INTERNATIONAL JOURNAL OF ADVANCED STUDIES IN COMPUTER SCIENCE AND ENGINEERING
IJASCSE VOLUME 5 ISSUE 11, 2016

Enhanced Image Retrieval using Distributed Contrast Model

Mohammed. A. Otair
Faculty of Computer Sciences & Informatics
Amman Arab University
Amman, Jordan

***Abstract*— Recent researches about image retrieval methods did not reach an advanced strategy that allows user to deal directly with images, but although there was a good inventory that makes it very close to develop new technologies for this issue. This paper present Enhanced Feature Sets Based Similarity Measures for Image Retrieval (called Enhanced Contrast Model-ECM), which deals directly with images features that will improve image retrieval process. Experiments results show that the ECM enhances the execution time with an average speed up of 5.685 over the classical contrast model.**

Keywords-component; Image retrieval; similarity; content based image retrieval (CBIR); contrast model.

I. INTRODUCTION *(HEADING 1)*

In the last few years the most important researches in retrieval models concerned about image retrieval, and how it is possible to retrieve images from databases by extracting the features of that images, which make it easy to retrieve images by only choose some features or colours that construct the image. The main idea in [1] was to develop a new algorithm, which takes some features of the required image and make some comparisons with the images features from a database -that already constructed by extracting images features- and by matching queried image features with images collection from the database in order to facilitate finding some similar images.

Many researchers [3, 4, 5, 7, 8, 9, 10, and 11] proposed many solutions in order to enhance the performance and efficiency of the images retrieval systems. However, each of which has some disadvantages and drawbacks. The proposed algorithm is designed based on distributed databases that will enhance the last algorithm by just classifying the images features collection into: grey and colored; and into six main parts as follow: Medical images, Sport images, Cars images, Earth images, People images, and other images.

II. BACKGROUND

The proposed Enhanced Contrast Model and content based image retrieval (CBIR) -image collection classification- was designed to exploit various features of the traditional contrast model. This section gives detailed description of the contrast model and CBIR.

Minkowski in [2, 12], mentioned that in any n-dimensional; the distance between x and y points can be computed using the following formula:

$$D(x, y) = \sqrt[r]{\sum_{i=1}^{n} |x_i - y_i|^r}$$

Where yi and xi are corresponding coordinates of objects x and y on the ith dimension and r is a parameter that determines the type of metric.

A. Similarity Measures and Image Retrieval

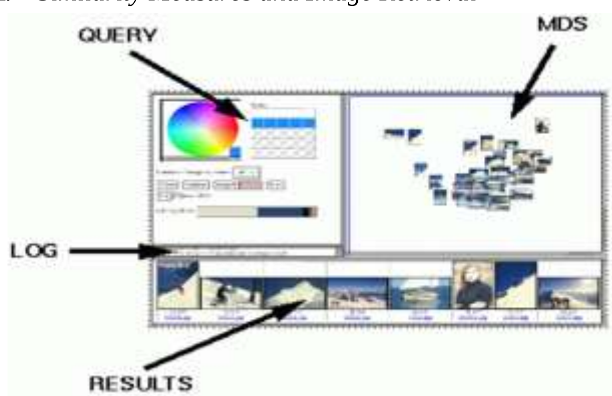

Figure 1. Matching and Similarity in content based image retrieval

As in figure 1, similarity measures (most of which are special cases of the Minkowski metric) are used to determine the similarity between the query image and the images in the database and rank order them. The retrieved sets of images through multidimensional scaling can be visualized by special systems such as content based image retrieval (CBIR) [12].

B. *The General CBIR Computational Framework*

The main trouble considered in this paper is the description of unrestricted submitted images by the user as queries to CBIR system in order to try to retrieve group of images that could be relevant to the images being queried. The retrieval and search procedures depend on the visual characteristics found in the images that included both the image database and the image query set.

Figure 2 depicted the framework of a CBIR system. The whole procedure begins with the building of the images database. A feature extraction algorithm processed the images before adding it to the database. A feature representation can be extracted using this algorithm by storing it into the database and computing the similarity. It is also used to manipulate the queried image and the images stored into the database. Then, the feature representations of the query and the database are compared using the similarity measures. The representations considered as "relevant" are retrieved to the user as a group of similar images.

Query is manipulated and displayed using GUI tools or using. However, the final result of the query description should be like the feature description that is applied by the database to index and save images. The query specification could be used to retrieve the similar images by extracting information from the user who submits the main characteristics of his interest.

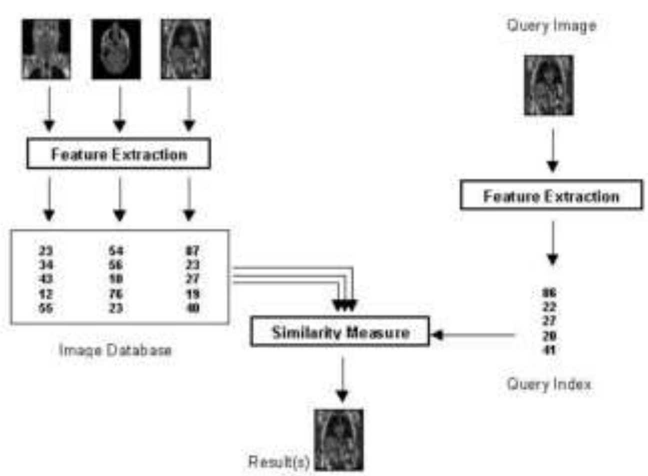

Figure 2. Framework of CBIR systems (Source: [6]).

C. *The Contrast Model (CM)*

People categorize, organize, and group things using the level of their uniformity and isolate them using the level of their diversities or disparity. Two objects considered are similar by several sciences investigations such as: behavioral, cognitive, psychology and other related multidisciplinary since hundred years. During this period of time, a huge number of models and theories have been proposed and experimented to demonstrate perceived similarity. One of these models is the contrast model of Tversky [2]. Tversky challenged the basic assumptions/axioms of the geometric/spatial models of similarity.

In a seminal paper, Tversky [2] mentioned that assumptions and measures of spatial/geometric models of similarity, in addition to formulate and test a set-theoretical model of substitutional of uniformity named the contrast model. According to his model, similarity judgment is a feature contrast task and the level of uniformity between two things is a linear synthesis of their shared and discriminatory features. On the other hand, the model assumes that two things are mostly relevant if they have many shared characteristics and minimum number of dissimilar features and vice versa.

Figure 3 illustrates the relations between two sets. Given two objects x and y and the feature sets of their competent are X and Y, the similarity between them is denoted by s(x,y), which is a linear function that measures the shared and discriminatory features [2], and is formulated by: $S(X,Y) = \theta f(X \cap Y) - \alpha f(X-Y) - \beta f(Y-X)$, where:

- $X \cap Y$ symbolizes the shared features between x and y,
- X-Y symbolizes features which belongs x and not in y (discriminatory features of x),
- Y-X symbolizes features which belongs y and not in x (discriminatory features of y),
- β, θ, and α are the relatively weights which represents the shared and discriminatory features and they are positive numbers,
- S is a range where $S(x, y) > S(z,w)$ if and only if $s(x,y) > s(z,w)$, such that x and y are more similar to each other than the other two objects,
- f function is considered as an additive because $f(X)+f(Y)=f(X \cup Y)$, Thus X and Y are dismantle ($X \cap Y = \emptyset$).

- Where X & Y are non-negative (≥ 0). *-The values of S (X, Y) range from 0 to 1

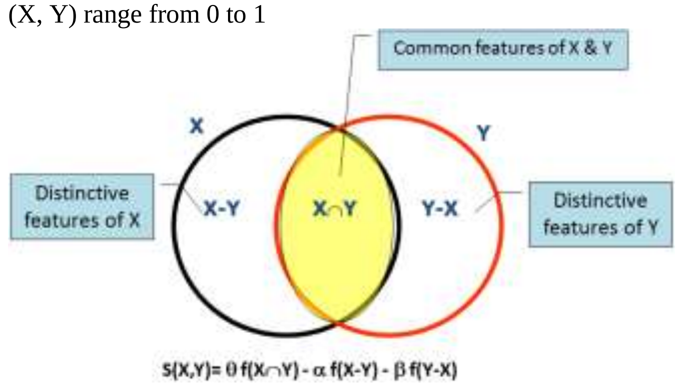

Figure 3. Graphical representation of the relation between two sets in Contrast Model.

According to [2], the experiments on the contrast model showed that the humans concentrate on their awareness more on the shared features when deciding the relevancy than when deciding the distinction of the objects.

III. The Proposed Model (The Enhanced Control Model- ECM)

The contrast model algorithm does not take into account the various types of images (Coloured or grey), and the category of the image (Medical, Sport, People, …., etc.). Thus, figure 4 shows that the enhancement can be done using the enhanced contrast model is by dividing the types of images into: Coloured or Grey. In each type or class the images will be categorized into six suggested categories as follow:

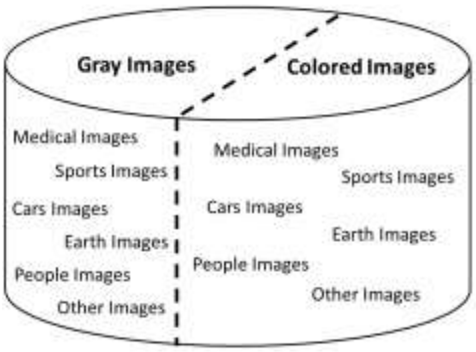

Figure 4. Images Features Database

1. Medical images: all relative images to medical issues.

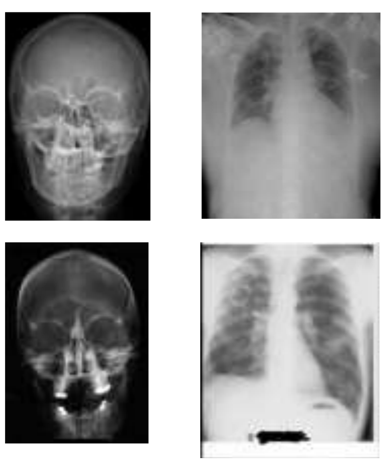

Figure 5. Examples of Medical Images

2. Sport images: any images that relative to sport like (football, basketball…).
3. Cars images.
4. Earth images (moon, stars, sun…).
5. People images.
6. Other images.

Figure 6 shows how the enhanced contrast model works. It firstly check whether the image is gray or colored, secondly try to find the similar images in all categories if it is a medical image or sport or belongs to another categories.

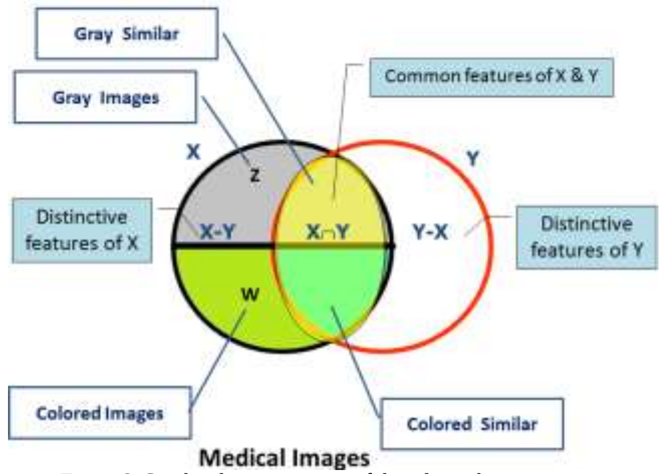

Figure 6. Graphical representation of the relation between two sets in the Enhanced Contrast Model

The similarity equations for the proposed model will be as follow:

(for the grey images)

$$S_g(X, Y) = \frac{(X \cap Y \cap Z)}{(X \cap Y \cap Z)+(X \cap Z)-(Y \cap Z)+(Y \cap Z)-X \cap Z)}$$

Where Z: denoted for the gray images.

(for the coloured images)

$$S_c(X, Y) = \frac{(X \cap Y \cap W)}{(X \cap Y \cap W)+(X \cap W)-(Y \cap W)+(Y \cap W)-(X \cap W)}$$

Where W: denoted for the colored images.

The above equations will be done on all medical images in the database. After that, if the queried image does not match or similar to any image into the medical images, then it will check if the image has similar image(s) into the rest categories (Sport, Car, People, …, etc.). If the queried image has similar image(s) into one category, then the searching process will be stopped and the similar image will be retrieved, which enhances the retrieval process by speeding it up.

Optimally, the proposed system gives users the ability to choose either gray or colored images, this will lend us to save more than 50% of searching time needed to find images from databases. Furthermore, if the user chose one of the six categories that we classify images into, then it will save more than 83% (searching in 1 category of 6 categories) of searching time too, and it will also help in retrieve the most similar ones.

IV. Experimental Results

Generally, precision and recall are the main measures for the retrieval system performance. However, the enhanced contrast model will be evaluated using the CPU runtime in seconds in order to show its efficiency over the classical contrast model.

MPEG database is used in the experiments, which consists of 480 images (240 coloured and 240 grey) that sample of MPEG image database. These images are different sizes from 1-8 MB, and they are categorized into the six classes equally (40 image in each class). The database and query images are matched using the contrast and enhanced contrast models. The experiments are done using Matlab R2010a on PC with 4 GB RAM and core i5 Pentium Processor (2.5GHz).

Table 1 depicts the results produced from the achieved experiments, which demonstrates the efficiency of the enhanced contrast model and average time of all retrieval processes as 0.864 seconds. For all sizes of queried images, these results depict that the performance of the enhanced contrast model is better than the classical contrast model.

The performance of the proposed ECM and CM are tested by firing the same 50 queries (5 queried images with each size: 1, 2, 3, 4, >4 MB).The average of CPU runtime is taken for each 10 queries with image size. The comparison of enhanced contrast model over the classical one is depicted in Table 1. In this table, the execution time classified based on the sizes of the images with the average runtime along 10 queried images with each size. The fourth and fifth columns in table 1 depict the speed up and the average of the speed up for all images' sizes.

TABLE I. COMPARISON OF THE EXECUTION TIME FOR CONTRAST MODEL AND ENHANCED CONTRAST MODEL

Queried Image Size	Average CPU Runtime (s) Contrast Model (CM)	Average CPU Runtime (s) Enhanced Contrast Model (ECM)	Speed Up =CM/ECM	Speed Up Average
~1 MB	2.32	0.487	4.764	
~2 MB	3.13	0.526	5.951	
~3 MB	4.22	0.705	5.986	5.685
~4 MB	5.76	0.913	6.309	
>4 MB	9.15	1.689	5.417	

According to table 1, the enhanced contrast model acquired an average speed up of 5.685 in compare with the classical contrast model when executing under the same circumstances (i.e. database, query, machine, …, etc.).

Figure 7 is a graphical representation for the first three columns of table 1, which shows the enhancement of the retrieval efficiency (Average of runtime CPU in seconds) of ECM over the classical CM with different queried image sizes ranged from 1 to 4 MB.

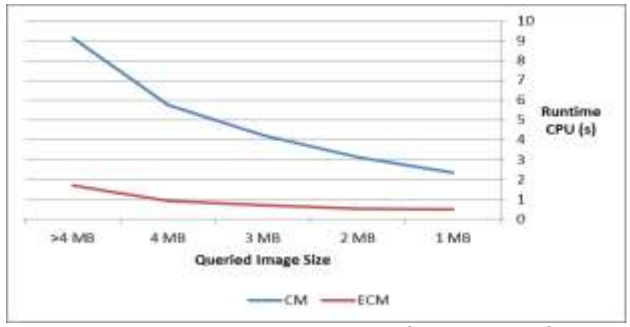

Figure 7. Runtime CPU Comparison between CM and ECM

Figure 8 show a sample of queried image and top 4 retrieved images using the ECM and CM. The precision and recall for the two methods are the same. However, the improvement is achieved by ECM only on the efficiency of the retrieval.

Fig. 8 Sample of feedback of retrieved images

V. CONCLUSION

The main part of information retrieval systems is the measures of matching and similarity, which should be taken into account and considered as a very important factor. The proposed model in this paper improved the contrast model algorithm by classifying the database into several levels, that speedup retrieving process. The proposed model allows users to make an advanced search from the database since they can choose type of image (colored /gray), and also to choose the subject of that image. However, the experiments were done under optimal situation where the numbers of images in each category are the same, and the numbers of images with each size are equal.

REFERENCES

[1] A. Rorissa, Y., "Feature Sets Based Similarity Measures for Image Retrieval". Ph.D. Thesis, University of North Texas, USA, 2005.

[2] A. Tversky, "Features of similarity". Psychological Review, 84(4), 1977.

[3] D. Pukale, H. Varsha, N. Shevate, N. Pawar, and P. Shendage, "Image Matcher – Content Based Image Retrieval System Using Image Sub Blocks and Indexing", International Journal of Computer Science and Information Technologies, Vol. 6 (3), 2015.

[4] G. Srilakshmi, K. Reddy, "Performance Enhancement of Content Based Medical Image Retrieval for MRI Brain Images Based on Hybrid Approach", International Research Journal of Engineering And Technology (IRJET), Vol. 2 (3), 2015.

[5] H. Heidari, A. Chalechale, "Parallel Implementation of Color Based Image Retrieval Using CUDA on the GPU", International Journal of Information Technology and Computer Science,Vol 1., 2014.

[6] J. Zachary, & S. Iyengar, "Information theoretic similarity measures for content-based image retrieval". Journal of the American Society for Information Science and Technology, Vol. 52, 2001.

[7] M. Eisa, "A New Approach for Enhancing Image Retrieval using Neutrosophic Sets", International Journal of Computer Applications, Vol. 95 (8), 2014.

[8] N. Khider, T. Alhasan, "Enhanced of Content-Based Image Retrieval (CBIR)", M.Sc. Thesis, University of Khartoum, Sudan, 2013.

[9] R. Chaudhari, A. M. Patil, "Content Based Image Retrieval Using Color and Shape Features", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 1(5), 2012.

[10] R. Lingadalli, N. Ramesh, "Content Based Image Retrieval using Color, Shape and Texture", International Advanced Research Journal in Science, Engineering and Technology, Vol. 2(6), 2015.

[11] S. Aloni, "An Efficient Approach for Image Retrieval Using Semantic Relevance", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Vol. 3(4), 2014.

[12] Y. Rubner, "Perceptual metrics for image database navigation", Unpublished doctoral dissertation, Stanford University, Stanford, CA, 1999.