Global Scientific JOURNALS

GSJ: Volume 8, Issue 4, April 2020, Online: ISSN 2320-9186
www.globalscientificjournal.com

Image Retrieval In Multimedia Database by applying enhanced color histogram.

HABIMANA Jean Bosco[1], Dr. Papias NIYIGENA[2]

Department of Information Technology, University of Lay Adventist of Kigali

ABSTRACT

This Paper was set to deal with "The lack of acute precision in retrieving images by using color histogram", which is the identified problem in how to retrieve images in multimedia database. Color histograms can potentially be identical for two images with different object content which happens to share color information. Conversely, without spatial or shape information, comparison objects of different color may be indistinguishable based solely on color histogram comparisons. As a way to find a solution to this problem color and shape were considered to make sure the retrieved images have little chance to give unexpected errors. The solution to the problem is made possible by the approach of "image retrieval by means of Enhanced color histogram". Image retrieval interface was used for querying, browsing, searching and retrieving closely, images from a large database of digital images. This database is prepared using color visual feature. The approach used in this system was "Enhanced color histogram", which represents the number of pixels that have colors in each of a fixed list of color ranges that span the image's color space, the set of all possible colors. Considering the results given by both color histogram and enhanced color histogram, enhanced color histogram was more effective for all queries than color histogram.

1. Introduction

This paper presents an approach to image retrieval in Multimedia Databases. Although the terms multimedia and database have very precise meanings, the concept of multimedia database is a collection of related multimedia data. The multimedia data include one or more primary media data types such as text, images, graphic objects (including drawings, sketches

and illustrations) animation sequences, audio and video. The concept of image has been widely investigated throughout the last century, both in the field of psychology and in that of computer science, attempting to define a theory consistent with the huge amount of experimental data. An important point, discovered by computer scientists only in recent times, like by Suzuki, WA (2010), is the discrepancy between the concepts of image in psychology and in computer science. In computer science, usually, the image has the target of recognizing an object under conditions of uncertainty. There is an object and a model of the same object: The system has to assess if the actual appearance of the object, different from the appearance of the model due to noise and distortion, is consistent with the model itself. Computer scientists, thus, have to define the class of possible transformations that an object can undergo. It is, therefore, possible, for computer scientists, to assess the comparison between two rotated images of a cube, since it is easy to model all the possible rotations of a cube in the space, but the task of assessing the comparation between a cube and a tetrahedron is a difficult one. And an even more difficult task is that of assessing if the image of a cube is more comparison to that of a tetrahedron or to that of a sphere P.G Young et al (2008). The concept of comparison for this scenario is that of psychologists, since the approach used in this research is to model the behavior of human mind in comparing perceptual stimuli. Therefore, in the following, is a brief review of some of the comparison theories presented by psychologists.

1.2. Conceptual framework

Comparison image retrieval system is an interface between high level system (the human brain) and low level system (computer). The human brain is capable of performing complex visual perception. Comparison image retrieval should be designed to match with human perception, but how humans judge the comparison between images is a topic of ongoing research.

Figure 1.1 shows the framework of the proposed image retrieval approach.

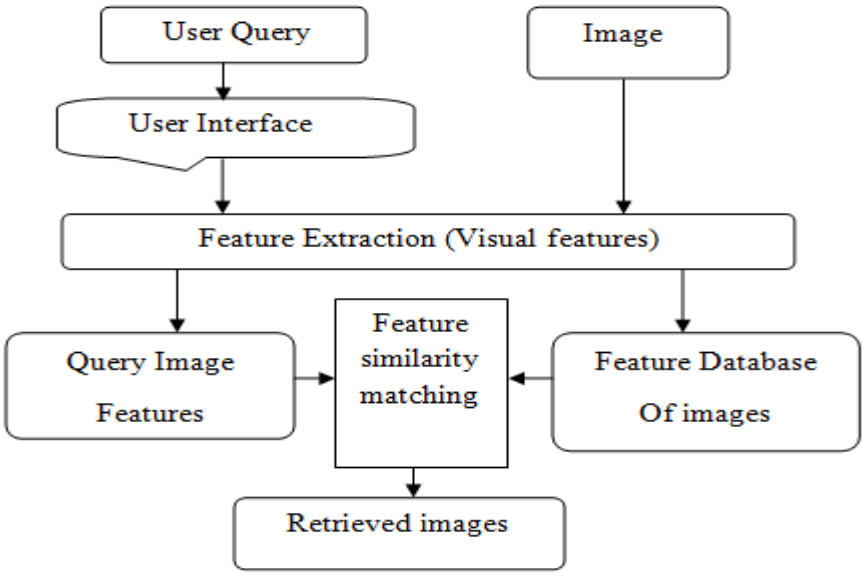

Figure 1.1: Proposed image retrieval approach

This figure represents the proposed approach to retrieve images to the query image in a multimedia database. The process starts with a query image being submitted for image retrieval, its color features are extracted and matching operation is performed between query image features and the images features stored in database . After matching operation is finished, all images from the database with features close to the query image are then retrieved from the database.

2.0. LITERATURE REVIEW

2.1. Introduction

A typical image retrieval system include three major components: i) feature extraction, ii) high dimensional indexing, iii) system design. Some work has been done on how to improve the performance of image retrieval by learning from user feedback. Relevance feedback from the user is used to search the target. This approach is based on a Bayesian framework that incorporates an explicit model of the user selection process Hanwang, Z. et al (2012). They show that the feedback system works much better than random chance in a queries setting. They also conduct a user-study of the performance of their system. Latika P., Manisha S., Kamal M. (2012); introduced a learning component in their system by using positive and negative examples which lets the system choose image groupings within and across images

based on color and texture cues. The user-chosen region defines the positive or negative examples which the system tries to generalize and label various parts of the scene. Positive and negative examples are used to first construct histograms Perrinet, L. U. and Bednar, J. A. (2015). Then, a scoring function is constructed from a comparison of the extent of overlap between the histograms. Learning is achieved by exploiting neural networks to partition the feature space into clusters of similar patterns, each pattern with a known label. The learning scheme is test on a texture database with images. The results suggest that the learning scheme improve the recall in the scope of 20 to 60 images. Kong Suran (2013) uses a two-layer model for a multimedia object.

The top layer contains different image features and their corresponding distance functions; the bottom layer contains the actual values of these features. Relevance feedback is learned in the two layers separately. For the top layer, the initial retrieval results come from overall results from different features and distance functions. Users are supposed to provide a rank list of the initial results based on their perception criteria. Then the best feature and distance function pair is found by finding the smallest difference between the initial ranks and users' ranks. Based on this best pair of feature and distance function, a new retrieval is returned. Although this approach seems to find the best, it assumes that users judge their ranks based on only one feature. For the bottom layer, they used the standard techniques in information retrieval.

2.2 Color Histogram

Apurva, S., Dr. Swati. (2015).define color histogram as representation of the distribution of colors in an image and also represents the number of pixels in digital image that have colors in each of a fixed list of color ranges that span the image's color space having the set of all possible colors. The color histogram can be built for any kind of color space but more often it is used for 3-D spaces like RGB or HSV. Intensity histogram may be used instead for monochromatic images. For multi-spectral images, where each pixel is represented by an arbitrary number of measurements (for example, beyond the three measurements in RGB), the color histogram is N-dimensional. Where, N being the number of measurements taken. Each measurement has its own wavelength range of the light spectrum, some of which may be outside the visible spectrum.

3.0. METHODOLOGY

3.1. Research design

The research design is the combination set of methods chosen for the empirical part of the study Bergold, Jarg & Thomas, Stefan (2012). Research design is a comprehensive process of the study and the problem formulation through dissemination of findings. This research used Enhanced color histogram methodology to perform comparison image retrieval in multimedia database; the design of this query system is simulated entirely using a programming language and interactive environment known as MATLAB. This programming language enables the user to perform computationally intensive tasks and to implement comparison image retrieval process. MATLAB is a software package for high-performance numerical computation and visualization. It provides an interactive environment with hundreds of built-in functions for technical computation, graphics and animation. It is also an efficient program for vector and matrix data processing. It contains ready functions for matrix manipulations and image visualization and allows a program to have modular structure. Because of these facts MATLAB has been chosen as prototyping software. It provides a suitable environment for image processing. Its built-in functions and syntax makes it a more versatile and faster programming environment for image processing. The design was very appropriate because due to its ability to elicit a diverse range of information about the area of study.

3.2. Sampling techniques

Samples should be as representative as possible, because too small sample is likely to yield under-estimated information that may not reflect the actual population characteristics or perceptions Meng, Xiangrui (2013). In situations where a population is too small to be sampled, it is logical to sample all the elements Deepan, P., et al (2015). This study purposively selected images as sample size which is divided into three classes: class of motorbike, class of airplane and class of persons. The study was purposively as it helped enough to get better information for the techniques (enhanced color histogram) to retrieve images.

3.3. Data collection procedures

The study use experimental method of data collection by choosing same images from the database. To conduct retrieval experiments using this system, a query image was chosen and the system returns all images in ranked order of perceived relevance to the query according to the search algorithms used. The research purposively collected images and grouped them into

three categories. All images were a depiction of the Google image period with snow, decorative items, magenta etc. Most images were very similar in color content, thus supporting a more rigorous testing of the system's ability to retrieve relevant images based solely on enhanced color histogram.

3.4. Implementation

MATLAB provides a customizable bag of features framework to implement a comparison image retrieval system. The following steps outline the procedure:

1. Browse query image
2. Select the Image for Retrieval
3. load the Images in the database
4. Search for Similar Images

These steps create an image retrieval system for searching the three different Categories of images in the database.

4.0.RESEARCH RESULTS AND DISCUSSION

This SIR system was implemented in MATLAB and SQL was used as a backend tool for database creation and management. The comparison image retrieval by a given query image involves searching the database for images. The retrieval results were a list of motorbike, airplane and persons images ranked by their similarities measure with the query image.

The images in the database were ranked according to their classes to the query image in numbering orders, and then the ranked images were retrieved. The computed comparison query images were ranked and retrieved according to the closest images; in addition, if the images were less than a certain threshold set, the corresponding original image was close or matches the query image. Precision P is defined as the ratio of the number of retrieved relevant images r to the total number of retrieved images n, i.e., P =r/n. Precision measures the accuracy of the retrieval and is defined

$$\text{Precision} = \frac{\text{No .of relevent image retrieved}}{\text{Total No .of image retrived}} = \frac{\text{r}}{\text{n}}$$

Recall is defined by R and is defined as the ratio of the number of retrieved relevant images r to the total number m of relevant images in the whole database, i.e, R=r/n. Recall measures the robustness of the retrieval.

$$\text{Recall}=\frac{\text{NO. of relevent image retrieved}}{\text{Total no. of relevent images retrieved}}=\frac{r}{n}$$

The databases of images of different classes respectively were used to check the performance of the developed algorithm. Classification of different classes in the database and the number of images in each class is depicted in table 1.1

Table 1.1: Classification for database of images.

Class	Persons	Motorbike	Airplane
No. of images	40	36	32

The table 1.1 indicates how the user of this query system can open a folder that contains images after opening the folder using GUI command which is the gui.fig the file name of this query system.

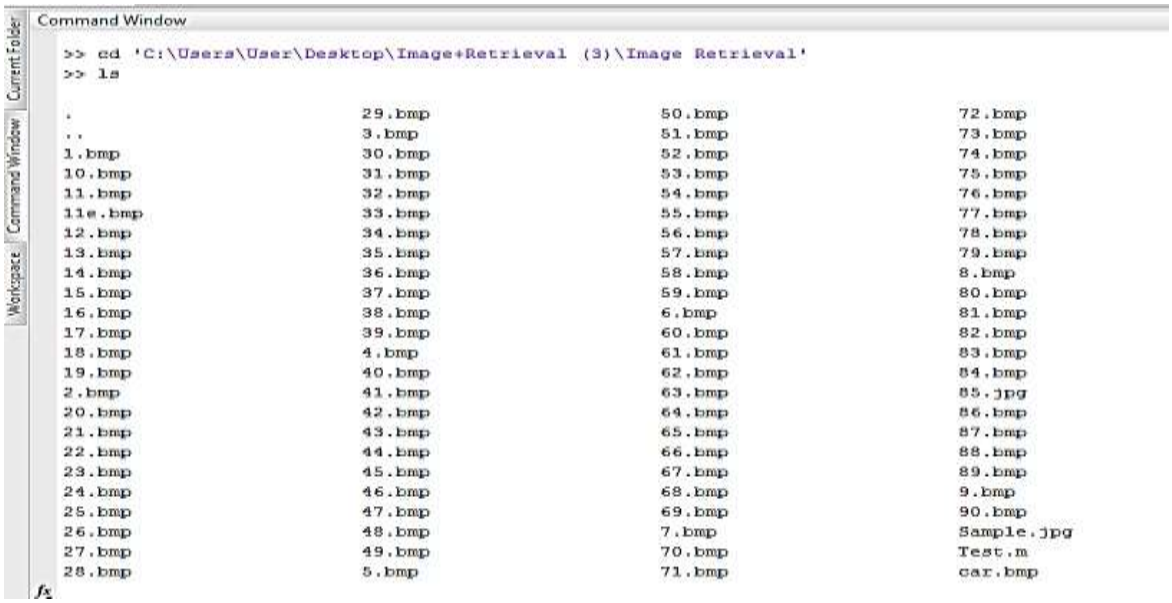

Figure 1.2: Query path of GUI.

Once user uses gui.fig to open the query system.Figures4.2 and 4.3 shows the first window (graphical user interface) that guides the user to all options for this query system, the window has the option of browsing image for comparison.

Figure 1.3: Input image from database.

Once a user opens GUI of the system a new window appears,figure 1.3. This window has four options which are respectively browse, load database, search and clear. “Browse” helps the user to choose the query image,after the query image is there then “load_database” helps in gathering all images that are similar to the query .Now that the user has loaded all images simalar to the input image he/she gets a message telling that database has successfully been loaded , “search option” plays a ro;2le of displaying all relavant images from the loaded images.At the end this operation the user can use “clear option” to end the query for the new one.

4.1.CONCLUSIONS

Based on the results, image retrieval approach in multimedia databases by applying enhanced color histogram produces a better performance of the query system. The results of the SIR approach give a great precision of relevant image of user perception, from this results the users can judge if the image is really what they expected to get.

4.2. FUTURE ENHANCEMENTS

Developments and studies are going on for further improvements in design and performance of “Image Retrieval Systems”. this research has shown that images retrieved by using the enhanced color histogram may not be semantically related even though they share similar color distribution in some results.

In the future, enhancements implementation can be on:

• Shape and Texture analysis.

Textures and shapes reveal important information concerning the fundamental visual elements in an image and characterizes the surfaces of many classes of objects.

• Image ranking in Euclidean Distance method.

Future research should address model selection. Potentially, if one was given a small labeled set or a query set greater than size 1, one could use standard cross validation techniques.

REFERENCES

1. Amandeep Khokher. et al (2011).Image retrieval: a state of the art approach for CBIR. *International Journal of Engineering Science and Technology*, 3(8), 6698-6704.
2. Avneet, K., Vijay, K. B., Navkirat, K. (2013).Color and texture based image retrieval: a proposed method. *International Journal of Research in Engineering and Technology*, 2(11), 498-501.
3. Cheruiyot, W., Guan, Z.T., Felix, M., Joseph, C. M. (2011). Query quality refinement in singular value decomposition to improve genetic algorithms for multimedia data retrieval. *Springer-Verlag.* 17.507–521. DOI: 10.1007/s00530-011-0231-3
4. Fazal,M., Baharum, B.(2013).The statistical quantized histogram texture features analysis for image retrieval based on median and laplacian filters in the DCT domain. *The International Arab Journal of Information Technology*, 10(6).
5. Hany, F. A., Gamal, A. El-Fishawy, N. (2013). Comparative Study on CBIR based on Color Feature.*International Journal of Computer Applications*, 78(16), 0975 – 8887.
6. Izquierdo, E.&Qianni. Z. (2013). Histology image retrieval in optimized multifeature spaces. *Journal of Biomedical and Health Informatics,*
7. Mangijao, S., Hemachandran, K. (2012). Image retrieval based on the combination of color histogram and color moment. *International Journal of Computer Applications,* 58(3), 27-34.
8. Manimala, S., Hemachandran, K. (2012). Content based image retrieval using Color and Texture. *Signal & Image Processing : An International Journal (SIPIJ),*3(1),39-57
9. Metty, M., Sarifuddin, M., Eri, P., Djati, K., Suryadi, H. (2014). Content based image retrieval using local color histogram. *International Journal of Engineering Research,*3(8), 507-511.