ABSTRACT

Shape is one of the most important visual features that characterizes an object and enable human beings to recognize it. In Content-based image retrieval, shape representation and description techniques are used to retrieve selected images effectively from large databases.

Availability of high speed processors, high capacity storage devices and high speed internet have resulted in large number of images being generated, transmitted and stored from across the globe in digital form. There is a growing interest in searching for a particular image in large collections of image databases. In order to search and retrieve a matching image, these images need first be represented and described based on certain features. Searching for images using shape features is an active research area. Many shape representation and description techniques are discussed in the literature. Shape representation and description techniques are broadly classified into two: Region based and Contour based. Region based systems typically use descriptors based on the regional or internal content. Boundary-based systems use the contour of the objects and usually give better results for images that are distinguishable by their contours.
CBIR system extracts the images that match the specified query image based on the image content. This research work proposes three novel shape based techniques for the representation and retrieval of images. Morphological contour representation algorithm for the representation and reconstruction of 2D binary shapes of digital images is proposed first. This algorithm uses representative lines of different sizes and types to cover all the significant features of the binary contour image. This is more efficient than that of most other shape representation algorithms including Morphological Skeleton Transform (MST) and Morphological Shape Decomposition (MSD).

Two other algorithms for image representation that enable effective Content Based Retrieval have also been proposed. Scale, Translation and Rotation invariance have also been addressed. Benchmark database used was that of Wang.

A novel similarity measure based on convex hull geometry (Convex Hull Area Ratio – CHAR) is proposed, to retrieve the exact match. This method which preserves scale and translational invariance employs an approach of comparing the Convex Hull geometry of the query image to that of the database image in terms of a relative metric which is denoted as the Convex Hull Area Ratio (CHAR). CHAR is the ratio of the area of intersection of the two convex hulls to the area of their union. Similarity measurement is performed and the maximum value of the CHAR indicates the closest match. Thus, the database images that are relevant to the given query
image are retrieved. Scale and translational invariance have been preserved by a suitable co-ordinate transformation.

Development of an effective algorithm that would preserve rotational invariance was also looked upon. Zhang and Lim in 2007 proposed a technique for region based shape representation and retrieval. In their method, the binary value of shape is sampled at the intersections of the circles and radial lines, which are overlaid at the centre of mass of the shape. The drawback of Zhang’s technique is that the shape points on the intersection are only considered. The number of concentric circles and radial lines on the polar grid need to be quite high, in order to get adequate sampling for a good match.

A new algorithm PRESS (Polar Raster Edge Sampling Signature) that forms an effective shape signature by taking into account all the prominent edge points by collecting them in specified number of radial and angular bins was also proposed. Scaling and rotational invariance are preserved in the CBIR system implemented using PRESS. In this technique, the extracted shape features are polar raster scanned into specified intervals in both radius and angle, using the proposed Polar Raster Edge Sampling Signature (PRESS) algorithm. Counts of edge points lying in these bins are stored in the feature library. When a query image is passed on to the system, the features are extracted in the similar fashion. Subsequently, similarity measure is performed between the query image features and the database image features.
and the database images that are relevant to the given query image are retrieved.

CHAR and PRESS algorithms have been successfully implemented and tested in a CBIR System developed for this purpose. The retrieval efficiency is evaluated by determining Precision-Recall values.

The main contributions of this research work are techniques for Morphological based Contour Representation and Reconstruction, Convex Hull Geometry based Representation and Retrieval and Polar Raster Sampling based Representation and Retrieval. A general purpose Content Based Image Retrieval System was also developed and the same was used to test retrieval efficiency of the proposed techniques.