ABSTRACT

Content-based image retrieval has been an active area of research over last decade. The goal is to create systems capable of interactively retrieving images that are semantically related to the user's query from a database. In this research, a content-based image retrieval system is presented, it supports querying by example to retrieve images from the images database according to their color and textural low level features.

The underlying techniques are based on the adoption of histograms of *YUV*, *YIQ*, *YC$_b$C$_r$*, *HSV* and *HSI* color models as color features. In addition, *Gray-Level Co-occurrence Matrix (GLCM)* and *Gray-Level Run Length Matrix (GLRLM)* are used to extract two sets of textural features. These features are used for retrieval tasks in separated and combined manners.

The various combinations of the extracted features lead to a large number of possible feature vectors; each describe some aspects of the image contents. The fuzzy concepts have been utilized to reduce the number of possible feature combinations; both the triangular and trapezoidal membership functions have been used for this purpose.

After reducing of possible feature vectors, and keeping those vectors which show high discrimination power, then, all images listed in the database are grouped (according to their feature vectors) by applying *K-means* clustering algorithm. The clustering of feature space is helpful to minimize search time. Euclidean Distance measure is used to assess the similarity (or the distance) between the query image and the images listed in database to find top N similar images.

Given a query image, the system first extracts the selected features to establish the query feature vector, and then compares it with the images belongs to most closest two database clusters. The most similar images are sorted, and then the top 8 images are retrieved as the query results. The

performance of retrieval system has been evaluated using two measurements (i.e., *precision* and *recall*).

The test results indicted that *HSV* and *HSI* color representations are the best when using single (around 0.63 for precision, and 0.1 for recall) or combined (around 0.65 for precision, and 0.1 for recall) types of features. The tests results indicated that color and texture features are not quite enough to retrieve complex images (such as natural images).

الخلاصة

.

.

.( )

.

. .

.

.

.

K-means

.( )

.

. ( )

)

(

( )

.

.( )

HSV   HSI

(0.1   0.63   )

.(0.1   0.65   )

.