International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 - 5161

Available at http://inpressco.com/category/ijcet

Research Article

Adaptive Image Retrieval Technique using Texture and Color Features

E Ramalakshmi[Â*] and Keerthi Lingam[Â]

[Â]Department of Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad,India

Accepted 30 May 2014, Available online 01 June2014, **Vol.4, No.3 (June 2014)**

Abstract

CBIR is the area in image mining system which performs retrieval based on the similarity defined in terms of extracted features with more objectiveness. It aims at searching image databases for specific images that are similar to a given query image. In CBIR the features of the query image alone are considered and this is the drawback of this technique. Thus, a novel method based on clusters is emerged that improves user interaction with image retrieval systems by fully exploiting the similarity information. In order to reduce the searching time of images from the image database, the query image will be classified in this method. The target image is selected optimally according to the color characteristics with feature vectors which represent typical color distributions. The proposed technique aims to reduce the searching time of image retrieval and hence it improves the performance of image retrieval system.

Keywords: *Image Retrieval, CBIR, Cluster Based.*

1. Introduction

In this digital era, multimedia data plays a very vital role in every field such as e-commerce, entertainment, education, medicine, aerospace and so on. With the increasing use of internet, there is an enormous volume of multimedia data available to the users. Even though, the huge availability of multimedia data like images, audio, video is useful and appreciable, sometimes it proves to be a bane as there are certain difficulties to gather the required useful data in an appropriate way. A huge volume of digital images are generated every day and if analyzed properly there is a lot of useful information available to the users. With the drastic increase in the multimedia databases, the usefulness of such information is dependent on how well it can be accessed; searched and how well knowledge can be extracted from it. Due to lack of proper extraction methodologies, the users are unable to extract the relevant information from the available databases (Hsu *et al*, 2002).

In the present scenario, the conventional user defined text searches are based on keyword, size, type, date and time of capture, identity of the owner etc. This search based techniques are successful but do not meet the user's final requirement in all cases. So, many researchers are concentrating on the search based on visual content i.e., finding the images similar to an input query image.

There has been a lot of research in text based retrievals, but image retrieval is also gaining its momentum with regard to both still and moving images. Efficient image database retrieval can be done only if we have a system that is able to automatically extract relevant features directly from the images stored in the database. So image mining proves to be efficient as it deals with complex operations like image retrieval, indexing and storing. There is a common misconception that image mining is an extension to data mining which is not so. Image mining is an upcoming research field and is still is in its extracting relevant knowledge from image data still remains a difficult task. The next few sections deal with a brief introduction to image mining and a broad explanation of the various phases involved in an image retrieval system.

The main objective of image mining is on research and technological activities for automated and user centered extraction of information from very large and heterogeneous image databases. Image mining is focused on extracting patterns, implicit knowledge, image data relationship or patterns which are not explicitly found in the images from databases or collections of images. Some of the methods used to gather knowledge are: image retrieval, data mining, image processing and artificial intelligence.

Image information mining is an interdisciplinary effort. Image retrieval is one of the phases in image mining. An image retrieval system is used for browsing, searching and retrieving images from a very large database of digital images. All image retrieval methods use some techniques of adding metadata such as keywords or descriptions to the images such that retrieval can be performed over the annotation words than images. Manual image annotation is slow, lengthy and costly.

Determining the complexity of image search system design is important and for this purpose it is crucial to

*Corresponding author: **E Ramalakshmi**

understand the nature of image data. The design is also influenced by factors such as expected user traffic and the diversity of user-base for a search system. In this aspect, search data can be classified into the various categories such as archives, domain-specific collection, enterprise collection, personal collection and web (Hsu *et al*, 2002).

(i) Archives – consists of large volumes of structured or semi-structured homogeneous data related to a specific topic.

(ii) Domain-Specific Collection – consists of large volumes of homogeneous collection accessed by controlled users with very definite objectives. Biomedical and satellite image databases are examples of such a collection.

(iii) Enterprise Collection - a heterogeneous collection of images that is accessible to users within an organization's intranet. In many special locations pictures may be stored.

(iv)Personal Collection - usually consists of a largely homogeneous collection and is generally small in size, accessible primarily to its owner and usually stored on a local storage media.

(v) Web - With an internet connection World Wkide Web images are available to everyone. These image collections are semi-structured, non-homogeneous and huge in volume, and are generally stored in big disk arrays.

2. Related work

2.1 Content based image retrieval

In understanding of image seeking behaviour and Research and practical developments in image retrieval have been major areas of development in information retrieval in the last decade.These methods allow image mining to have two different approaches. First, is to extract only from databases or collections of images, and second, dig or mine a combination of associated alphanumeric data and collections of images. An extension of data mining to image domain is not simply Image mining. Some of the current image retrieval techniques consist of Content-based image retrieval (CBIR) and Clustering techniques. CBIR have been progressed in four areas such as Global image properties based, relevance feedback based, region-level features based and semantic based.

A new technique called Image retrieval based on clusters is developed for improving user interaction with image retrieval systems by fully exploiting the similarity information. In this method query image is classified in order to reduce the searching time from the image database. According to the color characteristics, with compact feature vectors, that signify characteristic color distributions the target image is optimally selected.

The image mining researches can be classified into two types. One is image processing in which, it involves a domain specific application where the focus is in the process of extracting the most relevant image features into a suitable form (Fayyad U, *et al*, 1997; Kumar *et al*, 2012; A. Kitamoto, 2001) and the other is image mining in which, it involves general application where the focus is on the process of generating image patterns that may be helpful in the understanding of the interaction between high-level human perception of images and low-level features (Ordonez, C *et al*, 1999; O. R. Zaiane *et al*, 1998). So image mining improves the accuracy of images retrieved from image databases.

Image mining generally deals with the extraction of implicit knowledge, relationship between image and data, or other patters not explicitly stored from the low-level computer vision and image processing techniques. So image mining focuses in the extraction of patterns from a large collection of images, but the focus of image processing techniques and computer vision is in understanding or extracting specific features from a single image.

Existing techniques in image retrieval and classification concentrate on content-based techniques. Various systems like the QBIC (W. Niblack *et al*,1994), Retrieval Ware (Dowe, J,1993) and Photo Book (Pentland *et al*,1996) etc have different features, but are still used in particular domains. Jain et al apply color features combined with shape for classification. Smith and Chang apply color and the spatial arrangements of these color regions. Since perception is subjective, single feature is not sufficient and no single representation of a feature is sufficient. Hence multiple representations and a combination of features are required (Foschi *et al,2002*).

2.2 Color and texture based image retrieval

In the color based image retrieval, the RGB color model is normally used. Color images generally are in three-dimensional. RGB color components are taken from every image. Then for both query image and target images the average value of R, G and B values are computed. Then for each image three average values are stored and considered as features. Using these stored features the target image from the repository is retrieved with reference to the query image. Then according to their texture the top ranked images are re-grouped.

In the texture based approach the parameters gathered based on the statistical approach. The different texture parameters like entropy, dissimilarity, contrast, homogeneity, mean, variance and standard deviation of both query image and target images are calculated and the required image from the repository is extracted from the calculated values.

Then, the pre-processed images in the database are classified as low textured and high textured detailed images respectively. These classified images are now subjected to color feature extraction. The resulted values of both the query image and target images are compared by Euclidean distance method.

Content-based image retrieval (CBIR) is the application of computer vision to the image retrieval. CBIR try to avoid using of textual descriptions and instead retrieves images based on similarities in their contents (textures, colors, shapes etc.) to a user-supplied query image or user-specified image features. But the drawback in CBIR is the features of the query image alone are considered.

3. Proposed solution: Adaptive image retrieval

In this, a new method for image classification is formulated in order to reduce the searching time of images from the image database. The course content of image is grouped under two categories as:

a) High-texture detailed image
b) Low textured detailed image

Thereby, we can reduce the search space by half of what was earlier. If we go to more number of groups, they may lead unnecessary overlapping problems or may produce approximate results but not exact. So, the main focus on this classification is to make use of "textures" of an image as the texture-based classification is easy, simple and efficient for real time applications as well as segmentation based techniques.

The main objective of the work is to design an algorithm to create true databases from collections of multimedia data (specifically images) by mining content from the data offline in order to efficiently support complex queries at run-time.

Algorithm

1. Input the query image
2. Process the RGB components in the image
3. Cluster the images based on their RGB components.
4. Calculate the texture values for all the images in the database so that every image falls into one of the categories of low and high textured images.
5. Sort the results based on texture and color analysis.
6. Do the same procedure of finding the color and texture analysis for the query image.
7. Compare the results of query image with the remaining all images in a database.
8. Select the most similar image as a target image.

4. Methodology

The process of Image retrieval undergoes various phases as shown in Fig 1. Which are:

1) Query image selection iv. Sorting the results
2) RGB based clustering v. Similarity Comparison
3) Texture based clustering vi. Target image retrieval

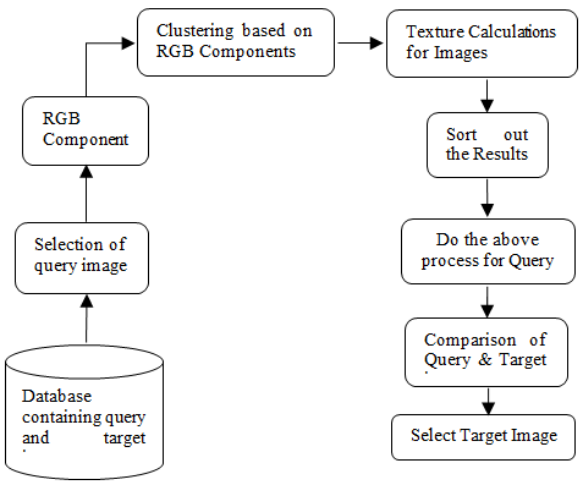

Fig 1: Process of Image Retrieval

4.1 Query image selection

The user who wish to search an image from the image repository, selects the query image based on his requirement. User can select any image from data base as shown in Fig 2.

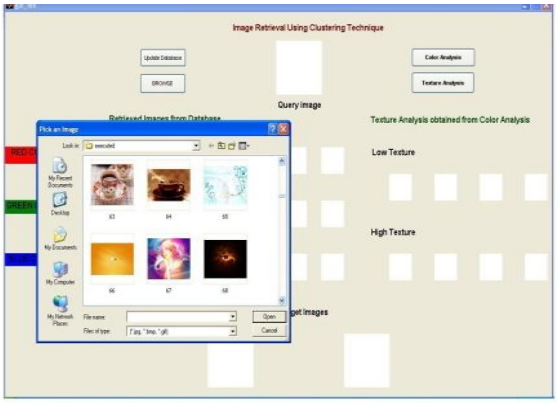

Fig 2: Selecting Query image

After selecting the query image we can view it on the GUI as shown in Fig 3 sothat we can check under which group the query image falls in to.

Fig 3: Selected Query Image

4.2 RGB based clustering and Texture based clustering

The average values for the RGB components were calculated for all images. Then the values are to be compared with each other in order to find the maximum value of the components. The set of images are clustered in to Red,Green and Blue planes as shown in Fig 4. Whenever the query image is given calculate the RGB components average values. Then compare this with the stored values.

The texture reflects the energy content of the image. If an image contains many high textures then the energy will be high as compared to that of low texture images. The set of images are classified into High and Low textures as shown in Fig 5.There are several texture parameters to be considered. Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image is calculated for the query and target images. Entropy is defined as $-sum(hc.*\log2(hc))$

where hc is the histogram count obtained from the histogram calculation.

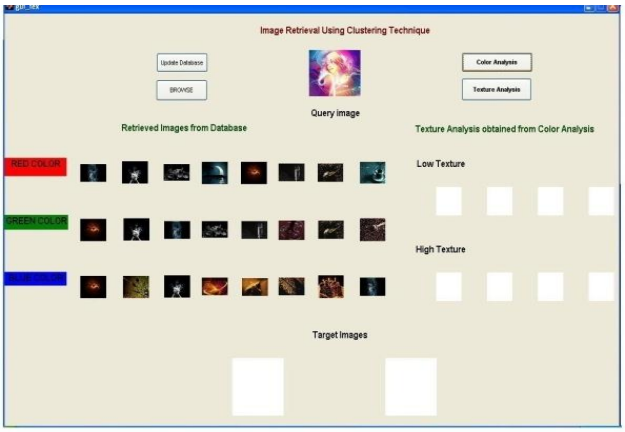

Fig 4: RGB based clustering

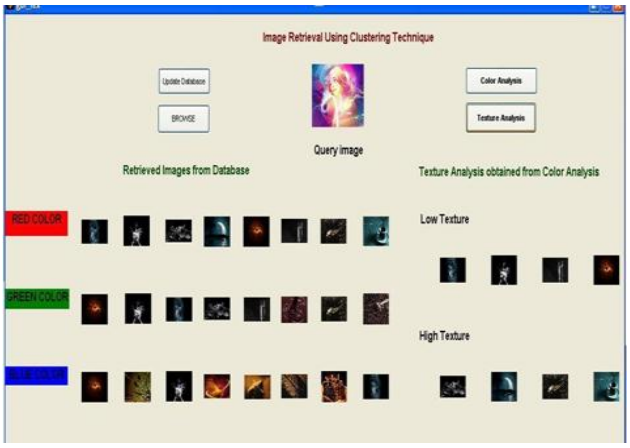

Fig 5: Texture based clustering

4.3 Clustering of Images

Clustering will be more advantageous for reducing the searching time of images in the database. Clustering methods also allow one piece of data to belong to two or more clusters rather than belonging completely to just one cluster. The target image is selected from the most appropriate cluster**.** The average values of Red,Green and Blue pixel values that fall into respective clusters are tabulated in Table1,Table2 and Table3.The texture values of images which are grouped in to low and high clusters are shown in Table 4 and Table 5.

Table1: Sample images falls in **Red cluster**

S.No	Image name	Red plane values
1	'82.jpg'	12.4721
2	'76.jpg'	17.1758
3	'80.jpg'	24.4583
4	'87.jpg'	25.0951
5	'68.jpg'	25.1947
6	'21.jpg'	26.2303
7	'95.jpg'	35.4466
8	'27.jpg'	39.9332

Table2: Sample images falls in **Green cluster**

S.No	Image name	Green plane values
1	'68.jpg'	8.7339
2	'76.jpg'	17.121
3	'82.jpg'	20.735
4	'80.jpg'	24.458
5	'21.jpg'	25.756
6	'38.jpg'	30.907
7	'95.jpg'	31.990
8	'48.jpg'	39.768

Table3: Sample Images Falls in Blue cluster

S.No	Image name	Blue plane
1	'68.jpg'	1.3779
2	'270.jpg'	11.8447
3	'76.jpg'	17.7014
4	'70.jpg'	20.1056
5	'93.jpg'	23.3522
6	'102.jpg'	23.6730
7	'52.jpg'	23.9821
8	'82.jpg'	24.2737

Table4: Sample images falls in Low Texture

S.No	Image	Low Texture values
1	82.jpg	1.6179
2	76.jpg	1.9702
3	21.jpg	2.0664
4	68.jpg	2.2532

Table5: Sample images falls in High Texture

S.NO	Image	High Texture values
1	80.jpg	2.2649
2	87.jpg	2.2886
3	95.jpg	2.4599
4	27.jpg	3.0057

4.4 Similarity Comparison

The retrieval process starts with feature extraction for a query image. The features for target images are usually pre-computed and stored as feature files. The same features are calculated for the query image as well. Using these features together with an image similarity measure, the resemblance between the query image and target images are evaluated and sorted. Similarity measure describes the resemblance in contents between two images. "Euclidean Distance" is used to measure the similarity comparison.

Euclidean Distance = sqrt[$(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2$ ...]

4.5 Target image retrieval

Collections of target images that are close to the query image are selected as the neighborhood of the query image.

Fig 6: Target image retrieval

This process is shown in Fig 6. A typical CBIR system by passes these two stages and directly outputs the sorted

results to the display and feedback stage. This system can be designed independent of the rest of the components because the only information needed for the system is the sorted similarities. This implies that this module may be used with a typical CBIR system regardless of the sorting method, image features that are considered and whether there is feedback or not. The mandatory requirement is a real-valued similarity measure satisfying the symmetry property.

Conclusion

In this system, a new image retrieval technique based on clusters is introduced in order to reduce the searching time space. Moreover, the RGB components of the color images are classified in different dimension in order to create Red, Blue and Green image clusters. Then based on texture analysis, images are clustered into two categories as high and low textured images. This project is implemented in MATLAB and observed from results that search time and search space is less than image retrieval from the entire database without clustering. So, using clustering technique of image retrieval, this paper compares the query and target images based on RGB components and texture analysis which reduces the searching time of image retrieval and improves the performance of the system.

References

Hsu, W., Lee, M. L., & Zhang, J. (2002). Image mining: Trends and developments. *Journal of Intelligent Information Systems*, 19(1), 7-23.

Fayyad, U. (1997), Data mining and knowledge discovery in databases: implications for scientific databases, Proceedings of *Ninth International Conference on Scientific and Statistical Database Management* pp. 2-11

Kumar, K. K., Bhutada, S., & Balaram, V. V. S. S. S. (2012), An Adaptive Approach to Relevance Feedback in CBIR Using Mining Techniques, Proceedings of *International Conference on Electronics Communication and Information Systems* Vol. 80.

Asanobu, k.i.t.a.m.o.t.o. (2001). Data mining for typhoon image collection, Proceedings of the *2nd International Workshop on Multimedia Data Mining* pp. 68-77.

Ordonez, C., & Omiecinski, E. (1999). Discovering association rules based on image content, *Research and Technology Advances in Digital Libraries*. Proceedings. IEEE Forum pp. 38-49.

O. R. Zaiane, J. W. Han et al.(1998) Mining Multimedia Data, *CASCON'98: Meeting of Minds*, pp 83-96.

Foschi, P. G., Kolippakkam, D., Liu, H., & Mandvikar, A. (2002), Feature Extraction for Image Mining, *Multimedia Information Systems* pp.103-109.

W. Niblack et al.,(1993) The QBIC Project: Querying Images by Content Using Color, Texture, and Shape, Proc. *SPIE*, vol. 1908, pp. 173–187,

Dowe, J. (1993), Content-based retrieval in multimedia imaging, *InIS&T/SPIE's Symposium on Electronic Imaging: Science and Technolog*, pp 164-167.

Pentland, A., Picard, R. W., & Sclaroff, S. (1996), Photobook: Content-based manipulation of image databases, *International Journal of Computer Vision*,18(3), pp 233-254.