy in comparison to color histogram.

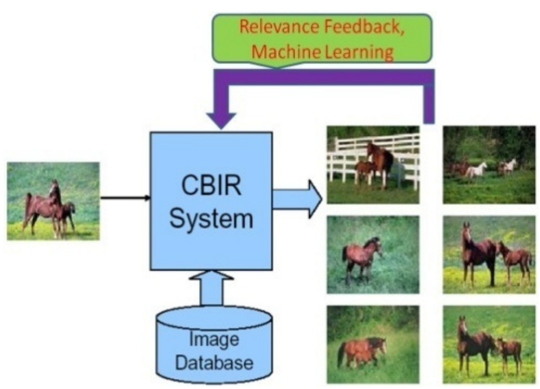

Fig 2: Result of a query image in CBIR.

Color autocorrelogram is a subset of color correlogram, which captures the spatial correlation between identical colors only. Since it provides significant computational benefits over color correlogram, it is more suitable for image retrieval. DCD is MPEG-7 color descriptors. DCD describes the salient color distributions in an image or a region of interest, and provides an effective, compact, and intuitive representation of colors presented in an image. However, DCD similarity matching does not fit human perception very well, and it will cause incorrect ranks for images with similar color distribution. Yang et al. presented a color quantization method for dominant color extraction, called the linear block algorithm (LBA), and it has been shown that LBA is efficient in color quantization and computation. For the purpose of effectively retrieving more similar images from the digital image databases (DBs), uses the color distributions, the mean value and the standard deviation, to represent the global characteristics of the image, and the image bitmap is used to represent the local characteristics of the image for increasing the accuracy of the retrieval system

In other hand, similar to the dynamic quantization-based histogram, MP7DCD also uses QSM with some modification to measure the dissimilarity between the query image and database images. However, QSM is not void of serious drawbacks. For instance, it does not match human color perception. In our work new similarity measure is proposed by the separation of RGB values.

Addition to this, in our present work we will introduce a new clustered and indexed database to increase the performance level and this will help to reduce the image retrieval time complexities. In the former works they used only clustering or indexing approach and not both approach together. As we use both approaches it increases the retrieval performance. Other great advantage we considered in our work is reduced time complexity for the retrieval.

3. Feature Extraction

Feature extraction is the foremost operation when considering Content Based Image Retrieval. The easiest feature in an image to comprehend is perhaps the color. Each pixel in a digital image consist of a color element. Color features are widely used to describe the image. The most common color space used is RGB color space where R stands for red, G represents green and blue by B .Color space is defined as a model for representing colors in terms of intensity Each pixel in the RGB image is represented by 24 bit. Each component is represented by a 8 bit value The color element value typically ranges from 0 to 255, in a grey-scale image where black is represented by 0, white by 255 and the values in between are the different shades of grey which ranges from black to white.

Many CBIR systems use color histograms as a color feature to represent the image. The color histogram is a graphical representation of an image that shows the distribution of colors. Distribution of each color is represented using bars and the height of each bar represents the number of pixels in the image having the same color. The x-axis of the histogram represents the number of colors being used in the image, while the y-axis represents the number of pixels. For normalizing the histogram bars similar to each other are grouped together and put them in the same bin. Thus it reduces the possibility of placing similar colors in different bins, but it will increase the possibility of assigning distinct colors to the same bins. So we are representing colors in an image using less number of bins thereby reducing information in the .But the advantage is, lesser space is needed to store the feature and the time for searching the features will decrease.

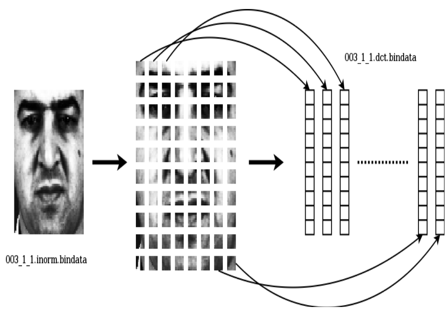

Fig 3: Extracted features of an image are stored in multi arrays

Therefore, there is a trade-off between the number of bins in the histogram and the information gained from the histogram. The trade-off is also between the processing time, storage space, and the accuracy of the results. Color histograms can be obtained by different methods. Traditionally, two known methods are used to calculate color histograms.

1. Global color histogram (GCH)
2. Local color histogram (LCH).

GCH method is the most commonly used method. GCH calculate the histogram for the entire image and two images are compared by the distance between the histograms of the two images. The drawback of this method is that it does not define the full information of the image. So the difference value wont provide the correct information and there is possibility, that two different images can be reported as similar by the system. LCH method takes into account the color distribution of regions. This method divides the image into fixed blocks and calculates the color histogram for each block individually. The image will be represented by these color histograms. To compare two images, using their histograms, we calculate the distance between a block from one image and another block from the second image in the same location. This method improves the efficiency of retrieving images more than using GCH. However, this method needs more computation, and it does not work well when images are translated or rotated. Figure 4.4 shows an example of color histogram for an image in RGB color space.

4. Bag of Words

To improve the search efficiency we are introducing the bag-of-words (BoW) model in a more detailed way and the steps involved. The BoW model proposed was relied by State-of-art large-scale image retrieval systems. BoW model in CBIR is mainly designed for the local descriptors of images which describe regions around the key points detected in the images. Different from global features describing a picture in a holistic way, one image can have a bunch of salient patches around key points. And local descriptor like 128-dimensional SIFT is demonstrated to be a good way to represent the characteristics of these patches. But once we extract such local descriptors for each image, the total number of them might be huge. The images are searched using the query image, and its consists of local descriptors. Then searching nearest neighbors using the local descriptors of query image are time consuming. We are proposing BoW to solve this problem by quantizing descriptors.

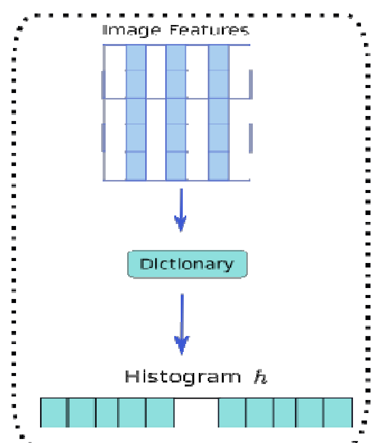

Figure 4: Single Dictionary: a single dictionary of visual words is generated from the pool of features, which is used to generate the histogram for the image.

Therefore, BoW is proposed as a way to solve this problem by quantizing descriptors into "visual words", which decreases the descriptors' amount dramatically. In this way, each image can be viewed as a long and sparse vector of words and therefore we can mimic text-retrieval systems, applying scalable indexing and fast search on this vector space. In BoW model, the following several procedures must be taken into consideration.

- Choose methods to detect key points and describe local patches around key points. This step can be viewed as preprocess of BoW model and it is mostly related to computer vision in this paper. How to detect and describe key points should be critical to the retrieval performance since all the following procedures are relied on this "data generation" step.

- Quantize local descriptors into "visual words". We call this procedure as vocabulary generation and local descriptor quantization. The fundamental difference of BoW model in text and image retrieval is that text words are sampled naturally according to language context; visual words are the outcomes of quantization. It is artificially generated and infers statistical information which has less semantic sense. So how to generate such a vocabulary remains an interesting question and different methods are proposed toward this direction.

- Indexing and search. Once we obtain the vocabulary and quantization method, we can mimic text-retrieval systems, applying scalable indexing and fast search on this vector space. Inverted file has been proved to be an efficient and common way to achieve this goal.

Although BoW model has shown excellent performance in image retrieval task, it still suffers from some problems. One major concern is that BoW model could not support spatial properties of local descriptors of images. Several papers were then published to improve this framework from different angles so that spatial information is taken into account. Another disadvantage is that key points are extracted from the grey level images and local descriptors do not contain any color information [9]. Moreover, they only grasp the local information, losing the overall distribution of visual information. The performance could be enhanced if we can combine global features and local descriptors together.

5. Experiment Results

To show the practical relevance of the proposed CBIR system, retrieval experiments are conducted. The technique is evaluated based on recall and precision. A recall rate can be defined as the number of relevant documents retrieved by a search divided by the total number of existing relevant documents (which should have been retrieved). A precision rate on the other hand is the number of relevant documents retrieved by a search divided by the total number of documents retrieved by that search.

Fig 5: Query Image

Fig 6: Search result of the input query image from a database based on similarity

6. Conclusion

Searching an image from a large image data base takes lot of time. Searching data based on features will gain tremendous amount of gain in time because very few features are stored for each image. Like wise bag of words also store lesser information of each image and hence very effective in searching of image .

References

[1] Long, F. H., H. J. Zhang and D. D. Feng, Fundamentals of content-based image retrieval, *in Multimedia Information Retrieval and Management - Technological Fundamentals and Applications*, Springer-Verlag, pp.1-26, New York, 2003.

[2] Lowe, D. G., Distinctive image features from scale invariant features, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.

[3] Kumar, A. and Sminchisescu, C. Support kernel machines for object recognition, in *ICCV*, 2007.

[4] Zhang, H., Berg, A., Maire, M. and Malik, J., Svm-knn: Discriminative nearest neighbor classification for visual category recognition, in CVPR, 2006.

[5] Yang, C. C., Prasher, S. O. , Enright, P., Madramootoo, C., Burgess, C., Goel, P. K., Callum, I., Application of decision tree technology for image classification using remote sensing data, in *Agricultural Systems* 76, pp. 1101-1117, 2003.

[6] K. Barnard, P. Duygulu, D. Forsyth. Modeling the Statistics of Image Features and Associated Text. Proc. Document Recognition and Retrieval, Electronic Imaging, San Jose, CA, Jan. 2002.

[7] Barnard, K., Duygulu, P., Forsyth, D., de Freitas, N., Blei, D., & Jordan, M. (2003).Matching words and pictures. JMLR, 3, 1107–1135.

[8] Bosch, A., Zisserman, A., & Munoz, X. Image classification using random forests and ferns. In ICCV.

[9] Bosch, A., Zisserman, A., & Munoz, X. Representing shape with a spatial pyramid kernel. In CIVR.

[10] Boureau, Y., Bach, F., Le Cun, Y., & Ponce, J. Learning mid-level features for recognition. In CVPR

Authors Profile

Shine Raj *G* is working as Asst. Professor in Computer Science department. He has more than 8 years experience in teaching and has published papers on data mining and image processing. His research interests include image processing, data mining, and image mining

Anver Muhammed K.M is working as System Administrator in Computer Science department. He has more than 10 years experience in System Administration and Implementation. His research interests include image processing and network security.