ABSTRACT

With the large multimedia databases becoming a reality in various domains, methods for organizing image databases and for efficient retrieval from them have become important. Conventional text based techniques have been used for image retrieval for many years. It has certain drawbacks such as language limitation and subjectivity of human perception. Furthermore, annotation process of these images is quite time consuming. Text is man's creation, while typical images are a mere replica of what man has seen since birth, interpretation of what we see is hard to characterize. This motivates the need of image retrieval based on the visual content of an image.

Content-based image retrieval (CBIR) is the process of retrieving desired images from a large collection on the basis of the visual contents of an image (such as color, texture, shape and spatial layout) that can be automatically extracted from the images.

Currently, most of the CBIR systems are based on the centralized computing model. Some are stand-alone applications while others are web based systems. A centralized system maintains central nodes to handle the query requests. It keeps the entire feature descriptor database in a centralized server. Upon retrieving the relevant images according to feature similarity measures, the content will be transferred directly from the content server to the requesting host. The drawback of the centralized system is its limited scalability for handling growing volumes of retrieval requests and large image databases.

In order to provide better scalability, we have proposed a decentralized system based on peer to peer (P2P) paradigm. Each node in a P2P network acts as a both, client for requesting image and a server for redistributing the images [23]. In P2P file systems, files are stored at the end user machines (peers) rather than on a central server and, as opposed to the traditional client-server model, files are transferred directly between peers. In general, P2P systems allow a decentralized sharing of distributed computational resources and contents of individual peers, which join and leave the network frequently. In P2P-CBIR, a peer submits a query image and the system retrieves the most similar images to the query image from across all peers. We foresee the advantages of using the P2P network for CBIR in several ways. First, with an increased number of users joining the

P2P network, the image collection will grow drastically due to individual contributions. This gives diversity and variety. Second, it overcomes the scalability problem of image retrieval by using a decentralized retrieval approach. Furthermore, the storage, information and computational cost can be distributed among the peers, allowing many individual computers to achieve higher throughput.

In this research, we have proposed an efficient CBIR system which works seamlessly in a P2P environment. Mainly, we focus on five main areas in the image retrieval problem and these are:

1. Presenting a new moment preserving technique which takes image color distribution into consideration during color feature extraction process.
2. Proposing an efficient algorithm that preserves the geometric consistency during image retrieval.
3. Developing a new relevance feedback approach based on uncertainty based sampling to further enhance the retrieval accuracy during successive user feedback.
4. Proposing a novel architecture for a P2P-CBIR system for decentralized sharing of distributed computing resources.
5. Developing a peer clustering and intelligent query routing strategy to handle query routing and to speed up the image retrieval task in a P2P network.

Our experimental results on COREL photo database of 10 thousand images show that the proposed feature extraction technique with FAST RANSAC algorithm significantly enhances the retrieval accuracy and saves about 17.43% of the time that conventional RANSAC spends during retrieval process. To further improve the retrieval results, we have presented a novel relevance feedback strategy based on uncertainty based sampling which enhances the average retrieval precision rate by a factor of 12% over that of a conventional relevance feedback loop. The proposed P2P-CBIR system is then verified by simulations with different parameters to investigate the performance change which is subjected to different network size and time To Live (TTL) of query messages. It is observed that the proposed system works seamlessly in Peer to Peer environment.