CHAPTER 2

IMAGE RETRIEVAL PROCESS

CBIR is a process of concerned with retrieving the desired images from a large collection on the basis of features (such as color, texture and shape) that can be automatically extracted from the images themselves. Image processing is any form of signal processing for which the input is an image, such as photographs or frames of video. The output of image processing can be either an image or a set of characteristics or parameters related to the image. Most image processing techniques involve treating the image as a two-dimensional signal and applying standard signal processing techniques to it. This chapter presents the details on image retrieval process and related concepts.

Users in many professional fields are exploiting the opportunities offered by ability to access and manipulate remotely stored images in all kinds of new and exciting ways. The problems of image retrieval are becoming widely recognized, and the search for solutions an increasingly active area for research and development. There are two types of approaches for image retrieval namely text based approach and content based approach. Text based image retrieval uses traditional database techniques to manage images. Through text descriptions, images can be organized by topical or semantic hierarchies to facilitate easy navigation and browsing based on standard Boolean queries. Figure 2.1 presents a sample architecture for text
based image retrieval and Figure 2.2 presents the architecture for image based retrieval.

![Diagram of Text Based Image Retrieval](image1)

**Figure 2.1 Text Based Image Retrieval**

![Diagram of Content Based Image Retrieval](image2)

**Figure 2.2 Content Based Image Retrieval**

Most text based image retrieval systems require manual annotation of images, since automatically generating descriptive texts for a wide spectrum of images is not feasible. Obviously, annotating images manually is a cumbersome and expensive task for large image databases, and is often subjective, context-sensitive and incomplete. Content-based image retrieval
Content Based Image Retrieval (CBIR) is the set of techniques for retrieving semantically-relevant images from an image database based on automatically derived image features such as color, texture and shape. Content based means that the search will analyze the actual contents of the image. The term content in this context might refer to colors, shapes, textures or any other information that can be derived from the image itself. The need for efficient content-based image retrieval has increased tremendously in many application areas such as military, education and web image searching.

Content Based Image Retrieval (CBIR) also known as Query By Image Content (QBIC) and Content Based Visual Information Retrieval (CBVIR). It is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. Without the ability to examine image content, searches must rely on metadata such as captions or keywords, which may be laborious or expensive to produce. Indexing is the attachment of a signature to an image. This index is external information standing for the image in the retrieval, so its relevance is crucial for the effectiveness of the search. Signature is a structured representation of the image, and is used as an index (it enables searches in a set of images). A histogram is a standard statistical description of a distribution in terms of occurrence frequencies of different event classes; for color, the event classes are regions in color space. An image histogram of scalar pixel values is more commonly used in image processing than is a color histogram.

Trademark Image Retrieval (TIR), a branch of Content-Based Image Retrieval (CBIR), plays an important role in multimedia information retrieval. The work proposed by the authors finds an effective solution by combining shape description and feature matching to improve effectiveness of retrieval process. This is done by effective shaping methods which includes
two shape descriptors. Then dissimilarity value between the feature vectors extracted from images is calculated. Finally, the shape description method and the feature matching strategy were combined to realize the proposed solution. The experiments were conducted on a large and standard image set of images (Heng et al. 2010).

Image texture, defined as a function of the spatial variation in pixel intensities (gray values), is useful in a variety of applications. One application of image texture is the recognition of image regions using texture properties. In the recent years, human interactive systems have given more attention using fuzzy relevance feedback (Kim Hui & Kui 2005). Contrary to the early systems, which focused on fully automatic strategies, recent approaches have introduced human-computer interaction. This system focuses on the retrieval of concepts within a large image collection. Here a user is assumed to look for a set of images, the query concept, within a database. The aim is to build a fast and efficient strategy to retrieve the query concept.

In Content-Based Image Retrieval (CBIR), the search may be initiated using a query. The top rank similar images are then presented to the user. Then, the interactive process allows the user to refine his request as much as necessary in relevance feedback iterations. To achieve the feedback process, first strategy is to focus on query concept updating. The aim of this strategy is to refine the query according to the user labeling. The user information consists of binary labels indicating whether or not the image belongs to the desired concept. The positive labels indicating relevant images of current concept and negative labels indicating irrelevant images. There are various approaches for refining the query such as query modification which computes new query by averaging feature vectors of relevant images, query reweighting which computes new similarity function between query and any picture in the database. In order to perform better refinement of similarity
function, optimization techniques can be used. They are based on mathematical criterion for computing the reweighting.

Performing an estimation of the query concept can be seen as a statistical learning problem, and more precisely as a binary classification task between the relevant and irrelevant classes. In image retrieval, many techniques based on statistical learning have been proposed. k-Nearest Neighbors, Naive Bayes Classifier, Gaussian mixtures and Gaussian random fields. In order to deal with complex and multimodal concepts, a statistical learning approach have adopted here. It focuses on statistical learning techniques for interactive image retrieval.

2.1 REPRESENTATION OF IMAGES

Images are represented in digital form as shown in Figure 2.3. Generally speaking, image content may include both visual and semantic content. Visual content can be very general or domain specific. General visual content include color, texture, shape, spatial relationship, etc. Domain specific visual content, like human faces, is application dependent and may involve domain knowledge. Semantic content is obtained either by textual annotation or by complex inference procedures based on visual content.

![Figure 2.3 Image Content Descriptors](image.png)
A good visual content descriptor should be invariant to the accidental variance introduced by the imaging process (e.g., the variation of the illuminant of the scene). However, there is a tradeoff between the invariance and the discriminative power of visual features, since a very wide class of invariance loses the ability to discriminate between essential differences. Invariant description has been largely investigated in computer vision (like object recognition), but is relatively new in image retrieval. A visual content descriptor can be either global or local. A global descriptor uses the visual features of the whole image, whereas a local descriptor uses the visual features of regions or objects to describe the image content. To obtain the local visual descriptors, an image is often divided into parts first. The simplest way of dividing an image is to use a partition, which cuts the image into tiles of equal size and shape. A simple partition does not generate perceptually meaningful regions but is a way of representing the global features of the image at a finer resolution (Heng et al. 2010). A better method is to divide the image into homogenous regions according to some criterion using region segmentation algorithms that have been extensively investigated in computer vision. A more complex way of dividing an image, is to undertake a complete object segmentation to obtain semantically meaningful objects (like ball, car, horse).

### 2.2 FEATURES FOR IMAGE

The features of image and its role in information retrieval process is given below in this section.

#### 2.2.1 Color Histogram

The color histogram serves as an effective representation of the color content of an image if the color pattern is unique compared with the rest of the data set. The color histogram is easy to compute and effective in
characterizing both the global and local distribution of colors in an image. In addition, it is robust to translation and rotation about the view axis and changes only slowly with the scale, occlusion and viewing angle. Since any pixel in the image can be described by three components in a certain color space (for instance, red, green, and blue components in RGB space, or hue, saturation, and value in HSV space), a histogram, i.e., the distribution of the number of pixels for each quantized bin, can be defined for each component. Clearly, the more bins a color histogram contains, the more discrimination power it has. However, a histogram with a large number of bins will not only increase the computational cost, but will also be inappropriate for building efficient indexes for image databases.

2.2.2 Color Correlogram

The color correlogram was proposed to characterize not only the color distributions of pixels, but also the spatial correlation of pairs of colors. The first and the second dimension of the three-dimensional histogram are the colors of any pixel pair and the third dimension is their spatial distance. A color correlogram is a table indexed by color pairs, where the k-th entry for (i, j) specifies the probability of finding a pixel of color j at a distance k from a pixel of color i in the image.

2.2.3 Texture

Texture is another important property of images. Various texture representations have been investigated in pattern recognition and computer vision. Basically, texture representation methods can be classified into two categories: structural and statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular. Statistical methods, including
Fourier power spectra, co-occurrence matrices, Shift-Invariant Principal Component Analysis (SPCA), Tamura feature, Wold decomposition, Markov random field, fractal model, and multi-resolution filtering techniques such as Gabor and wavelet transform, characterize texture by the statistical distribution of the image intensity.

2.2.4 Feature Extraction

The performance of the CBIR system usually depends up-on the features adopted to represent the images in the data-base. The proposed CBIR technique will use both the human perception as well as machine level perception. Proposed system also uses a recently retrieved image library for the retrieval of the system. The proposed Image retrieval system consists of two steps namely feature extraction and retrieval phase. We focus on considering four popular features, namely, shape, texture, color and homogeneity features. The shape, color, texture and homogeneity features are the low level features used in this CBIR for retrieval. The image features are either extracted from the whole image or from the regions. As it is found that the users are mostly interested in specific region as compared to the entire image, the pro-posed system extracts shape, color and texture features region wise. After the completion of the feature extraction, the query image features compared with the database image features and recently retrieved image library.

2.2.5 Low Level Feature Extraction

The visual contents of an image are analyzed by using the low level features such as shape, texture and color of the image. As it is found that the users are mostly interested in specific region as compared to the entire image, the pro-posed system extracts shape, color and texture features region wise.
2.2.6 **Shape feature extraction**

An efficient and robust representation of shape feature plays an important role in image retrieval. These features should also be independent of different characteristics such as translation, rotation, and scaling of the shape. To extract the shape feature from the image, initially, the image in RGB color space is converted to gray scale image. RGB color is a format for color images and it represents an image with three matrices of sizes matching the image is reduced to 128. The LGB vector quantization algorithm is used in pro-posed technique to obtain the set of different colors which will represent image colors in lab space (with respect to mean square error).

2.3 **LEVELS OF RETRIEVAL**

Content based image retrieval has different levels. Level 1 is categorized based on color, texture and shape features. Here the images are compared based on low-level features and no semantics involved. Level 2 deals with the semantics. Level 3 deals with retrieval with abstract and subjective attributes. In contrast to the text-based approach, CBIR operates on a totally different principle, retrieving stored images from a collection by comparing features automatically extracted from the images themselves. The commonest features used are mathematical measures of color, texture or shape; hence virtually all current CBIR systems, whether commercial or experimental, operate at level 1. A typical allows users to formulate queries by submitting an example of the type of image being sought, though some offer alternatives such as selection from a palette or sketch input. The system then identifies those stored images whose feature values match those of the query most closely, and displays thumbnails of these images on the screen.