Entropy-Based Indexing On Color And Texture In Image Retrieval

Anirban Das

The exponential growth of image data that are being generated makes it imperative to use computers to save, retrieve and analyze images. The problem of image retrieval has been an active area of research since early 70's. In order to make the best use of information in images, we need to organize the images so as to allow efficient browsing, searching and retrieval.

The basic two approaches for image retrieval are text-based and visual-based. Early image retrieval techniques were generally based on textual annotation of images rather than visual features. In other words, images were first annotated with text and then searched using a text-based approach from the traditional database management systems. In the approach using visual features of images, Content-based Image Retrieval (CBIR) focus on the intuitive and efficient methods for retrieving images solely based on the information contained in the images.

Image indexing and its application to content-based image retrieval is a challenging research topic in the field of image processing. In content-based image retrieval system, the user sets the query image by choosing a similar image from the "Sample Query Set" and a query can also be set by identifying the desired object from the segmented query image. The image retrieval system evaluates the similarity of each image in its data store to the query image in terms of textural and color characteristics and returns the retrieved images with ranking within a desired range of similarity.

Several attempts have been made in the recent past to improve the process of image retrieval, but a very few of them are information theoretic. In this research, we have defined a fuzzy entropy function to measure the information in the images. This definition provides more flexibility on the information gain function without violating the properties of entropy function. The fuzzy entropy function is used to extract information from images by optimizing the function through four tunable parameters.

The query paradigm “Query By Object” (QBO) provides the segmented image to the user for identification of the object of interest. The image segmentation and object identification are very important part of image retrieval system. We have applied Improved Mountain Clustering technique for color segmentation and we apply the same clustering technique on the entropy values obtained from the fuzzy response of the pixel for the purpose of texture segmentation.

Feature extraction is very crucial step in image retrieval system to describe the image with minimum number of descriptors. The basic visual features of images include color and texture; we have attempted to capture the color information by its color entropy. We have devised a fuzzy-based method to compute the fuzzy response of each pixel due to its neighborhood texture pattern and we calculate the texture entropy from fuzzy response. The underlying technique of the image retrieval system is based on the adaptation of an entropy-based approach to color and texture analysis. An optimal set of three features - color entropy, texture entropy, and the dominant color of image are extracted so as to render the feature vector for each image maximally informative, and yet to obtain a low vector dimensionality for efficiency in computation. Furthermore, we consider the entropy-based color and texture entropies as constituting the feature vector for indexing the images. The indexing scheme is effectively used to generate an intermediate result set for any query image and this result set contract the search space by including only similar images.

We propose a similarity measure between images using color entropy, texture entropy and the CIELab value of dominant color. In defining similarity, the L_2 distance between the entropy values and L_2 distance between the dominant colors are considered separately and they are tied by linear combination. These distance/similarity measures are used to rank the retrieved images. It is shown through experimental results and precision-recall analysis that the indexing technique and content-based retrieval system are effective in terms of retrieval and scalability.