2. Review of Literature

2.1 Overview

Due to increase in the use of digital images in almost all fields it needs efficient mechanism to store, manage and retrieve required images from the database. The text based image retrieval most of the time provides the ambiguous retrieval result which needs to be worked out with alternative. That is why it is always preferred to make use of the inherent content of the image itself to extract its unique characteristic. This unique characteristic can be used for matching the query image in the database this whole scenario leads us towards Content-based image retrieval (CBIR) [1-52]. CBIR is a technology that basically assists to organize digital picture archives by their visual content. By this definition, an image similarity function to a robust image annotation engine falls under the purview of CBIR. The CBIR system makes use of similarity measuring methods [1-52,196-197] to match the query image feature with the feature database. The term Content-Based Image Retrieval (CBIR) seems to have originated by T. Kato in 1992. [2] Since then, the term CBIR has been used to describe the process of retrieving desired images from a large collection on the basis of syntactical image features.

The techniques, tools and algorithms that are used in CBIR originate from various fields such as statistics, pattern recognition, signal processing, and computer vision. Due to the wide use of digital images in various fields’ experts of these fields such as, computer vision, machine learning, information retrieval, human-computer interaction, database systems, Web and data mining, information theory, medical images, satellite images etc. are contributing to CBIR [1]. There are various CBIR application [12] systems like Image Rover [15], VisualSEEK [16] have been developed by various researchers. The online search and retrieval of digital images for World Wide Web have been developed by many researchers. These systems are based on the textual and visual content of images [29].

The CBIR finds its wide application in various biometric systems [53, 64] like face recognition [56-60, 63, 66, 68-71, 103, 188], finger print recognition [53-55, 61, 62, 64, 81, 190, 191, 193], iris recognition[72-79, 200, 201], palm print recognition [121, 192], speaker recognition [67, 102, 118, 162,
163, 187, 189], finger Knuckle recognition system [107], signature recognition system [65, 106, 122-124]. Apart from this CBIR is very much applicable in detecting tumor in the medical images [99,127-130] of the body part taken through MRI scan, CT scan etc. The medical field finds the use of CBIR in mammography [82, 108-117], MRI Images [83-84] etc. The CBIR can be utilized for the searching the required trademark from the large database of trademarks [24-28], video content analysis [21], and database indexing [23].

The CBIR basically focuses on the inherent features of digital images. These features are color, shape and texture. Researchers around the world are working and have developed CBIR system based on these three features: color [3, 8, 18, 32, 34, 35, 37, 38], shape [25-28,204] and texture [48, 51,185]. As far as the color feature is concerned there are various color spaces [164-184] like Kekre’s LUV Color space [32], YCbCr [34, 37] etc are used to extract colors of the image by various approaches like row mean, column mean, diagonal mean of color components[49] present in the image. There are many researchers who have used the color histogram [38], color bins for the same to extract the color feature for CBIR system.

The object shapes like curve, contour, boundary, edges, lines and points available in the image can be used as the feature of the image. These object shapes are extracted using various high pass filters like canny edge detector, Prewitt, sobel edge detector. These extracted shapes are stored in the database by means of histogram information or shape number.

The texture pattern: There is a wide variety of features used for texture analysis. These uses coarseness, contrast and directionality. These textures can be enhanced by co-occurrence matrix features. Textures can be described in terms of periodic, oriented and random components. Gabor filters or Gaussian derivatives are often used for describing texture .The use of Rotation invariant Gabor filters for extracting the texture feature for a remote sensing information retrieval system. Various mathematical transforms can be applied to extract the texture patterns present in the image.

The image feature extraction for CBIR has been proposed by various researchers in both spatial domain [119] and frequency domain [52, 86-
The spatial domain includes the methods like BTC [31-37], Thresholding [39], Hashing [40], Energy compaction [41], Color averaging [49], Image tiling [41], Row mean, column mean variances [46], Histogram [22, 27, 35, 38], Vector quantization [42]. The feature extraction techniques in the frequency domain makes use of various mathematical transforms like DCT [51,87-89], DST, Walsh [50,90-91], Kekre transform[17], Slant transform [94], Hartley transform, DFT and wavelets like Haar[146], Harrlet [200], Kekre wavelet[146, 192] and plane sectorization [133-157]. We have found that the various methods implemented for extracting the texture features in order to get the good retrieval is limited. We have proposed new methods of conceptualizing the planes of transformed images and their wavelets. These newly conceptualized planes are sectored in various sectors to extract image features. The Figure 2.1 Shown below shows the area of focus (blocks shaded with red color) for the CBIR system we have proposed in this thesis.

![Image feature extraction for CBIR](image)

**Figure 2.1. The proposed CBIR system**

Following section discusses the various existing approaches of CBIR system implemented by various researchers in both spatial and frequency domain.

### 2.2 Existing systems.

#### 2.2.1 Spatial domain [160,161]
A. Block truncation coding (BTC) [30-37,47,194,203]

The extraction and retention of the visual information of the image has been the focus of most of the image coding methods developed so far. The information such as DCT coefficients of JPEG can be used for indexing of images and various recognition systems. However since JPEG and other similar methods are not explicitly designed for image indexing purpose so models and features have to be derived from the complicated and computational complex transform coefficients, causing the expansion of data. For example, even though color is an excellent cue for image indexing. However it is difficult to exploit color information from the transform coefficients without decoding. On the other hand, non-transform based image coding can have image features such as color more easily available. Block truncation coding (BTC) is a relatively simple image coding technique developed in the early years of digital imaging more than 29 years ago. [32].

Block truncation coding (BTC) was first developed in 1979 for grayscale image coding. The method divides the image into R, G, and B components. Then computes inter band average image (IBAI) which is the average of all components (R, G, and B) and mean of inter band average image is taken as threshold. The bitmap is then created by comparing each pixel with this threshold value. If a pixel in the inter band average image is greater than or equal to the threshold, the corresponding pixel position of the bitmap will have a value of 1 otherwise it will have a value of 0. Two mean colors one for the pixels greater than or equal to the threshold and other for the pixels smaller than the threshold are also calculated [32].

Boosting this BTC based CBIR as BTC-LUV and Spatial BTC-LUV is done in [32]. Here Kekre’s LUV color space is considered. In BTC-LUV feature vector is computed by considering L, U and V planes of the image independently. While in Spatial BTC-LUV, the feature vector is composed of four parts. Each part is representing the features extracted from one of the four non overlapping quadrants of the image. These methods are tested on the database of 1000 images and the results show that the precision is improved in Spatial BTC-LUV and recall is better in BTC-LUV. If both precision and recall are considered together Spatial BTC-LUV outperforms the other methods discussed in the paper [32]. Other techniques like BTC With Local Average Thresholding [31], Augmented Block Truncation Coding
Techniques[33], Amendment of Block Truncation Coding with YCbCr Color Space[34] has been experimented by various researcher along with this; BTC has been used for extracting and using the color histograms [35] and Multileveled BTC using Sundry Color Spaces has been tried in CBIR[36,37]. The concept of BTC can be used with various approaches like Gradient Operators and Slope Magnitude Technique for Shape Feature Extraction for CBIR [47] and Fusion of Gabor Magnitude [194].

B. Thresholding[31,39]
Using BTC with Local Average Thresholding [31], Bins using Global and Local Thresholding of Images [39].

C. Color averaging[49]
The paper [44] presents six innovative content based image retrieval (CBIR) techniques based on color averaging. The color averaging methods used by the researcher in this paper are row mean, column mean, forward diagonal mean, backward diagonal mean, row and column mean and forward and backward diagonal mean. The proposed CBIR techniques in this work are tested on generic image database consisting of 1000 images of 11 classes and COIL image database consisting of 1080 images of 15 classes. [44], the concepts of Row mean, column mean and diagonal mean consideration have been shown in the Figure given below.

**Figure 2.2. Sample Image Template (with size nxn) showing row & column mean vector [44]**
The Figure 2.2 shows the steps followed to calculate the row mean and column mean method given in the work [44]. The row mean is generated by means of calculating the average of all n pixels (R, G, and B) present in each row and stored in the separate vector. Similarly the column mean is calculated by taking the average of each column separately and stored in the separate vector. These newly generated vectors later on are used by the author to generate the feature vector.

![Figure 2.2](image)

Figure 2.3 Sample Image Template (with size nxn) showing forward diagonal mean vector [44].

The Figure 2.3 focuses on the sample image template of size nxn showing the forward diagonal mean vector calculation procedure in detail. The average of each color component of particular pixels (taken in the diagonal fashion) in the image has been taken into the consideration for calculating the mean. These average values are used later on for generating the feature vector for CBIR.

Following are the various techniques applied by the author in this work [44].

**A. All Image Coefficients**

In this approach all image pixels are considered as feature vector and Euclidean distance is used as similarity measure in RGB plane to find the best match.
B. **Row Mean of Image (RM)**
In this method row mean of image is calculated (as depicted in the Figure 2.2 given above) and used as feature vector. Euclidean distance is used in RGB plane to find the best match.

C. **Column Mean of Image (CM)**
This method advocates to have feature vector composed of column mean of pixel color values of image (Shown in the Figure 2.2 given above). Euclidean distance is used as similarity measure in RGB plane to find the best match.

D. **Row & Column Mean of Image (RCM)**
This approach considers both row and column mean of image together as feature vector. Euclidean distance is used as similarity measure in RGB plane.

E. **Backward Diagonal Mean of Image (BDM)**
In this approach backward diagonal mean of color components present in the image is considered as feature vector. Euclidean distance is used as similarity measure in RGB plane to find the best match.

F. **Forward Diagonal Mean of Image (FDM)**
Author has considered the forward diagonal mean of color value of the pixel present in the image for feature vector generation (as shown in the Figure 2.3 as given above). Euclidean distance is used in RGB plane to find the best match.

G. **Forward & Backward Diagonal Mean of Image (FBDM)**
In this part author has used the combination of both forward and backward diagonal mean of color values of each pixels in the image for feature vector generation.

These all techniques of color averaging have been tried in various manners like Amelioration of the color averaging with the help of Even and Odd image parts [49]. The approach of row mean and column mean can be tried on various approaches of the pixel selection like image tiling [41] which is discussed as below.
D. Image tiling[41]

The image splitting divides the image into non overlapping parts. Then row mean and column mean of each part are obtained. After applying transform on these, feature sets can be obtained to be used in image retrieval. The work presented in the paper [41] considers 1, 4, 16 and 64 non overlapping parts as exemplified in Figure 2.4 given below.

The Figure 2.4 shows that a single image splitted in 64 non overlapping parts. Similarly any given image can be splitted/tilled into number of non overlapping parts. Later on this can be utilized to generate the feature vector.

Continuing further with the concept of exploiting the row and column it has also been experimented with the statistical methods like standard deviation and variance which is given as below.

E. Standard Deviation of Row mean, column mean variance[46,205]

This approach to content-based image retrieval (CBIR) represents each image in the given database by a vector of feature values extracted is nothing but “Standard deviation of average vectors of color distribution of rows and columns of images”. The paper [205] discusses the feature extraction for the texture descriptor that is ‘variance’ and ‘Variance of Variances’. First standard deviation of each row and column mean is calculated for R, G, and B planes. These six values are obtained for one
image which acts as a feature vector. Secondly variance of the row and column of R, G and B planes of an image has been calculated. Then six standard deviations of these variance sequences are calculated to form a feature vector of dimension six.

- **Basic Algorithm**
  This approach calculates 'Standard deviation of variance' of rows and columns rather than calculating the row mean and column mean values. The Euclidean Distance is used as primary similarity measure to calculate the distance between the query image and images in the given database. Threshold values are determined by taking 10%, 20%, 25%, 30% and 40% of maximum Euclidean distance. The steps followed by the author are given below.

- **Implementation Details.**
  2. Calculate the variance of rows and columns of each plane of an image.
  3. Calculate the ‘standard deviation of variance’. (Whatever sequences we obtained in above step, i.e. sequences of variance for rows and columns per R, G, and B planes) calculate the standard deviations which give the feature vector of six dimensions.
  4. Calculate the Euclidean Distance between the query image and images in the database.
  5. Select those images where the distances are less than a preselected value threshold $T$ obtained by taking 25% of the maximum Euclidean distance.

The outcome of the experiments depicts the better performance of Precision and Recall both for standard deviation of Mean sequence than the standard deviation of variance sequences.

Apart from this approach of row and column values there are other approaches like use of histogram to draw out the important information from the image. This information can be used to generate the unique feature vector for better performance in CBIR. The methods experimented based on the histogram plotted for the images are discussed below.

**F. Histogram.**
The use of Histogram as the technique to extract the essential information from the given image for generating the feature vector has been very useful
in CBIR. Most of the research done in the spatial domain makes use of this approach in various ways. Some of these techniques are using directional detail histograms[22], Relational histograms for shape indexing[27], Color Histogram’s Moments based Features[35], color histogram processing[38] and accessing histogram Bins using Global and Local Thresholding of Images[39].

The extraction of the shape features from the image and using that along with the histogram has been tried in [27]. This work uses two dimensional histograms that encode both the local and global geometric properties of the shapes. The pair wise attributes considered in this work are the directed segment relative angle and directed relative position. This approach advocates the simultaneous use of the relational and structural constraints, derived from an adjacency graph, to find the histogram contributions. The retrieval capabilities of the method suggested in this work has been investigated for various queries. The experiment suggested in this work investigates the robustness of the method to segmentation errors as well. The observation made by the researcher in this work is that relational histogram of pair wise segment attributes presents a very efficient way of indexing into large databases. The optimal configuration is obtained when the local features are constructed from six neighboring segments pairs. Moreover, a sensitivity analysis reveals that segmentation errors do not affect the retrieval performances [27].

- **Color Histogram [38]**

  The work presented by the one of the researcher in paper [38] makes use of color extraction and its comparison using the three color histograms, conventional color histogram (CCH), invariant color histogram (ICH) and fuzzy color histogram (FCH). The conventional color histogram (CCH) of an image indicates the frequency of occurrence of every color in an image. The appealing aspect of the CCH is its simplicity and ease of computation. There are however, several difficulties associated with the CCH focused by the author of the paper. The problems are as follows. The first of these is the high dimensionality of the CCH, even after drastic quantization of the color space. Another problem with the CCH is that it does not take into consideration color similarity across different bins and cannot handle rotation and translation. To address the problem related to the rotation and translation researcher has proposed an invariant color histograms (ICH)
based on the color gradients is used and to address the problem of spatial relationship fuzzy linking color histogram (FCH) is used.

- **Histograms Bins [39]**
  The experiment presented by the author of the paper [39] makes use of the color histogram plotted for the given image. The given image is separated in R, G and B planes and then each plane is separated into two parts (that is intensities above average and below average) by applying local and global Thresholding. This helps to obtain six different planes. For every pixel in the image Red, Green and Blue values are checked to find out whether it falls in the plane below the average or above the average and assigned flags to form bins which holds the count of pixels of the particular category. These bins are used as feature vector components of an image of size 8. The Euclidian distance has been used as the primary similarity measure to calculate the distance between feature vectors of database images and the query image. The threshold value is obtained based on which similar images can be retrieved. The suggested approach has been applied to an image database consisting of 525 BMP images which includes 75 Flower 75 Barbie 75 Mountains and 75 Sunset, 75 Puppy, 75 Parrot and 75 Mickey images [39]. The implementation detail of the algorithm given by the author is explained below.

**Formation of bins and Feature Vector Extraction [39]**
Each pixel of the image for which the feature vector is being extracted, its Red, Green and Blue values are compared with the Local threshold (R,G,B) or Global threshold (R+G+B)/3 values to check whether it is above or below the threshold. If the value is less than the respective threshold then the flag is assigned with the value ‘0’ to it else flag is set to ‘1’. As three flag values (R_flag, G_flag and B_flag) to be set for each pixel either to ‘0’ or ‘1’ based on criteria whether the pixel value falls below or above threshold 8 different combinations of these three flags (R_flag, G_flag and B_flag) with two values(0 and 1), from 000 to 111 are formed. These eight bins formed can be used to extract the feature vector of the image in it. These three flags for each pixel of the image are checked. For instance If R_flag = 0, G_flag = 0 and B_flag = 0 then that pixel goes in bin0, if flag values are 001 respectively then it goes in bin1, if it is 010 then it goes to bin2 likewise eight bins are formed from bin 0 (000) to bin 7 (111) holding the pixel counts of eight different
categories. This is how feature vector extracted of all images to be compared in these eight bins.

- **Implementation details given by the author [39].**
  2. Find local average of each component as calculating the average value of the each plane i.e. R, G and B. (R_avg, Gavg, and Bavg).
  3. Split every image component using local average to obtain R_above, R_below, G_above, G_below and B_above, B_below.
  4. If the pixel falls in the category above average value as checked in step 3 then set the flag to ‘1’ to it else set the flag to ‘0’. 
  5. The steps 1-4 are repeated for all pixels of each plane and eight bins are formed.
  6. The number of pixels in each bin gives the feature vector of 8 components.
  7. The primary similarity measure used in Euclidian distance.

- **Image Retrieval Using BINS approach (Global Thresholding)**
  Similar to the above algorithm (Bins formation using local Thresholding), the global average of the three components (R, G, B) of the image are calculated. Rest of the steps followed for bin formation and feature vector generation are same as followed for local Thresholding approach. The system is tested for different query images and the set of retrieved images obtained as a result are with the distance lesser than the preselected value of threshold.

- **Determining the threshold value**
  The Euclidean distances obtained for both approaches (local and global averaging) are sorted in ascending order. The threshold values are selected with the consideration of the maximum distance in the list. The threshold values selected are as given below.

  First, threshold we have is 10% of the maximum ED.

  Second, threshold value is 20% of the maximum ED.

  Third, threshold is 25% of the maximum ED and
Fourth, threshold is 30% and then 40% of the maximum ED.

Retrieval results obtained for 25% of Maximum Euclidean distance has been considered as the standard threshold for both approaches.

**G. Phong shading[19]**
In the paper [19] features are extracted after applying Phong shading on input image. Phong shading is basically used for flattering out the dull surfaces of the image. The features are extracted using color, texture & edge density methods. Following flow graph depicts the steps followed for feature extraction and similarity comparison [19].

![Flow graph](image_url)

**Figure 2.5. The phong shading framework for CBIR.**

Phong shading is a well known method for producing realistic shading but it has not been used by real time systems. The phong shading is basically used to enhance the visual quality of an image to extract features more accurately. The Phong model describes the interaction of light with a surface, in terms of the properties of the surface and the nature of the incident light. The Phong model reflected light in terms of a diffuse and specular component together with an ambient term. The intensity of a point on a surface is taken to be the linear combination of these three components.
Phong shading calculates the average unit normal vector at each of the polygon vertices and then interpolates the vertex normal over the surface of the polygon after calculating the surface normal it applies the illumination model to get the color intensity for each pixel of the polygon. One of the most important features that make possible the recognition of images by humans is a color. Color is a property that depends on the reflection of light to the eye and the processing of that information in the brain. The main method of representing color information of images in CBIR systems is through color histograms. For extracting the color features, the color histogram is used as it is independent of image size & orientation. Here the RGB color histogram is used. For extracting these feature first the histogram for red, green & blue planes are found. Texture is that innate property of all surfaces that describes visual patterns, each having properties of homogeneity. It contains important information about the structural arrangement of the surface, such as; clouds, leaves, bricks, fabric, etc. The texture features describes the relationship of the surface to the surrounding environment. A texture is characterized by a set of values called energy, entropy, contrast, and homogeneity. In short, it is a feature that describes the distinctive physical composition of a surface. The following formulas are used to calculate the features.

\[
\text{Entropy} = \sum_i \sum_j P(i,j) \log P(i,j) \quad (2.1)
\]
\[
\text{Contrast} = \sum_i \sum_j (i - j)^2 P(i,j) \quad (2.2)
\]
\[
\text{Energy} = \sum_i \sum_j P^2(i,j) \quad (2.3)
\]

Along with that Gx and Gy matrices of sobel edge detector is convolved to detect edges present in the image.

**H. Vector quantization [10,11,42,48]**

Vector Quantization [10, 11] is basically the clustering algorithm, where the image is divided into pixel windows. These pixel windows give the texture information of the image. Smaller the window size finer texture details will be represented. Bigger pixel window gives coarse texture details of image [42]. There are various researchers who have used the vector quantization for the feature vector generation in the CBIR. One of the papers [42] suggests the selection of medium size of window such as 2x2 or 3x3. Here the image is divided in 2x2 pixel windows because since the color images are
used, the window is then converted into vector of size 12 to form training vector set. For 3x3 window size the vector dimension will be 27 increasing computational complexity in that ratio. The Kekre’s Median Codebook Generation (KMCG) algorithm is applied on this set. Finally on KMCG code vectors, DCT is applied to get image signature (feature vector) with energy compaction [42]. The feature extraction steps suggested in the paper [42] are explained as given below.

- **Feature Extraction[42]**
The feature vector space has 256x12 numbers of elements. This is obtained using following steps of Kekre’s Median Codebook Generation (KMCG) algorithm.

1. Image is divided into the windows of size 2x2 pixels (each pixel consisting of red, green and blue components).
2. These are put in a row to get 12 values per vector. Collection of these vectors is a training set.
3. The training set is sorted with respect to first column. The centre value of first column is used to divide the training set in two parts.
4. Further each part is then separately sorted with respect to second column to get two centre values.
5. The process of sorting is repeated till 256 centre values are obtained.
6. Using these centre values as code vectors, Codebook of size 256x12 is generated.
7. Codebook is then converted to 1-Dimension of size 3072 and DCT is applied on this to get the feature vector of size 3072.
8. The codebook is stored as the feature vector for the image. Thus the feature vector database is generated.

- **Query Execution**
The codebook of size 256x12 for the query image is extracted using Kekre’s Median Codebook Generation Algorithm. The feature vector of query image can be obtained by applying DCT on this codebook. Then this feature set is compared with other feature sets in feature database using Euclidian distance as similarity measure.

As compared to taking complete DCT of the image, this proposed method takes DCT of codebook only, which saves tremendous number of
computations. As the size of codebook is same for all images, there is no need to resize them for feature vector extraction or query execution, which is necessary in case of other transform based CBIR techniques.

A block of 2x2 color pixels is considered for vector thereby making the training vector dimension equal to 12. The method suggested by the author has been implemented on Intel processor 1.66 GHz, 1GB RAM machine to obtain results. The database consists of total 1100 images of varying sizes from 128x128x3 to 384x256x3. There are 12 different categories with different number of images from 50 to 100 images in each class: Tribal’s, Beaches, Monuments, Dinosaurs, Elephants, Horses, Flowers, Buses, Sunsets, Cartoons, Barbie Dolls, and Puppies. Five query images per class are selected randomly for testing the method.

In other approach [48] of feature vector generation using vector quantization uses Linde-Buzo-Gray (LBG) and Kekre’s Proportionate Error (KPE) algorithms for texture feature extraction [48]. The image is first divided into pixel blocks of size 2X2, each pixel with red, green and blue component. A training vector of dimensions 12 is created using this block. Collection of such training vectors is a training set. To generate the texture feature vector (size of codebook 16X12) of the image, popular LBG and KPE algorithms are applied on the initial training set. Results are compared with the Gray Level Co-occurrence Matrix (GLCM) method. The GLCM is computed in four directions for $\delta=00$, $\delta=450$, $\delta=900$, $\delta=1350$. Based on the GLCM four statistical parameters energy, contrast, entropy and correlation are computed. Finally a feature vector is computed using the means and variances of all the parameters. The steps for extracting texture features of image using GLCM can be given as below.

1) Separate the R, G, B planes of image.
2) Repeat steps 3-6 for each plane.
3) Compute four GLCM matrices (directions for $\delta=00$, $\delta=450$, $\delta=900$, $\delta=1350$)
4) For each GLCM matrix compute the following statistical features
   Energy (Angular second moment), Entropy (ENT), Correlation (COR), Contrast (CON).
The proposed method requires 89.10% less computations compared to the GLCM method. The LBG and KPE based image retrieval techniques give higher precision and recall values than GLCM based method [48].

2.2.2 Frequency domain

A. Basic transforms [87-124,155-157]

The transformed images are used by various researchers to generate the unique feature vector for CBIR. There are lot of work done based on basic transforms so far like use of Kekre’s Transform over Row column Mean and Variance Vectors [46], use of Walsh Transform over color distribution of Rows and Columns of Images[50], Color-Texture Feature extracted from DCT on VQ Code vectors obtained by Kekre’s Fast Codebook Generation for better retrieval [51], using Partial Coefficients of Transformed Image[52], using DCT on Row Mean, Column Mean and Both with Image Fragmentation[85], application of Kekre’s Transform on Each Row of Walsh Transformed VQ Codebook[86], Kekre’s Transform Applied on Each Row of Walsh Transformed VQ Codebook[119], application of DCT on Kekre’s Median Codebook for texture feature extraction[120], use of DST[149], Sectorization of Kekre’s transform [151], Sectorization approach applied on various transforms and wavelets [155-157]. Some of these approaches have been explained in the detail as given below.

I. Hashing[40][43]

Image hashing is one of the techniques used to generate hash value for each images in the database. These hash values generated for images can be used for content based image retrieval, image database indexing, and image authentication, avoiding and mitigating the tampering of digital images. In the information era, the increasing availability of multimedia data in digital form has led to a tremendous growth of tools to manipulate digital multimedia. To ensure trustworthiness, multimedia authentication techniques have emerged to verify content integrity and prevent forgery. Traditionally data integrity issues are addressed by cryptographic hashes or message authentication functions, which are key dependent and sensitive to every bit of the input message. As a result, the message integrity can be validated when every bit of the message is unchanged. While this sensitivity usually meets the need to authenticate text messages, the definition of authenticity for multimedia is not as straightforward. Multimedia data can allow for lossy representations with graceful degradation. The information
carried by media data is mostly retained even when the multimedia has undergone moderate levels of filtering, geometric distortion, or noise corruption. Therefore, bit by bit verification is no longer a suitable way to authenticate multimedia data, and a media authentication tool that validates the content A number of media specific hash functions have been proposed for multimedia authentication. A multimedia hash is a content based digital signature of the media data. To generate a multimedia hash, a secret key is used to extract certain features from the data. These features are further processed to form the hash.
The hash is transmitted along with the media either by appending or embedding it to the primary media data. At the receiver side, the authenticator uses the same key to generate the hash values, which are compared to the ones transmitted along with the data for verifying its authenticity. In addition to content authentication, multimedia hashes are used in content based retrieval from databases. To search for multimedia content, naïve methods such as sample by sample comparisons are computationally inefficient. Moreover, these methods compare the lowest level of content representation and do not offer robustness in such situations as geometric distortions. Robust image hash functions can be used to address this problem.
A hash is computed for every data entry in the database and stored with the original data in the form of a lookup table. To search for a given query in the database, its hash is computed and compared with the hashes in the lookup table. The data entry corresponding to the closest match, in terms of certain hash domain distance that often accounts for content similarity, is then fetched. Since the hash has much smaller size with respect to the original media, matching the hash values is computationally more efficient [40].

- **The use of Hashing algorithm in CBIR [43]**
The proposed method of image hashing advocates to generate the hash value for each image in the database and then the Hash value for the query image is calculated. The hash value of query image is used to calculate the hamming distance with the hash values of images in database. Finally the retrieval of images from the database with minimum hamming distance of same class and cross class is checked to calculate the precision and recall rates of retrieval. The Hashing technique proposed follows following steps to generate the hash value for each image.
Step1: Fast Fourier Transform of image is calculated.
Step2: Extracting the real and imaginary part of the Fourier complex
Numbers of the image calculated in step1. Step3: Cartesian to polar
Coordinate conversion
Step4: Feature vector generated based on the angles of complex plane used
for plotting complex numbers using Euler's formula?
Step5: 4 feature vectors of 16 bits are attached together to generate the 64
bit feature vector.
Step6: Generating the gray code of these 64 bit feature vector to get the
final hash.

- **Feature vector generation using Hashing [43]**
  Every complex number can be represented as a point in the complex plane,
  and can therefore be expressed by specifying either the point's Cartesian
  coordinates (called rectangular or Cartesian form) or the point's polar
  coordinates (called polar form). Complex numbers can be plotted on the
  complex plane. This theory helps us to generate four feature vectors based
  on the complex plane as shown above. The real and imaginary parts of
  complex numbers of images generated by the fast Fourier transform is
  checked for the angle of complex plane lies in. Corresponding intensity
  values of images are taken into first feature vector whose complex numbers
  lies in between angle 0 to 90 on the complex plane. Similarly four feature
  vectors are generated by taking four quadrants of the complex plane.

- **Fractional Energy of Column Mean of Row transformed Image
  with Six Orthogonal Image Transforms [14].**
  This work presents techniques based on feature vectors as fractional
  coefficients of column mean of row transformed images using all basic six
  transforms namely DCT, Walsh, Haar, Slant, DST, and Hartley transforms.
  Here the advantage of energy compaction of low frequency coefficients in
  transform domain is used to minimize the feature vector size per image by
  considering the fractional coefficients of column mean of row transformed
  image. The feature vectors are extracted in six different ways from the
  column mean of row transformed image, with the first being considering all
  the coefficients of column mean of row transformed image and then six
  reduced coefficients sets (as 50%, 25%, 12.5%, 6.25%, 3.125%, 1.56%,
  25% of complete column mean of row transformed image) are considered as
  feature vectors. The all basic transforms are applied on the color
components of images for extraction column mean of row transformed RGB plane respectively. These six reduced coefficients sets for RGB feature vectors are used, instead of using all coefficients of transformed. These techniques are implemented on a database consisting of 1000 images of 10 classes.

**B. Contourlet transform[9]**

This work makes use of the contourlet transform with Unique properties of directionality and anisotropy as a powerful tool for feature extraction of images in the database. Improved results in terms of computational complexity and retrieval efficiency over recent work based on Gabor-Zernike features based CBIR system are observed by the author in this paper. The author has used two distance measures viz., Manhattan distance and Euclidean distance are used as similarity measures in the proposed CBIR system. The work proposed in this paper observes the superiority of Manhattan distance over Euclidean distance in terms of average retrieval rate.

Multiscale and time-frequency localization of an image is offered by wavelets. But, wavelets are not effective in representing the images with smooth contours in different directions. Contourlet Transform (CT) addresses this problem by providing two additional properties viz., directionality and anisotropy.

**Directionality:** The representation should contain basis elements oriented at a variety of directions, much more than the few directions that are offered by wavelets.

**Anisotropy:** To capture smooth contours in images, the representation should contain basis elements using a variety of elongated shapes with different aspect ratios.

Contourlet transform is a multi scale and directional image representation that uses first a wavelet like structure for edge detection, and then a local directional transform for contour segment detection. A double filter bank structure of the contourlet is used for obtaining sparse expansions for typical images having smooth contours. In the double filter bank structure, Laplacian Pyramid (LP) is used to capture the point discontinuities, and then followed by a Directional Filter Bank (DFB), which is used to link these point
discontinuities into linear structures. The contourlets have elongated supports at various scales, directions and aspect ratios. This allows contourlets to efficiently approximate a smooth contour at multiple resolutions. In the frequency domain, the contourlet transform provides a multiscale and directional decomposition. Contourlet transform is simple and flexible, but it introduces redundancy (up to 33%) due to the LP stage. These properties of CT, i.e., directionality and anisotropy made it a powerful tool for content based image retrieval.

- **Proposed Algorithm[9]**

The basic steps involved in the proposed CBIR system includes database processing and resizing, creation and normalization of feature database, comparison and image retrieval. Steps of the proposed algorithm are as follows.

1. Decompose each image in the Contourlet domain
2. Compute the standard deviation (SD) of the CT decomposed image on each directional sub-band. Standard deviation is given as

\[
\sigma_k = \frac{1}{\sqrt{M \times N}} \sum_{i=1}^{M} \sum_{j=1}^{N} (W_k(i,j) - \mu_k)^2
\]  \hspace{1cm} (2.4)

Where \( W_k \) = Co-efficient of \( k^{th} \) CT decomposed sub-band  
\( \mu_k \) = Mean value of \( k^{th} \) sub-band  
\( M \times N \) = Size of the CT decomposed sub-band

Resulting standard deviation (SD) vector is given as

\[
f = [\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n]
\]  \hspace{1cm} (2.5)

3. Normalise the SD vector to the range [0 1] for every image in the database.

\[
f_{CT} = \frac{f - \mu_f}{\sigma_f}
\]  \hspace{1cm} (2.6)
Where \( \mu_f, \sigma_f \) are the mean and the standard deviation of \( f \). This normalized SD vector \( f_{CT} \) is used to create the feature database.

4. Apply query image and calculate the feature vector as given in the steps 2 to 3.

5. Calculate the similarity measure using Manhattan distance as given below:

\[
D = \sum_{i=1}^{M} |f_q - f_i|
\]  

(2.7)

Where \( D \) is the Manhattan distance between the feature vector of query image and every image in the database. \( f_q, f_i \) are the normalized SD feature vectors of the query image and database image respectively.

6. Retrieve all relevant images to query image based on minimum Manhattan distance.

This query is then processed to compute the feature vector as in equation (2.5) and (2.6). The distance is computed. The distances are then sorted in increasing order and the closest sets of images are then retrieved. The top ‘N’ retrieved images are used for computing the performance of the proposed method. The retrieval efficiency is measured by counting the number of matches.

Spatial and spectral features of the images can be explored for image retrieval in CBIR systems. Due to the local nature, it is difficult to detect edge and texture orientations using spatial methods [24]. Spectral methods based on multiscale directional transforms viz., wavelets have fixed number of directions. They are also inefficient to capture edges and smooth contours in natural images. The Contourlet Transform (CT) is a directional transform capable of capturing contours and fine details in images. The contourlet expansion is composed of basis functions oriented at variety of directions in multiple scales with flexible aspect ratios. With this rich set of basis functions, the contourlets can effectively capture smooth contours (Edge and Texture orientations) that are the dominant features in images in the database.
The performance of the CBIR system is dependent on the feature vector that represents the image in the database [9]. Important characteristics of contourlet transform viz., directionality and anisotropy are explored in this work. Normalized standard deviation calculated in each sub band of the CT decomposed image is used as feature in the feature vector representing the image. The author suggests the performance of retrieval system can be further improved by considering the energy of each sub band in the CT decomposed image, in addition to standard deviation. However, the computation time increases further due to the increase in size of feature vector. The feature vector of the CT decomposed image varies if the database consists of images with different rotations. Hence, It is suggested by the author that a CBIR system with rotational invariance is to be explored.

C. Energy compaction [17,41,206]

The work presented here makes the use of energy compaction to its fullest to generate the feature vector for CBIR. Kekre Transform [17, 46, 86, 123, 132, 151, 206] applied on an image transfers the high frequency components at the higher end and low frequency components towards lower side. This can be used as an advantage in image retrieval by eliminating the coefficients which do not contribute significantly. The energy compaction method therefore aids in reducing the feature vector size, which gives faster retrieval. This leads to the another approach of the feature vector generation as explained below.

• **Average Energy Plot[41, 206]**

Average Energy plot depicts the average energy compaction done by Kekre Transform on all the database images. The average energy plot can be obtained for each of the above specified feature vector techniques, i.e., row mean and column mean. Firstly, all the feature vectors are arranged into a two dimensional array. Now, the average value feature vector is computed by adding corresponding values and dividing it by number of feature vectors. The number of feature vector used depends on the splitting technique used as explained in section 3. The database considered consists of 1000 generic images. Hence if single part technique is used, 1000 feature vectors, either row mean or column as per selection, are obtained. For 4 parts technique
four features, one for each part, are obtained. Therefore, the two dimensional feature vector array will consist of 4000 vectors. Once the average vector is obtained, Kekre Transform is applied on it. By application of Kekre Transform, the high frequency components are obtained at the higher side of the vector. Further, its terms are squared to obtain positive values of energy. In the following subsections calculation of average value feature vector for row mean and column mean is discussed.

**i. Row mean feature vector**
Here, row mean is considered as the feature vector. Hence, the row means of all the parts of all the images in the database are arranged in a two dimensional array and average row mean is calculated. The size of row mean depends on the image splitting technique [41] used. For an image if size of row mean is \( N \) when 1 part method is used, the size of row mean is \( N/2 \) when 4 parts method is used and four such row mean vectors can be obtained per image. Now, Kekre Transform is applied to obtain transformed row mean vector.

**ii. Column mean feature vector [46]**
In this method, column mean is considered as the feature vector. Hence, the column means of all the parts of all the images in the database are arranged in a two dimensional array and average column mean is calculated. The size and number of column means depends on the splitting technique as in case of row means. Now, Kekre Transform is applied to obtain transformed column mean vector. This vector depicts the values of the average energy plot for all images in the database considered together. Further, after calculation of the transformed average feature vector, the values are added to get a cumulative vector.

**D. Sectorization of basic transforms[133-144]**
The new and innovative idea of conceptualization of planes from transformed images has been suggested. These planes are further sectored in various numbers of sectors like 4, 8, 12 and 16 sectors. The sector mean and sector density values are considered to generate feature vector. These feature vectors are experimented with augmentation of additional component consisting of average DC values of the transformed image. This method of sectorization has been applied on all seven basic transforms namely DFT,
DCT, DST, KT, Hartley, Walsh and Slant. These methods have shown quite good retrieval performances for the image database of size 1055 applied on. The retrieval results of the approach of sectorization has been measured by means of three newly developed approaches of Precision-recall cross over point (PRCP), Length of initial string of relevant images retrieved in the output (LIRS) and Length of string of images to retrieve all images of the class (LSRR). These approaches have never been used by anyone else before. The work suggested in these papers makes use of the sum of absolute difference as the similarity measure along with the well known Euclidian distance. The experiments performed show the better performances of sum of absolute difference over the Euclidian distance.

E. Wavelets[7,104,159,195,199]

The wavelet analysis procedure is to adopt a wavelet prototype function, called an analyzing wave or other wave. Other wavelets are produced by translation and contraction of the mother wave. By contraction and translation infinite set of functions can be generated. This set of functions must be orthogonal and this condition qualifies a transform to be a wavelet transform. Thus a transform is qualified to be a wavelet transform if and only if it satisfies the condition of orthogonality. Thus there are only few functions which satisfy these conditions. Principal advantage of wavelet is that they provide time-frequency localization. Temporal analysis is performed with a contracted, high-frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low-frequency version of the same wavelet. The exact shape of the mother wave strongly affects the accuracy and compression properties of the approximation.

Because the original signal or function can be represented in terms of a wavelet expansion (using coefficients in a linear combination of the wavelet functions), data operations can be performed using just the corresponding wavelet coefficients. And if you further choose the best wavelets adapted to your data, or truncate the coefficients below a threshold, your data is sparsely represented. This sparse coding makes wavelets an excellent tool in the field of data compression [7,104].

The Haar wavelet is also the simplest possible wavelet. The technical disadvantage of the Haar wavelet is that it is not continuous, and therefore not differentiable. This property can, however, be an advantage for the
analysis of signals with sudden transitions, such as monitoring of tool failure in machines.

The Haar Wavelet’s mother wavelet function \( \varphi(t) \) can be described as [195]:

\[
\psi(t) = \begin{cases} 
1, & 0 \leq t < \frac{1}{2} \\
-1, & \frac{1}{2} \leq t < 1 \\
0, & Otherwise
\end{cases}
\]  

Based on the concept of Haar wavelet the generation of Haar wavelet pyramid can be used for feature extraction in CBIR as given in the paper [159]. The concept of the same is discussed below.

**J. HAAR WAVELET PYRAMID [159]**

The Haar Wavelets [159,199] of particular image at different levels, when considered together gives Haar Wavelet Pyramid. Here the author has considered the first level of Haar Wavelet pyramid for that the Haar transform has been applied on the given image. This process gets the approximation, horizontal, vertical and diagonal components of the image. The approximation components of first level Haar Wavelet is considered to be transformed with Haar T. to get second level Haar Wavelet. The Haar Wavelet pyramid of a sample images are shown in Figure 2.6 given below. Where the baby image is decomposed into three levels of Haar Wavelet pyramid as Haar Wavelet level-1, Haar Wavelet level-2 and Haar Wavelets level-3. The approach is followed further to proceed with the various levels of the Haar wavelet pyramid and the feature vector is generated for the use in the CBIR. The feature extraction approach followed in this work has been explained in the following section in detail.
Feature Extraction [159]
Here the approximate components of Haar Wavelet level-1, Haar Wavelet level-2, ..., Haar Wavelet level-7 are obtained for every image in the database and haar transforms of respective sizes are applied on them, the results are stored as feature vectors for respective image. At level-1 Haar Wavelet the feature vector size is N/2xN/2. At level-2 Haar Wavelet the feature vector size is N/4xN/4 and so on. Thus the feature vectors for up to level-7 Haar Wavelets are extracted and the feature vector database is generated. The Gray-Haar Wavelets are extracted from gray images (average of red, green and blue components is taken as gray). Then the Haar Wavelets of Red, Green and Blue planes of images are extracted and considered as Color-Haar Wavelets of respective images for various levels.

Query Execution [159]
Here the feature set of Haar Wavelet level-p is extracted as a feature set for query image using proposed technique of Haar Wavelet generation. Then these are matched with Haar Wavelet level-p feature vector database using Euclidian distance as similarity measure. As compared to applying complete Haar transform on the image, the author of the paper comments that this proposed method takes fewer computations to extract the feature set and gives better precision and recall values. For image of size NxN complete Haar needs 2N2log2(N) additions and for Haar Wavelet of level-p the
number of additions needed are $2(N/2p)^2 \log(N/2p)$ as the size of feature vector would be $(N/2p) \times (N/2p)$. This gives tremendous reduction in query execution time using higher Haar Wavelet level.

- **Implementation [159]**
  This method [159] is implemented in Matlab 7.0 on Intel Core 2 Duo Processor T8100, 2.1 GHz, 2 GB RAM machine to obtain results. To check the performance of proposed technique precision and recall has been used. To test the proposed CBIR techniques using Haar Wavelets, the databases of 1000 variable size images spread across 11 categories of human being, animals, natural scenery and manmade things are used. Here all images were resized to 256x256x3 before using Haar Wavelets for feature extraction. Five queries were selected from each category of images, so in all 55 queries for every Haar Wavelet level and complete Haar T. are fired on the database to test the proposed CBIR techniques. To measure the retrieval effectiveness, author has used the precision and recall as statistical comparison parameters for our proposed technique of CBIR.

**II. Sectorization of Wavelets[147,148,152,153]**

The method of using wavelet transforms to get the image feature extracted has already been used by various researchers as discussed earlier. The plane sectorization concept followed by us for the basic transforms is extended for the wavelets and discussed in chapter 4 and chapter 6 in detail.

**Generation of Wavelet transform matrix for any basic transforms**

Generation of Wavelet transform matrix for any basic transforms like DCT, DST, DFT, Hartley etc. of size $N \times N$ from its basic matrix of size $N \times N$ is possible as given in [104]. However our images are of size 128x128 and hence we require Transform matrix of size 128x128 where 128 is not a square. To simplify this situation, we propose an algorithm to generate wavelet transform from the basic transform matrix of 8x8 and 16x16. The 128x128 Wavelet transform matrix can be generated from 16x16 and 8x8 orthogonal matrices of the particular basic transform. First 16 rows of Wavelet transform matrix are generated by repeating every column of the basic transform matrix of dimension 16x16, 8 times. To generate next 17 to 32 rows, second row of basic transform (8X8) is translated using groups of 8 columns with horizontal and downward shifts. To generate next 33 to 48 rows, third row of basic transform (8X8) matrix is used in the same manner.
Like wise to generate last 113 to 128 rows, 8th row of basic transform (8x8) matrix is used. Note that by repeating every column of the basic transform matrix 8 times we get global components. Other components are generated by using rows of basic transform matrix of size 8x8 giving local components of the basic transform Wavelet. This concept of Wavelet transform matrix generation has been applied for various transforms like DCT, DST, Hartley, DFT, Walsh, Kekre Transform and slant transform to extract the feature vectors for the purpose of retrieval check in Content Based Image Retrieval (CBIR). The concept of sectorization has been applied on the Wavelet transformed images in the Database. These newly introduced wavelets have been experimented for all three types of transforms that is full, column wise and row wise.

2.3 Summary

As per our extensive literature review we have observed that all CBIR researchers have considered only the magnitude values of the transformed images for feature vector generation. The Discrete Fourier Transformed (DFT) image consists of complex co-efficient values. It has both the magnitude and Phase part in it. None of the researchers have been found exploring the possibilities of the phase angle of it for the purpose of feature vector generation. In this thesis phase angle has been used to partition the complex transform plane into number of radial sectors. Using sectorization exhaustive study has been performed for many transforms and their wavelets in subsequent chapters.