On Feature Extraction And Its Application To Content - Based Image Retrieval

Thesis Submitted to

Jadavpur University

by

Minakshi Banerjee (Bagchi)

Center for Soft Computing Research, Indian Statistical Institute

203 B. T. Road, Kolkata, India

On Feature Extraction And Its Application To Content-Based Image Retrieval

A thesis submitted to the Jadavpur University in partial fulfillment of the requirements for the award of degree of

Doctor of Philosophy

2008

by

Minakshi Banerjee (Bagchi)

Center for Soft Computing Research, Indian Statistical Institute

203 B. T. Road, Kolkata - 700108

India

under the supervision of

Prof. M. K. Kundu

Machine Intelligence Unit, Indian Statistical Institute

Kolkata - 700108

and

Prof. P. K. Das

Dept. of Computer Science and Engineering, Jadavpur University

Kolkata - 700032

To my parents

Acknowledgments

I express my sincere thanks to my supervisor Prof. M. K. Kundu who introduced me to the world of images, visions and patterns. With a deep sense of gratitude I remain obliged for his unconditional support and guidance during the course of my research work. I greatly acknowledge his valuable conceptual suggestions during the research work. Prof. Kundu gave me the freedom to explore a variety of topics. Whenever I struggled, his generous support always came at just the right time. It is my privilege to work under the supervision of a person like him of profound knowledge and humble personality.

My sincere thanks goes to my joint supervisor Prof. P. K. Das of Jadavpur University for his guidance, valuable suggestions, moral support and encouragement time and again during the research period. I owe a lot to him for providing me constant support throughout the last few years.

I would also like to extend my sincere thanks to the distinguished Professor S. K. Pal of ISI, for his valuable conceptual suggestions, encouragement and moral support. I remain obliged for all the supports he has provided me, during the research period.

I want to thank the professors of the Machine Intelligence Unit : Professor C. A. Murthy, Prof. Sushmita Mitra, Dr. B. Uma Sankar for their constructive suggestions and comments. Also my thanks goes to Dr. Sambhunath Biswas, Prof. Ashish Ghosh and Dr. D. P. Mondal. I thank Dr. Sanghamitra Bandopadhyay, Dr. Rajat De, Dr. Amita Pal and Dr. Mausumi Acharyya for valuable suggestions and encouragement. I extend my hearty thanks to Dr. Pabitra Mitra for supporting with conceptual suggestions and expertise, during the time he was within us. I am also thankful to Dr. Pradipta Maji for

helping me in my research work by providing fruitful academic suggestions.

I would never forget the company I had from my fellow research scholars B. L. Narayan, Sarif Kumar Naik, Sriparna Saha, Ranajit Das, Mouli Das, Manish Chaudhury and friends. In particular, I am thankful to Dr. Swati Choudhury with whom I had several fruitful discussions. She provided valuable advises and helped me during the research work. I enjoyed my work and and had fun with the company of my dear friends Swati, Maya di during the period. I am also thankful to Priyank Bagrecha for helping in developing the programs. Special note of thanks goes to Mr. Joydev Gupta, Mr. Indranil Dutta, Mrs. Maya De, Mrs. Niyati Das, Mr. Sanjay Das, Mr. Samir Das, Mr. Biswanath Porel, Mr. Surja Bhattacharya for providing me all kinds of facility whenever needed.

I would like to thank http://bergman.stanford.edu/cgi-bin/$www\_wavesearch$ for creating SIMPLIcity database and COREl dataset and making it available for research purpose.

I am also grateful to the Department of Science and Technology, for providing fellowship under Women Scientists Scheme (grant No. SR/WOS-A/ET-111/2003) to continue the research work. My vote of thanks goes to the Machine Intelligence Unit, Center for Soft Computing Research for providing all facilities to perform the research work.

I thank my parents and all elders in my family, who taught me the values of hard work and patience. I am also grateful to my in-laws and my husband who provided continuous family support and valuable personal advices during my research period. I would like to share this moment of happiness with my sons Debapratim and Debangshu and all my family members and well wishers.

Contents

List of Figures

List of Tables

Settings for fancy header

Chapter 1

Introduction and Scope of the Thesis

Accessing of contents from digital libraries is becoming very popular in different application domains with the advent of the World-Wide Web. This necessitates the development of cost effective techniques that support effective search through the content of large digital archives. Images are one of the most important components of the widely accessed multimedia data. Images are widely used in every sphere of life, such as item catalogs, education, biomedicine, commerce, crime prevention, etc. It is often required to find desired images from a very large repository or to measure closeness between images available in some collections. As a result, development of suitable image retrieval techniques has emerged as a growing field of research in recent years [187], [119].

In the recent past, Content-Based Image Retrieval (CBIR) techniques have become popular [189], [43], [127], [108], [70], [91], [103], [36] for retrieving relevant images from an image database by measuring similarity between the automatically derived features (color, texture, shape, etc.) of the query image and that of the images stored in the database [187]. Image retrieval system is broadly divided into two main types namely, annotation or text-based image retrieval and Content-Based Image Retrieval(CBIR) [192]. Traditional image retrieval systems used keywords as labels to characterize images. However characterization of images with text labels is difficult in the case of large databases and for complex images. The same image may be perceived differently by different persons at

different times. Image data is very subjective in nature and there is no agreed vocabulary for manual annotation of these data. As a result, the traditional text-based image retrieval systems suffer from the following drawbacks : (a) Enormous man-hour requirement (b) error arising due to variability of individual perception. These difficulties call for a new and alternative approach known as Content-Based Image Retrieval.

CBIR is aimed at efficient retrieval of relevant images from large image databases, based on automatically derived image features. Features like color, texture, shape, sketch, spatial relationship between features, etc. are taken as queries for measuring visual similarities between the query image and those stored in the database.

Although, a fully automatic CBIR system is desirable but development of such a system is a very difficult task. Human beings capture perceptual and semantic meanings from an image easily but it is difficult to achieve automated interpretation close to human level understanding. Images with high feature level similarities (i.e., similarities between color, texture, shape, etc.) may be different from the query in terms of semantics. The discrepancy between low level visual contents like (color, texture, shape, etc.) and high level semantic concepts like (sunrise, picturization of some special events, etc.) is called semantic gap. Minimization of semantic gap is an open challenging problem in CBIR. To bridge this gap, user feedbacks may be used in an interactive manner which is popularly known as ”relevance feedback” mechanism.

Using a fully automatic CBIR system one cannot achieve satisfactory results. Therefore it cannot replace wholly the traditional text based retrieval system in the near future. However, it is possible that a text based retrieval system can be used in combination with a CBIR system, to enhance the retrieval accuracy without totally depending on human annotation. At the same time it takes care to minimize the burden on a user.

The present thesis deals with the development of several algorithms for feature extraction of digital images and its application to perform CBIR. Use of Fuzzy logic, Wavelet and Support Vector Machines have been demonstrated from feature extraction perspective. An image is characterized in terms of the extracted features. A fuzzy relevance feedback

framework has been proposed to evaluate the extracted features from user's relevance rating in order to improve the results.

Finally, the developed algorithms have successfully been applied for retrieving relevant images similar to the queries, from real life image database. Ultimately, the described works would be presented through a chapter. The said chapter is organized as follows. Section 1.1 gives a formal description of the main aspects of images used in image retrieval and a brief description of a CBIR system. Section 1.2 describes the main research related issues in CBIR. Section 1.3 discusses about the motivation and the focus of the problems. Section 1.3.1 describes the problem definitions. Section 1.3.2 provides a short overview of different CBIR approaches. Section 1.3.3 narrates the scope of the thesis.

1.1 Main characteristics of the image used in retrieval

The automatic interpretation of semantically meaningful information from low-level visual features like color, texture shape, etc. is the focus of interest for most research on image retrieval. The user's requirements can be specified by querying from a database in the form of attributes like color, texture, shape, spatial relationships, a sketch, an example image or in terms of keywords.

Visual contents of an image

The distinguishing aspects of visual information retrieval is that it is based on the availability of a representation of visual content. A digital image consists of picture elements with finite size(pixel). These pixels carry information about the brightness of a particular location in an image. Such a digital image is represented by a two-dimensional matrix whose elements are mare integer numbers corresponding to the quantization levels of the brightness scale. Similarity define the relation between image pixels or image features regardless of its perceptual causes. When there is much data in the matrix, the processing time required to evaluate similarity between images at pixel levels is very high. Better representation of an image may be obtained if global features are derived from the orig-

inal image matrix first. Such representation (at macro level) is more concise and occupies less memory. Content descriptors may regard perceptual features like color, texture, shape, structure, spatial relationship, etc. For the purposes of content-based retrieval, an integrated view on color, texture, and local geometry is very useful, because only an integrated view on such properties can provide the means to distinguish among hundreds of thousands different images.

The visual contents of an image can be defined from (a) Low level properties like, color, texture, shape and spatial relations. (b) Semantic properties : that correspond to objects, scenes, impression, emotions and meanings associated with the combinations of low level properties.

The important visual features may be described as follows :

Color features : Color makes the image take values in a 3-D color vector space. Commonly used color space include RGB, Munsell, CIEL*a*b*, CIEL*u*v*, HSV (or HSL, HSB) and opponent color space. There is no agreement, on which color space is the best. RGB color representation a good choice since that representation was designed to match the input channel of the eye. RGB-representations are in wide-spread use. They describe the image in its literal color properties. Others approaches use the Munsell or the Lab-spaces because of their relative perceptual uniformity. The Lab representation is designed so that the Euclidean distance between two colors representations models the human perception of color differences. The HSV-representation is often selected for its invariant properties. The hue is invariant under the orientation of the object with respect to the illumination and camera direction and widely used in different retrieval applications.

Shape features : The shape properties capture conspicuous geometric details in the image. Use of Shape features are best way to enhance object-specific information contained in images. Edges and salient points are often used to characterize shape properties.

Salient feature calculations lead to sets of points with known location and feature values capturing their salience. Corners or commonly the high curvature points are important salient points used to capture structural information within an image. With the use of edges and salient points the information of the image is condensed into just a limited number of feature values. Extraction of edges and salient points leads to a grouping of the data based on psycho visual perception and may be used to avoid the brittleness of strong segmentation into regions.

Edges : Edges are used to locate the changes in intensity functions. Edges are pixels where the intensity function changes abruptly. An edge is a property attached to an individual pixel. It is calculated from the image function behavior in neighborhood of that pixel. It is a vector variable with two components magnitude and direction. The edge magnitude is the magnitude of the gradient and the edge direction is the direction perpendicular to the maximum gradient. There are several popular edge detectors which use spatial and frequency filtering techniques [75], [86].

Corners : Corners are very good interest points. Corner points lie on sharp edges and are generally defined from curvature.

Texture features : Texture contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. Effective representation of textures can be made from the statistical and structural properties of the brightness pattern.

Spatial relations : Within an image, the information between the semantically important parts of image-objects may be derived in terms of relations like, inside, top, left to, right to, etc. Descriptions by means of relational structures are appropriate for higher levels of image understanding.

1.1.1 Query specification

The image semantics i.e, the meaning of an image has different levels depending upon the degree of abstraction used to describe it. As a result, the different query categories which an efficient CBIR is expected to handle may be broadly classified into 3 levels, based on the level of complexity.

Level 1: This uses primitive features like color, texture, shape or spatial relationships of image elements that could be extracted automatically from the images without the help of external knowledge base.

Level 2: This level uses derived features involving some degree of logical inferencing about the identity of objects depicted in the image. For example, retrieval of objects for a given type (find a double decker bus in the image) or retrieval of individual objects or persons (find the picture of the Eiffel tower).

Level 3 (high level semantics): This involves a significant amount of high level reasoning about the meaning or the purpose of the objects or scenes depicted.

In the present state -of -the -art, the research in CBIR has mostly dealt with level 1, i.e, with low level properties. However, efforts are also being made to achieve success in retrieval, based on level 2 and 3 i.e, the semantic aspects [119].

1.1.2 Main components of a CBIR system

The main modules of a CBIR system are as follows : (1) The user interface. (2) The feature extraction subsystem. (3) The indexing subsystem. (4) The query processing subsystem. (5) The feature matching subsystem (6) Relevance feedback module [30].

1. User interface: The user interface facilitates interaction between the user and the system, specially when presenting a query to the database and viewing the retrieval results.

2. Feature extraction subsystem: A subsystem for extraction of visual properties, for proper representation of images. Image processing and pattern analysis are mainly used

in order to detect visual properties and compute their measures. As mentioned earlier that, features may be grouped into two broad categories. (a) Low level features (b) High level features. Low level features are in general the features computed from the pixel values. High level features mostly involve semantics of the entire image. In some cases, semantics can be extracted automatically from images, based on a combination of a low level features through a suitable set of rules. Semantic information may be partly supplied as textual annotation in general for all those concepts that cannot be extracted automatically.
3. Indexing subsystem : The indexing subsystem provides an efficient way to access the features of each image in the database [124]. If the database is very large, then suitable index structures are required for careful management of the feature vectors. Hashing tables and signature files are the commonly employed indexing methods used to index keywords or alphanumeric strings.
However, the visual properties like color, texture, shape, etc, are modeled as points in a multidimensional feature space. Among the several multidimensional point indexing methods, the dynamic data structures, R-tree, K-d tree are popularly implemented in CBIR for feature vector indexing from a large database.
4. Query processing subsystems : Performs necessary operations on the features extracted from the query image such as, (a portion of the image may be submitted as a query). To initiate a query, the user selects which features or parameters are important for a particular case, where the user's query may be in terms of an image, a sketch or keywords.
5. Feature matching subsystem : This module is used to perform similarity matching between each image in the database and the query image. The similarity based image retrieval is the task of re-ordering database images, according to some specified distance measure like, Euclidean distance, Mahalanobis distance, etc.
6. Suitable relevance feedback : Automated image analysis is important because it provides information about the image without costly human interaction. In spite of its attractive advantages, a fully automated system cannot generate satisfactory results. This is because images with high feature similarities to the queries may be different in terms

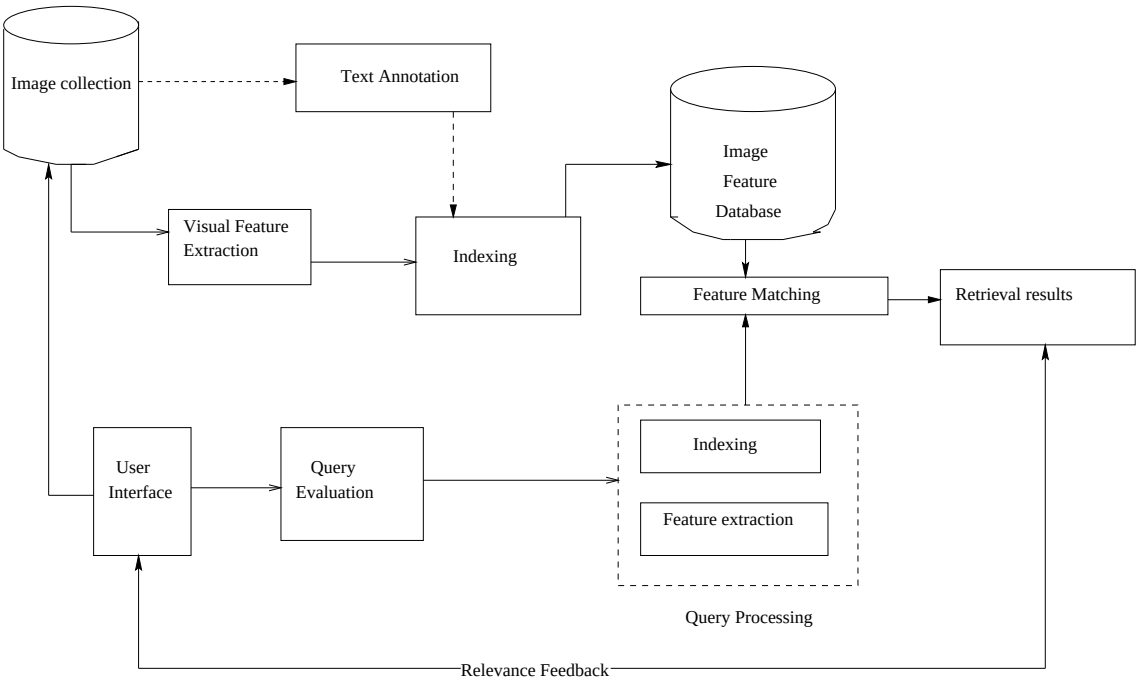


Figure 1.1: Building Block of a typical CBIR system.

of semantics. To solve the problem of mapping, a high-level concept in terms of low level features is known as minimization of semantic gaps. To bridge this gap, human and computers should interact to refine high level queries, with representations based on low level features. The iterative process of refinement of queries and results guided by user's feedback is known as "relevance feedback" [178], [224], [80], [45], [214], [192], [117].

1.2 Main research related issues in CBIR

Research in CBIR is experiencing high potentiality in the recent past. It has opened scope in multiple domains, some of which are mentioned below.

1. Suitable feature extraction model : Research in image retrieval can be exploited to define new ways for automatic representation of images, so that high level concepts can be mapped together with the extracted low level visual features [107], [144], [43], [70], [103], [36], [50], [147]. However, retrieval systems based on conceptual representation

of semantic aspects are very few and research in this area is still in its infancy [114].

2. Suitable relevance feedback : Suitable Relevance feedback mechanism is important for an effective image retrieval system [178], [224], [80], [45], [214], [192], [117], [73]. Several techniques have been proposed to enhance the retrieval accuracy of the system and has become one of the important issues in CBIR research.

3. Suitable data model and knowledge structures: Suitable data model [30], [180] and knowledge structures are required to organize the extracted visual information [60]. This has become an important research issue.

4. Similarity models : Effective similarity models are used to measure similarities between the extracted features [179], [198]. Currently used similarity models, are mostly based on the evaluation of proper distance function in the metric feature space. The popularly used metrics are Euclidean, Mahalanobis distance, etc. These models are not satisfactory enough in many cases. Psychological similarity models that fit human similarity judgment more closely must be investigated.

5. Effective indexing : Operations allowing efficient indexing of high dimensional features, and development of dynamic structures that allow changes in the database without the need for rebuilding the whole index for large database, are important research issues [162], [105], [103].

1.3 Motivation and objectives

The accuracy of an image retrieval system strongly depends on the proper representation of visual contents. Under this perspective, the feature extraction is the most important component of the retrieval architecture. Various image classes that compose the database should be clearly separated in feature space. Once such separation is achieved, the remaining components become fairly easy to design. Main objective of feature extraction is data reduction which means representing visual patterns by a minimal set of features or properties. Although minimum number of features are desirable in developing a CBIR

system but in such cases the accuracy of retrieval may not be satisfactory. In general, a trade off is often being made between the computational cost involved in considering more features and the retrieval accuracy. In this line, the thesis is organized to show the retrieval performance by adding different type features, which are discussed in the subsequent chapters.

Efficient feature extraction should exhibit the following important properties :

Invariance and **perceptual relevance**

Invariant transformations are those which are robust to changes in imaging conditions, like change in illumination, linear transformations like rotation, translation and scaling. Perceptually relevant transformations mimic in some way the properties of the human visual system. This gives rise to challenging issues like handling incomplete query specification, incomplete image description, variability of sensing conditions (sensory gap) and object states, uncertainties arising due to perturbation of data, individuality in perception mechanism, etc.

1.3.1 Problem definition

This thesis makes contributions that are related to CBIR. Emphasis has been given mainly on feature extraction for efficient CBIR, looking into some of the practical difficulties in CBIR. There may be multiple interpretation of the same visual data by different users. As a result, relevancy of the retrieval are strongly related to prediction accuracy. Therefore the system should support multiple features, multiple representation of features according to its relative importance, learning mechanism guided by user's feedback, handling of impreciseness arising due to incomplete or perturbed data, etc. As an attempt to solve some of these problems, use of Soft decision making through application of fuzzy logic in the feature extraction and feature evaluation module for relevance feedback has been made. Applications of Wavelets and Support Vector Machines have also been explored for efficient characterization of images.

The basic outlines of different contributions described in this thesis are as follows :

(a) Design of a Content-Based Image Retrieval system based on fuzzy edge based features. Multilevel edge maps consisting of weak, medium and strong edge points (characterized through gradient membership values) are generated from a gray level image. Using the fuzzy edge maps, a fuzzy compactness feature vector is computed, which is subsequently used for measuring the similarity between the query and the database images. The algorithm is also used to retrieve logo images.

(b) A fuzzy-set theoretic approach for detection of gray level corners is proposed. To test the robustness of the proposed algorithm, the results are tested on different images which have undergone changes due to blurring, varying illumination, etc.

(c) A classification based corner detector is proposed, using Support Vector Machines on the extracted fuzzy edge data.

(d) Design of an image retrieval system by extracting significant color features from visually significant points (i.e., in the locality of the true corner points). A fuzzy entropy based mechanism to evaluate the features from the retrieved results guided by user's relevance feedback is also proposed.

(d) Perform segmentation to extract significant colored-textured regions using wavelet packet frames and compute fuzzy spatial relations between the segmented regions to implement a color image retrieval scheme.

1.3.2 A short review on different CBIR approaches

Color, texture, shape, spatial relations, etc. are the most widely used features in CBIR applications. A short description about application of various features is provided in the following subsection. We also present the State -of- the -art in CBIR by giving a brief discussion on some of the popular commercially available CBIR systems.

Retrieval by color based representations :

Color feature is one of the most widely used visual features employed in Image retrieval.

Much research has been devoted towards generation of robust and efficient color features, so that the features be able to present the following properties like, perceptual similarity, low complexity, low dimensions without affecting the retrieval accuracy, changes in imaging conditions of the database images, etc.

Each pixel of the color image can be represented as a point in a 3D color space. Commonly used color space for image retrieval include RGB, Munsell, CIEL*a*b*, CIEL*u*v*, HSV (or HSL, HSB) and opponent color space. There is no agreement, on which color space is the best. The effective and efficient computation of color indices generally require sufficient reduction in the number of colors to represent the color contents of an image. This may be achieved through color space quantization.

Image retrieval based purely on color distributions may not achieve satisfactory results. Integration of color based features with other features may be used to produce effective results.

Some of the popular color quantization techniques are mentioned as follows : Smith and Chang [188] have partitioned the HSV color space into 166 bins, placing more importance on hue (18 levels) than on value and saturation (taking three levels each). Median filtering has been performed on each of the HSV color space as a pre-processing step, to eliminate noise and emphasize on prominent color regions. Gagliardi and Schettni [68] have proposed the use of multiple description of colors. In their quantization method, the CIELAB color space is divided into 256 subspaces (categories). The colors which remain perceptually the same, are labeled with its own linguistic tag, according to ISCC-NBS color naming system. Images quantized by this method are further clustered in a set of different equivalent classes representing different basic colors (black, gray, white, red, etc.) to produce a coarse unsupervised segmentation. The significant color components which are extracted after effective quantization of the color spaces can be represented in many different ways to represent the image content.

Among the different color content representation of images, Color Histogram is the most traditional way of describing low level color properties of images and has been widely

used in image retrieval applications, [139], [30] [164], [195], [92].

Color histogram is obtained by discretizing image colors and count how many pixels belong to each color. A color histogram H is a vector [h_1, ..., h_n] in which each bin h_j contains the number of pixels having the color j in the image. Color histograms are fairly robust to translation and rotation about the view axis and changes slowly with the scale, occlusion and viewing angles. Similarity between image histograms can be measured in terms of the sum of squared differences (L_2 metric), or the sum of absolute values of differences (L_1 metric). Swain and Ballard have proposed Histogram Intersection measure (L1 metric) as a similarity measure for the color histogram [196]. To take into account the similarities between similar but non identical colors, Niblack et al. [62] introduced a weighted Euclidean metric in comparing the histograms. In order to reduce noise sensitivity, Sticker and Orengo [193] proposed the use of cumulated Color Histograms and obtained better results over conventional color Histograms. In [193], boundary histograms has been used to encode the lengths of the boundaries between different discrete colors, in order to take into account geometric information in color image indexing. Clustering methods can be used to determine the k- best colors in a given space, where each of the best color will be taken as a histogram bin. Selective bins that capture the largest possible pixels numbers is proposed in [74]. Histogram comparison may not generate satisfactory results, in the case of image database containing large number of images. To solve this problem, the joint histogram technique is introduced in [160].

As color histogram lacks spatial information of pixels, different images can have similar color distributions. This problem has been handled by many researchers, [92], [90]. Mandar et al., have made the use of color correlogram. 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. A different way of incorporating spatial information into the color histogram, was color coherence vectors (CCV), which partitions each histogram

bucket based on spatial coherence. Each histogram bin is partitioned into two types, i.e., coherent, [160] if it belongs to a large uniformly-colored region, or incoherent, if it does not. Such an approach gave better results in the case of image database containing large number of images. Other similar approach includes application of motif cooccurrence matrix (MCM) [93]. Conceptually, the MCM is quite similar to the color cooccurrence matrix (CCM), however, the retrieval using the MCM is better than the CCM since it captures the third order image statistics in the local neighborhood.

A new content-based image retrieval method using the color and pattern histogram that is adaptive to the block classification characteristics has been proposed in [195]. A new and effective image indexing technique that employs local uni-color and bicolor distributions and local directional distribution of intensity gradient is proposed in [164]. Histograms with adaptive bins have been proposed in [113], which performed better compared to existing methods for histogram retrieval in terms of good accuracy, small number of bins and efficient computation.

Changes in illumination conditions modify the colors of an image and, affects the results of the color based algorithms. Swain and Ballard [196] have suggested the application of color constancy algorithm to illumination independent features. Their idea was modified by Funt and Finlayson in which [67] histograms of the derivative of the logarithm, code the length of the boundaries between colors, and proves an illuminant independent description of an image. Gevers and Smoulders have proposed various models and checked their invariance under changing conditions [70]. One of the color models proposed in [71], assumes dichromatic reflection and white illumination, and is independent of the viewpoint, surface orientation, illumination direction and intensity. It also discounts the effect of shadows and shadings. These algorithms are used satisfactorily for object recognition [69].

Color moments have been successfully used in many retrieval systems like, QBIC, [62], [194]. The first order(mean), the second (variance) and the third order(skewness) color moments have been proved to be efficient and effective in representing color distribution

of images. Although color moments provide a compact representation, it has lower discrimination ability compared to other sophisticated features. Color moments can have good application in the first pass, where the search space can be reduced before other sophisticated color features are used for retrieval. In the Recent past, color moment (CM) and the color variance of adjacent pixels (CVAP) has been combined to raise the quality of similarity measure for image retrieval in [44]. An image retrieval system which expresses the semantics associated with the combination of chromatic properties of color images is proposed in [56].

Other approaches for color image indexing are mentioned as follows: Binaghi et al., [33] proposed three fuzzy sets corresponding to similarity in lightness(L*), hue(h*) and chroma(C*) color dimensions. The similarity between two colors say, i and j is defined as a function of difference between the membership values, in a given color space.

Wavelet transforms have been used for effective color representation for image retrieval. In [37], multiresolution wavelet transform has been used in a modified CIELUV color space to compute multiresolution image signatures. The multiresolution wavelet decomposition, obtained using Haar wavelets are coded in signatures of predefined lengths, which are compared in the retrieval phase by applying a similarity measure, based on supervised learning.

An effective saliency-based color image indexing method has been proposed recently, in [110]. A color/texture signature by using jointly the well-known color correlogram extended to salient features and rotated wavelet filter responses is used as features for image retrieval. Experimental results are conducted by adopting a global salient approach and a local salient approach to show the effectiveness of the proposed scheme. A recent method used for assessing the contribution of Color in Visual Attention, for image retrieval has been proposed in [95]. Visual attention is the ability of a vision system, to rapidly detect potentially relevant parts of a visual scene , on which higher level vision tasks may be performed. A methodology consisting in comparing the computational saliency map with human eye movement patterns to assess the quantitative contribution of chromatic

features in visual attention is proposed.

Retrieval by texture based representations :

Texture contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment [30]. Effective representation of textures can be made from the statistical and structural properties of the brightness pattern. Textures have wide applications in pattern recognition and computer vision and also finds its way in image retrieval applications. Basically, texture representation methods can be classified into two categories: structural and statistical [75]. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules [125]. They tend to be most effective for regular type of texture patterns. The texture properties like coarseness, contrast, directionality, regularity and roughness, etc. are mostly represented based on statistical models. The Statistical methods, including Fourier power spectra, co-occurrence matrices [83], Tamura feature [197], Wold decomposition [64], Simultaneous Autoregressive model [135] Markov random field [57], etc. including multi-resolution filtering techniques such as Gabor [10] and Wavelet [3] transform, characterize texture by the statistical distribution of the image intensity. A brief review on texture representation and properties have been presented in [221]. Coefficients of a 2-D transform are used to indicate the correlation of a brightness pattern in the image. Coefficients of wavelet transforms can also be used to represent frequency properties of a texture pattern. Wavelet transforms have been widely used for texture representation [3], [4]. Different modeling of textures have widely been used in image retrieval applications. Computational approximation to meaningful visual texture properties or psychological aspects of texture like coarseness, contrast, directionality, line-likeness, regularity, roughness, etc. was proposed by Tamura et al., [197]. A similar representation [118] was obtained according to Wold decomposition corresponding to repetitiveness, directionality and complexity percepts. Retrieval based on texture similarity using Wold transform based approach, is used in Photobook system [161]. The texture

representation based on Wold transform is aiming to detect the presence of periodic textures, texture orientation, etc. Representation of textures at different resolutions, in order to model different granularities, with SARmodel has been proposed in [135] and known as MRSAR. As Tamura textural features capture the psychological aspects of texture, it has been successfully used in image retrieval application in MARS [151] and QBIC [62]. A Gabor wavelet decomposition model has been used for image retrieval by Ma and Manjunath [132]. The performance is reported fairly comparable to MRSAR method.

Some of the recent approaches are mentioned as follows. Recently, a statistical approach, using a novel ordinal co-occurrence matrix framework for the purpose of content-based texture retrieval is proposed in [159]. A new method for color texture retrieval using color and edge features is proposed in [213].

Due to the attractive advantages of multiresolution characteristics of wavelets, representation of textures using wavelets are extensively used in image retrieval applications. Wavelet transform leads to the development of adequate tools to characterize different scales of textures effectively. A novel histogram-based technique that is robust to the changes in image illumination levels, where retrieval is performed by comparing the parameters of histograms of the wavelet subbands is proposed in [131]. In the recent times, Cosine-modulated wavelet transform based technique for extraction of texture features has been proposed in [101]. The major advantages of Cosine-modulated wavelet transform is that, it involves less implementation complexity and produces good filter quality. A recent approach for texture image retrieval is proposed by using a new set of two-dimensional (2-D) rotated wavelet filters (RWF) and discrete wavelet transform (DWT) jointly. A new set of 2-D rotated wavelet improves characterization of diagonally oriented textures [100].

A very recent approach on retrieval of structured and random texture has been proposed by [52]. In this approach, a multiscale directional filter bank (MDFB) is designed to suppress the aliasing effect as well as to provide rotation -invariant features for texture characterization. Secondly, an entropy-based measure on energy signatures is proposed

to classify structured and random textures. A recent effective texture descriptor invariant to translation, scaling, and rotation for texture-based image retrieval applications is proposed in [185]. The proposed descriptor is obtained by first calculating the power spectrum of an original texture image for translation invariance and then the power spectrum image is normalized for scale invariance. Finally, modified Zernike moments are calculated for rotation invariance. Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler distance is proposed in [59]. A recent texture retrieval scheme, based on rotation invariant texture classification using Gabor wavelets has been proposed in [10]. Wavelet correlogram has been used for image indexing and retrieval in [144]. Rotation invariant features using Gabor wavelets, in which redundancy problem due to non-orthogonal decomposition of Gabor wavelets has been addressed is proposed in [145].

Image retrieval by shape based representations :

Shape features of objects or regions have been widely used in many content-based image retrieval systems [55], [172], [172]. In general, shape description can be categorized into either (a) Boundary-based methods. (b) Region-based methods. Boundary-based methods include polygonal approximation [11], finite element models [181], Fourier-based shape descriptors, etc. [9], [96]. Region-based methods include statistical moments [97], geometric properties of regions, like area compactness, elongatedness, etc. [58]. A good shape representation feature for an object should be invariant to translation, rotation and scaling. A survey on shape analysis techniques has been performed in [120]. Fourier descriptor is popularly used as one of the boundary based shape descriptors. Fourier descriptors describe the shape of an object with the Fourier transform of its boundary [219], [9]. The Fourier transforms of contour representations generate sets of complex coefficients, representing the shape of an object in the frequency domain. A recent work on Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance is proposed in [24]. A recent approach using Fourier descriptors in multiple scales, which improves

the shape classification and retrieval accuracy in compared to ordinary Fourier descriptors is proposed in [107].

Invariant moments are also popularly used as shape descriptors. A set of seven moments invariant to rotation, translation and scale have been proposed by [89], to describe shapes completely. A solution considering the issue of invariance with Zernike Moments for object recognition has been proposed in [97]. Mehre et al., have compared the performances of different shape representations in [140]. A digitization process, that preserves several properties of the object boundary, such as convexity and inflection in presence of noise or a boundary discontinuity is proposed in [77].

Representation of shape as modal deformations of a prototype object has been considered in [161]. In this method, shape description using finite element model and shape comparison between two FEM shape representations are made from deforming one elastic shape model to align with the other. Bimbo et al., have considered in [31], the shape similarity which accounts for the amount of deformation of a sketch. Such technique is based on elastic matching of sketched templates over the shapes. A retrieval method by shape similarity using local descriptors is proposed in [27] where shapes are partitioned into tokens in correspondence with their protrusions, and each token is modeled according to a set of perceptually salient attributes. Shape indexing is then obtained by arranging shape tokens into a suitably modified M-tree index structure. Evaluating shape similarity from finding corresponding between the points on two shapes has been solved in [26]. The correspondence problem has been solved by attaching a shape descriptor and the shape context to each point, where the shape context at a point captures the distribution of the remaining points relative to it. Given the point correspondences they also estimate the transformations that best aligns two shapes. A very recent approach, in which a contour based method for region based shape representation and retrieval is proposed in [218]. In this approach, a one dimensional signature function is generated from a two dimensional region shape and used for image retrieval. Shape representation from orientation of edge map has been proposed by Fariboez et al., in [128]. A shape-based image retrieval

method using salient edges has been proposed in [81]. A content-based image retrieval algorithm based on information of coded blocks which are used to find matches from an image database is proposed in [53]. A recent work using information of edge color histograms in images retrieval applications has been proposed in [204]. Combination of edges and interest points are also used for efficient feature detection, to increase the accuracy of the image retrieval system. A very recent work on object-based image retrieval using a method based on visual-pattern matching has been proposed in [54]. In this approach, visual pattern is obtained by detecting the line edge from a square block using the moment-preserving edge detector. A voting scheme based on generalized Hough transform is proposed to provide object search method, which is invariant to the translation, rotation, scaling of image data, and also invariant to orientation and position. A new adaptive edge detection approach based on predictive error values, for image content analysis has been proposed in [215], where the performance is shown to be better than some of the standard edge detectors. Co-efficients of wavelets have also been used for shape representation. A shape based image retrieval based on adaptive wavelet lifting approach has been proposed in [150]. In this approach, the feature vectors are computed based on moment invariants of detail co-efficients produced by adaptive lifting schemes. A recent approach, for affine invariant shape representation, combining approximation coefficients from two different wavelet families have been proposed in [175].

Graph based Shape matching approach has been used by many researchers. A hierarchical representation for shape analysis at multiple scales has been performed in [32] by Bimbo et al. In this approach, shape similarity is measured from traversing the graph from coarse to fine similarity matching. Graph based Shape matching done on structural descriptions, using a set of rules called shape grammars is done in [184], [182]. A recent geometry-based image retrieval scheme that makes use of projectively invariant features like cross ratio has been proposed in [168]. The results have been shown to be effective in retrieving images having man-made objects rich in polygonal structures like buildings, rail tracks, etc. Retrieval irrespective of view point and illumination changes is obtained using this

approach. A convexity measure for recognizing rectilinear structures has been proposed by [225]. This shape measure is applied on image database consisting of images, rich in structural description.

Although several approaches for measuring shape similarities have been proposed, but determining the similarity between 3D shape from 2D images, is a fundamental problem. A technique for extracting features for shape similarity for 3D object models has been presented in [191]. The method includes extraction of spatial arrangement from a 3D object surface followed by extraction of 2D shape features from the projection images using curvature distribution of model surfaces. 3D Zernike invariants used as descriptors for content based 3D shape retrieval has been proposed in [149]. These moments are computed as the projection of the functions defining the object onto a set of orthonormal functions within the unit ball and then used as features. Shape matching becomes a challenging problem when the object is occluded. This aspect of shape representation is useful in many applications. Recently a problem on partial shape matching using directed acyclic graph (DAG) has been proposed in [109].

Retrieval by Spatial similarity :

Spatial relations between objects in an image can contribute significantly to the description of its content. The relative positions like left, right, below above, etc. to an object may be used to capture meaningful semantic information from an image. Freeman [66] defined 11 primitive spatial relations between two objects like (left of, right of, above, below, behind, in front of, near, far, inside, outside, surround, etc). It has also been recognized that they are best described in an approximate (fuzzy) framework [102].

A new CBIR system which evaluates the similarity of regions one to the other, taking into account the spatial configuration of different regions is proposed in [63]. In this approach, the spatial structure of the regions is represented by means of fuzzy spatial relations, like horizontal and vertical disposal. A very recent histogram representation method called R-Histogram that extends the histogram of angles by incorporating both angles and la-

beled distances for representing topological spatial relations like, inside and overlap for its application in image retrieval is proposed in [208]. An image retrieval method, where multiple regions and their spatial relationship are used in comparing images and Hausdorff Distance (HD) is used to estimate spatial relationships between regions is proposed in [99]. Matching Spatial Relations Using db-Tree for Image Retrieval has been proposed in [116].

Spatial relationships modeled in the framework of fuzzy sets for defining topological properties like (set relationships, adjacency) and metrical relations (distances, directional relative position) have been analyzed and reviewed in [35]. Geometry based representation and graph matching has been popularly used to represent spatial relations which is proposed in [103]. Fuzzy spatial similarity has been used to assess the retrieval accuracy of the techniques.

Retrieval by Combination of features :

A single feature cannot be sufficient enough to generate a meaningful characterization of an image. As each feature tries to capture a particular aspect of an image content, a combination of features is generally used for more meaningful description of image content. A weighted histogram based combined approach, where the shape of an object is represented by histograms of edge directions and color with color histograms is presented in [91]. A region based approach, considering homogeneous regions with similar color and texture is proposed in Blobworld [43]. In their approach, image segmentation is performed by fitting a mixture of Gaussians to the pixel distribution in a joint color texture position feature space. Image querying is done based on the regions of interest. The QBIC [62] system supports users to retrieve images by color, texture and shape. In their method, the color features computed are, the 3D average color vector of an object or the whole image in RGB, YIQ, and Munsell color space, a 256-dimensional RGB color histogram. The shape features consist of shape area, circularity, eccentricity, major axis orientation and a set of algebraic moment invariants. The texture features used

in QBIC are modified versions of the coarseness, contrast, and directionality features proposed by Tamura. QBIC allows queries based on example images, user-constructed sketches or/and selected color and texture patterns. Virage is another CBIR system, that supports visual querying based on color, texture and spatial-relationship. The Virage Engine [12] provides a set of general primitives, such as global color, local color, texture and shapes. Apart from these, various domain specific primitives can be created when developing an application. Queries can be performed on various user-defined combinations of primitives. Photobook [161] implements three different approaches to constructing image representations for querying purposes, each for a specific type of image content such as, faces, 2D shapes and texture images. The matching algorithms include Euclidean, Mahalanobis, vector space angle, Histogram intersection, Fourier peak and wavelet tree distance as the distance metrics. Combining color and shape invariant features for image retrieval has been proposed in [72]. An approach to represent spatial color distributions using local principal component analysis (PCA), followed by a symmetry based saliency map and an edge and corner detector is proposed in [87]. An integrated region matching approach, where the regions are characterized by color, texture, shape, location and application for semantic sensitive classification of picture libraries, has been proposed in SIMPLIcity [206]. Wavelet based salient points are evaluated for image retrieval using the color and textural features in [121]. The concept of image-transform bootstrapping using transforms in the image space for application in the cases of scene matching has been proposed in [126]. A very recent image retrieval mechanism based on joint modeling of semantic label and visual feature distributions is illustrated in [41]. The proposed method is shown to be fairly robust to parameter tuning. Recently Content-based image retrieval (CBIR) system with relevance feedback, which uses the algorithm for feature-vector (FV) dimension reduction, is described in [217]. Instead of all FV components describing color, line directions, and texture, only their representative members describing FV clusters are used for retrieval. A memory learning framework for effective image retrieval, which uses combination of low-level features and learned semantics has been

proposed in [80]. A CBIR system which automatically generates a set of modifications by manipulating the features (color, texture, shape) of the query segments from relevance feedback is proposed in [8].

Retrieval using Fuzzy set theoretic approach :

In image interpretation, computer vision, structural recognition, etc. the management of imperfect information and imprecision constitutes an important issue. This calls for the framework of fuzzy sets. A fuzzy set framework is used for representing features that are intrinsically vague. If the objects are precisely defined, their relationships can be defined and computed in a numerical (purely quantitative) setting. On the other hand if the objects are imprecise, as is often the case, when they are extracted from images, then the semi-quantitative framework of fuzzy sets proves to be useful for their representations. In CBIR systems, the queries that are used to retrieve images can be broadly classified as primitive (based on features such as color, shape, and texture), logical and abstract (mostly Sketch-based and linguistic queries). Logical and abstract queries are sometimes known as semantic queries. Digital images are inherently fuzzy. Even a region appearing as homogeneous to human vision system has a graded composition in terms of gray levels [200]. Evaluating similarity between images is also a fuzzy concept owing to differences arising due to individual perception. It can be shown that fuzzy sets can be used to model the vagueness that is usually present in the image content, user query, similarity measure, relevance feedback, etc. Fuzzy set theory can be useful in building a more versatile CBIR system that can handle queries of different types. A fuzzy logic approach, UFM (unified feature matching) for region-based image retrieval is proposed in [51]. In their approach, an image is represented by a set of segmented regions, each of which is characterized by a fuzzy feature (fuzzy set) reflecting color, texture, and shape properties. As a result, an image is associated with a family of fuzzy features corresponding to regions. The overall similarity between images is defined as the similarity between two families of fuzzy sets. Fuzzy set theory is effectively used to

describe an image retrieval system called FIRST (Fuzzy Image Retrieval System) [103] which can handle exemplar-based, graphical-sketch-based, as well as linguistic queries involving region labels, attributes, and spatial relations. FIRST uses Fuzzy Attributed Relational Graphs (FARGs) to represent images, where each node in the graph represents an image region and each edge represents a relation between two regions. The given query is converted to a FARG, and a low-complexity fuzzy graph matching algorithm is used to compare the query graph with the FARGs in the database. The use of an indexing scheme based on a leader clustering algorithm avoids an exhaustive search of the FARG database. Quantitative Analysis of Properties of Spatial Relations of Fuzzy Image Regions has been proposed in [102]. Applications of Fuzzy Logic in Intelligent Systems for analysis of fuzziness in image information and scene analysis have been proposed by Pal et al., [152]. An Image retrieval scheme which expresses the semantics associated with the combination of chromatic properties of color images is proposed by Pala Bimbo et al.,in [56]. In their approach, fuzzy sets are used to represent low-level region properties such as hue, saturation, luminance, warmth, size and position. A formal language and a set of model-checking rules are implemented to define high-level properties which address concepts such as the perceptual quality of colors and the sensations that they convey. An image retrieval technique based on Fuzzy set theoretic evaluation of color similarity has been proposed in [33]. A very recent fuzzy approach for image retrieval on the basis of color features is presented in [34]. A fuzzy color is proposed, by using linguistic labels for representing the color information in terms of hue, saturation and intensity. Manipulation of fuzzy objects in an object-relational database system is also provided in [138] for handling flexible queries on the database. An image retrieval system Based on Fuzzy Color Histogram Processing is proposed in [212]. A fuzzy color histogram based image retrieval system is proposed in [79] to address the concerns like, the sensitivity to noisy interference such as illumination changes and quantization errors. The approach presents a new color histogram representation, called fuzzy color histogram (FCH), by considering the color similarity of each pixel's color associated to all the histogram bins through

fuzzy-set membership function. Better results have also been shown using this method, than using conventional approaches of color histograms. The problem of discretization in digital space has been handled with fuzzy set theoretic approach in [186]. The method presents several measurements on digitized 2D and 3D objects with fuzzy borders. The performance of surface area, volume and roundness measure estimators for digitized balls with fuzzy borders, are obtained from fuzzy segmentation of objects. The method is used as better estimate of analogous quantities of the corresponding real objects, than those obtained for, crisply segmented objects. An image retrieval system, posing queries at semantic level is proposed in [104]. The method proposes a novel fuzzy approach for mapping the fuzzy database while extracting the color features from image and assigning the weights to this fuzzy content when calculating the similarity between the query image and the images in database.

1.3.3 Scope of the thesis

The objective of the thesis is to develop some suitable low-cost feature extraction algorithms for specific application in Content-Based Image Retrieval. Emphasis have mainly been made, on developing appropriate methodologies using soft decision making with Fuzzy Logic, in the feature extraction and feature evaluation modules in order to handle certain imprecise situations which are addressed in Chapters 1, 2, 3, 4.

Wavelets have been recently used as a powerful tool for extracting significant features both in spatial and frequency domain. There are various applications of wavelets in the area of texture analysis [106], [6] such as document image analysis [5], remote sensing images [7], applications in the area of data hiding [129], etc. It has been decided, to choose wavelets for exploring its suitability for extraction of features in CBIR applications. Use of wavelet packets for segmenting images into homogeneous regions have been explored. Some Fuzzy Topological properties have also been considered to capture the spatial relations between the segmented regions. These properties are used to characterize an image in chapter 5.

Support Vector Machines are widely used in various application domains such as pattern recognition, datamining [158], [142] image processing etc. [143] for its implicit learning capabilities. Support vector machine has been used as a tool for application in image processing. This has been used in extraction of gray level corners in chapter 3.
The organization of the thesis is depicted in the flow diagram of Fig. 1.2 outlining the contributions of the thesis. The detailed results and investigations are summarized under each chapter heading.

1.3.4 A brief introduction about different tools used in developing algorithms

A brief discussion about Fuzzy logic, Wavelets, Support Vector Machines, which are used in different stages in developing the algorithms has been furnished. Although, a detailed discussion on the above- mentioned topics are beyond the scope of this work, only the relevant portions that will be required in the rest of the thesis will be discussed.

Wavelet :

The term **wavelet** means **small wave**

The basic idea of wavelet transform is to represent an arbitrary function $f(x)$ as a set of linear combination of a set of wavelets or basis functions [130]. These basis functions are obtained from a single prototype called Mother wavelet by dilations (scaling) and translation called shift. The mother wavelet is a prototype for generating the other windowed functions. Small means the window of the function is of finite length (compactly supported). Wavelet transform use basis functions to analyze and reconstruct a function. Wavelet cut up data into different frequency components and then study each component with a resolution matched to its scale.

There are two types of wavelet transform (a) Continuous wavelet transform (CWT) (b) Discrete wavelet transform (DWT)

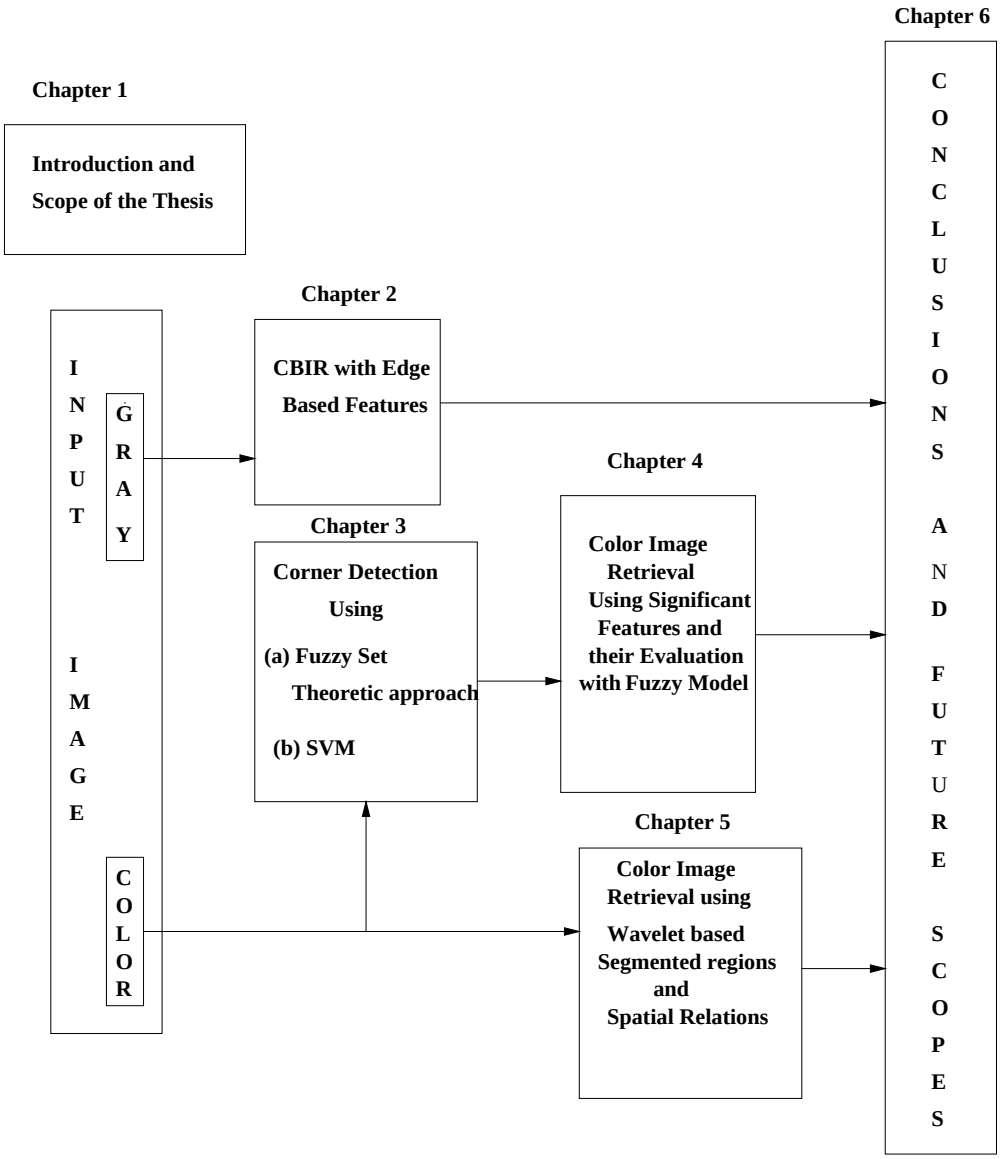


Figure 1.2: Organization of the thesis.

The CWT is given by,

$$CWT_x^{\psi}(\tau, s) = \Psi_x^{\psi}(\tau, s) = \frac{1}{\sqrt{\mid s \mid}} \int x(t)\psi^*(\frac{t-\tau}{s})dt \tag{1.1}$$

$\tau \rightarrow$ translation $s \rightarrow$ scale parameter

$\psi(t)$ is the mother wavelet.

It is necessary to discretize the transform for computational purpose. The full discrete wavelet expansion of a signal that forms an orthonormal basis for $L^2(R^2)$ is given as,

$$\psi_{j,k} = s^{-j/2}\psi(s^{j/2}t - k\tau) \tag{1.2}$$

The effect of discretizing the wavelet is that the time scale space is sampled at discrete intervals. s is usually chosen as 2 and $\tau = 1$ so that the sampling of the frequency axis and the time axis correspond to dyadic sampling.

Multiresolution analysis (MRA) analyzes the signal at different frequency with different resolution. A brief description about dyadic wavelet transform and packet transform is presented in the Appendix.

Fuzzy Sets :

A fuzzy set A in a set of points $R = \{r\}$ is a class of events with a continuum of grades of membership. It is characterized by a membership function $\mu_A(r)$ which associates a real number $\mu_A(r) \in [0, 1]$ with each element of R. $\mu_A(r)$ at r represents the grade of membership of r in A. Formally, a fuzzy set A with its finite number of supports $r_1, r_2, ..., r_t$ is defined as a collection of ordered pairs

$A = \{(\mu_A(r_i), r_i), i = 1, 2, ..., t\}$,

where the support of A is an ordinary subset of R and is defined as

$S(A) = \{r | r \in R and \ \ \mu_A(r) > 0\}$.

Here $\mu_i = \mu_A(r_i)$, the grade of membership of r_i in A, denotes the degree to which an event r_i may be a member of A or belong to A. It is to note that $\mu_i = 1$ indicates the strict containment of the event r_i in A. If, on the other hand, r_i does not belong to A then $\mu_i = 0$.

Fuzzy logic is based on the theory of fuzzy sets [216]. Unlike classical logic, it aims at modeling the imprecise (or inexact) modes of reasoning and thought processes (with linguistic variables) that play an essential role in the remarkable human ability to make rational decisions in an environment of uncertainty and imprecision. This ability depends, in turn, on our ability to infer an approximate answer to a question based on a store of knowledge that is inexact, incomplete, or not totally reliable. In fuzzy logic everything, including truth, is a matter of degree. Zadeh has developed a theory of approximate reasoning based on fuzzy set theory. By approximate reasoning we refer to a type of reasoning that is neither very exact nor very inexact. This theory aims at modeling the human reasoning and thinking process with linguistic variables in order to handle both soft and hard data, as well as various types of uncertainties. Many aspects of the underlying concept have been incorporated in designing decision-making systems.

Because fuzzy sets are a generalization of the classical set theory, the embedding of conventional models into a larger setting endows fuzzy models with greater flexibility to capture various aspects of incompleteness or imperfection (*i.e.*, deficiencies) in whatever information and data are available about a real process. Assignment of membership functions of a fuzzy subset is subjective in nature, and reflects the context in which the problem is viewed. It cannot be assigned arbitrarily. In many cases, it is convenient to express the membership function of a fuzzy subset in terms of standard S and π functions [155]. It is to note that fuzzy membership function and probability density function are conceptually different. Probabilities convey information about relative frequencies of objects while fuzzy membership represents similarities of objects to imprecisely defined properties.

Support Vector Machines :

Support Vector Machines are a general class of learning architecture inspired from statistical learning theory, that performs *structural risk minimization* on a nested set structure of separating hyperplanes [202]. A principle method for choosing a learning machine for a given task, is the essential idea of structural risk minimization. Given a training data, the SVM training algorithm obtains the optimal separating hyperplane in terms of gener-

alization error. For a labeled training data say, $\{x_i, y_i\}$, where i = 1.... l, $y_i \in \{-1, 1\}$, $x_i \in R^n$. Suppose we have some hyperplane which separates the positive from the negative examples (a separating hyperplane). The points x which lie on the hyperplane satisfy w . x + b = 0, where w is normal to the hyperplane, $|b|/||w||$ is the perpendicular distance from the hyperplane to the origin, and $||w||$ is the Euclidean norm of w. Let $d_+ + (d_-)$ be the shortest distance from the separating hyperplane to the closest positive (negative) example. The margin of a separating hyperplane is defined to be $d_+(d_-)$. For the linearly separable case, the support vector algorithm simply looks for the separating hyperplane with largest margin.

1.3.5 Chapter 2 : Image Retrieval Using Edge Based Features [19], [18]

In Chapter 2, A CBIR system based on fuzzy edge based features has been proposed. The work has been started with the CBIR problem for application in the case of gray level images. Among the various low level features, shape based features have good image retrieval applications [30]. Image retrieval using edge based features have been proposed in [42], [223]. However, the common problem in most of the edge related methods are that the features are sensitive to noise and transformations like rotation, translation, scale, etc. Moreover, in CBIR the users are more interested in retrieval of results according to similarity (closeness) than equality (exactness). Also digital images are inherently fuzzy in nature [200].

Motivated from these ideas, fuzzy set theoretic approach has been adopted for appropriate handling of the uncertainties, arising in the feature extraction process.

The task involves extraction of the candidate edge pixels from the border regions between the uniform intensity surfaces termed as Plateaus. The extracted candidates are assigned membership values, by measuring the local contrast in the neighborhood, in order to characterize its edge strength. The edge points are characterized as strong, medium and weak

type based on the assigned membership values. Thresholding above different membership values results in multilevel fuzzy edge map consisting of weak, medium and strong edge points. Fuzzy Topological features are computed from each thresholded plane, for constructing the feature vector.

The key advantage of the proposed feature is that the Topological features computed from each plane is invariant to rotation, translation and scaling by definition. The local contrast information is also embedded in the feature value in addition to the shape property. It is expected that relevant images would generate better results, when matched at each level. Euclidean distance is computed between the feature vectors to evaluate similarity between images.

The algorithm is tested with different query examples of gray scale images. The algorithm performs well under transformations like rotation, translation, scale and noise. The algorithm also works fairly well in case of retrieving logo images from (USPTO) trademark databases.

1.3.6 Chapter 3 : Extraction of Gray Level corners : Perceptually Significant Points of Interest [15], [14], [17]

Chapter 3, focuses on feature extraction. To achieve still better representation of images, efforts have been made to obtain a subset of edges which may be salient and and rich in structural information. Points having high curvature on the planar boundaries, which are commonly known as corners can be used efficiently to represent different objects of an image. Corners characterize visually significant portions of an image and play an important role in shape perception. Features in the locality of the corner regions can be used to represent an image.

Classical corner detectors [84] are mostly used for the purpose of characterizing images [121]. Classical corner detectors have some drawbacks when applied to various natural images for computation of local features, because visually significant points are

not always exactly at corners and the position of corners may also vary under several real life situations.

Real life image data are always imprecise due to inherent uncertainties that arise in the imaging process, such as defocusing, wide variations of illuminations, etc. This results into blurred, over exposed or under exposed conditions. As a result, precise localization and detection of corner is difficult under various imperfect situations. Motivated from these ideas, a fuzzy set theoretic approach has been used for developing the corner detection algorithm. Dominant curvature points can be represented as entities of a fuzzy set where exact corner point is also a member.

The proposed method involves extraction of the fuzzy edge map. The points on the edge map are categorized as straight edge points, corner type points, etc. taking into account the variations in the local properties with its neighbors. By applying fuzzy rules on the categorized points, the significant curvature regions are separated out. Cluster of points near the curvature junctions were obtained, whose centroid almost depict the true corner point. These set of points are taken for computation of local features and considered as the source of similarity evaluation.

Results have been found satisfactory when tested on images that had undergone varying imaging conditions like blurred, overexposed, underexposed, etc. The results are also compared with standard algorithms of Harris detector [84] and SUSAN method [190] to prove the efficiency of the proposed feature extraction algorithm.

Although the proposed fuzzy rule based approach presents a simple corner detection strategy, it lacks learning capabilities. In order to incorporate implicit learning capability for obtaining good performance for a wide range of images, a Support Vector Machine is designed using the fuzzy edge map as input. The critical points in the classification problem correspond to the corner points in the proposed method.

1.3.7 Chapter 4 : Retrieval of Color Images using Significant Features and their Evaluation with Fuzzy Model [16], [21], [23], [22]

In chapter 4, considering the extracted visually significant points based on fuzzy approach as obtained in chapter 3, a reduced representative set of the original image is obtained. With an aim to enhance the accuracy for retrieval of relevant images, the extracted cluster of high curvature points are considered for computation of features. Natural images are mainly heterogeneous, where different parts of the image have different characteristics. As a result, global features are inadequate to generate an effective representation of an image. Local features around visually significant points can be used for generating discriminating features for image matching and retrieval applications [141], [121].

Considering the extracted point set as the reduced representative of the original image, illumination, viewpoint invariant color features are computed from the extracted locations of each component plane for evaluating similarity between images. A relevance feedback mechanism, using fuzzy entropy based feature evaluation is provided for interactive enhancement of retrieval accuracy. The user marks the significant and insignificant outcome of the retrieved images. Based on this feedback, individual features are weighted and next set of results are retrieved.

The performance of the proposed methodology are tested on different type of images from a general purpose image database. The results obtained from the proposed method are also compared with the results obtained from [51], [176]. Successive improvement of the results after relevance feedback has also been checked. The results are also validated with MPEG-7 standard visual content descriptors.

1.3.8 Chapter 5 : Image Retrieval Using Wavelet Based Segmented Regions and their Fuzzy Spatial Relations [20]

In chapter 5, a region based approach for image retrieval has been proposed. Although, the features computed from the visually significant points of interest can serve as a considerably good measure for capturing the salient aspects of an image, but to capture the properties of each object and to generate a representation quite rich in semantic information, the objects need to be isolated. This requires suitable segmentation algorithms capable of segmenting the image into individual regions.

An algorithm to segment an image into fuzzy regions, based on the coefficients of multiscale wavelet packet transform has been developed. The features comprise the colored texture properties of the segmented image and the spatial relationship between the images. The objective is motivated towards better capturing of the semantic information associated with the scene. It has also been shown that the system performs fairly well when tested on general purpose image database. The results are tested on various example images, and compared with [176], [206] to study the efficiency of the system in the case of retrieving semantically relevant images.

Conclusion

The concluding remarks along with the scope for further research are presented in Chapter 6.

Appendix

In chapter 7, Mathematical analysis on wavelet packets used in the proposed feature extraction has been discussed.

Chapter 2

Image Retrieval Using Edge Based Features

2.1 Introduction

The common problem in Content Based Image Retrieval (CBIR) is selection of features. Image characterization with lesser number of features involving lower computational cost is always desirable. Features like color, texture, shape, spatial relationship among different objects and their combinations are generally used for the computation of multidimensional features [78]. Edge is a strong feature used for characterizing the shape of different objects within an image. Therefore, investigation of a CBIR system where edge features are considered for retrieval purpose may lead to an important research issue.

In this chapter, an image retrieval scheme using edge based features has been proposed. Application of the proposed method has been made to retrieve gray level images and logo images. Image retrieval based on matching between shapes, is broadly classified as region based and boundary based methods. Some of these methods have been discussed in Chapter 1. Region based methods mostly segment an image into different homogeneous regions, where the regions may correspond to different objects within an image. Such pre-segmentation is necessary whenever a scene is characterized by different

objects in it. However, segmentation becomes less important when there are large number of fragmented objects in a scene or no specific objects. In region based segmentation, the cost of computation depends upon the method used, whereas accuracy depends upon the apriori information of the total number of classes present in the scene. Boundary based methods utilize the properties of boundaries for characterizing an image.

As an alternative to segmentation and detailed object representation, the objective of research presented is this chapter, is to develop new algorithms, which explore structure of different objects in an image. A technique for extracting edge map of an image is proposed, which is followed by computation of Topological feature (like fuzzy compactness). This global feature incorporates gray level as well as shape information of the edge map, which is subsequently used in image retrieval applications.

From the available literature in CBIR it is seen that, there are methods [91], [223], [205] in which edge points of a scene can be used for indexing images. Edges characterize gray level discontinuities of an image. Edges that appear due to discontinuities between the regions are generally of strong value. Edges that appear due to variation within regions or due to defects of imaging give rise to weak edges. Region boundaries are characterized by strong edges and approximately represent the shape of the regions. Using directional histogram of edges, Jain and Vailaya [91] proposed a technique for shape comparison. However, histograms are not always very discriminative features for comparing images, because dissimilar images may also give rise to similar histograms. Ravela et.al proposed a method to characterize visual appearance for determining global similarity in images with geometric features [171]. Zhou and Huang [223] proposed structural features extracted from edge maps for CBIR. Histogram of blurred images are also used for defining shape features because these are strongly influenced by the original shapes of different regions in it [58]. Bruno et al., [38] defined multiscale families of feature calculated from the convolution of the input image with a set of multiscale Gaussian function. A recent work on edge extraction has been reported by Chang et al., [46]. In their approach, the edge orientation of each block in DCT domain has been classified in total five di-

rectional edge patterns, which are subsequently used in image matching applications. A recent shape-based image retrieval method using salient edges has been proposed in [81]. The parametric edge description, using general construction of functions whose moments serve to locate and characterize step edges within an image has been proposed in [163]. Using edge detection method, Wang et al., [205] presented a technique in which edges of the images can be extracted by wavelet transform. The edge images are then classified based on the invariant moments.

Although minimum number of features are always a desirable property for characterizing images, but single feature may not be sufficient for achieving desired accuracy. Color and texture based similarity measures without shape information may fail to produce desired results. Image retrieval only on shape information is difficult, in case the objects are occluded or having different projections onto different 2D planes. Sensitivity to noise may even cause inaccuracy in shape description. In CBIR, shape information is generally combined with other features like, gray level information, color, texture, etc. for characterizing an image. Combining shape and color in invariant fashion have been reported in [70]. Achard et al., proposed a shape based feature combining gray level information [2] by thresholding the image at different gray level. Compactness of the object is computed over each gray level for shape comparison. This method is computationally very expensive and cost is proportional to the number of gray levels considered.

One important aspect, which need to be addressed, is how to model the vagueness that is usually present in the image. As a result, the trend in CBIR research is to develop efficient features for simpler characterization of images and related matching techniques so that it can handle real life images [187], [78]. Real life images are inherently fuzzy due to several uncertainties arising in the imaging process [200]. Moreover, measuring visual similarities between images highly depend upon subjectivity of human perception of image content. As a result, fuzzy image processing for extracting visual features find an important place in image retrieval applications. It is desirable that, the selected features should not be computationally expensive, should be noise tolerant and invariant to

rotation, translation, scaling, etc.

Motivated from these ideas, a new feature extraction algorithm has been developed, which exhibits some of the mentioned properties. The proposed feature generates an overall description of an image, considering significant information of the edge maps. Using these edge maps, the issue of evaluating overall similarity between real images without segmenting the images has been addressed. The performance of the features has also been tested on images having objects of regular shapes e,g. (logo images of USPTO database).

The remaining Sections are organized in the following way. Section 2.2 describes briefly the mathematical model used in this work. Section 2.3 represents the detailed methodology of the proposed algorithm. Section 2.4 deals with the implementation details of the algorithmic steps. The experimental results are discussed in Section 2.5. Section 2.6 gives a conclusion.

2.2 Image properties used in proposed feature extraction

In real life images, inherent fuzziness mainly arises due to different imaging conditions. Even for perfectly homogeneous objects the corresponding images will have graded composition of gray levels [200] due to imperfection of imaging. As a result, uncertainties exists in locating the exact position of the boundary which separates different objects within an image. These vagueness are required to be modeled in the feature extraction process.

The basis steps involve extraction of the possible edge candidates using the concept of Top and Bottom of the intensity surface of a smoothened image. The extracted edge candidates are assigned gradient membership value to characterize its edge strength, within the range of (0.0. to 1.0). This is done by measuring the relative intensity differences over a small neighborhood, around a pixel. The selected points are categorized as Weak, Medium and Strong edge pixels based on their gradient membership values. Multilevel

thresholding is performed to segregate the edge pixels. As a result, fuzzy edge maps are obtained which consist of different type of edge pixels, at different thresholds. Intuitively it appears that more realistic matching between two images may be obtained, when the quantitative information as obtained from different type of edges match simultaneously. The feature vector elements include fuzzy compactness value of the edge map obtained at different levels. This feature is invariant to Rotation, Translation and Scaling by definition. The dissimilarities between two images are evaluated by measuring the Euclidean distance between the feature vectors.

In the following subsection, various membership functions used in fuzzy set theoretic approach and some of the fuzzy geometrical properties which are used in the process of feature extraction will be discussed in brief.

2.2.1 Image as fuzzy sets

An image X of $M \times N$ dimension and L levels can be considered [216] as a fuzzy subset A' in a space of points X=$\{x\}$. Each point in X can be characterized by a membership function $\mu_A(x_{mn})$ which can have a value in the interval [0,1]. $\mu_A(x_{mn})$ denotes the degree to which an event x_{mn} belong to A.

A = $\{(\mu_A(x_{mn}), x_{mn}\}$ $m = 1, 2, ..M, n = 1, 2, ..N$ where $0 \leq \mu_A(x_{mn}) \leq 1.0$.

This kind of image representation is useful in handling the uncertainties arising out of gray level as well as spatial digitization [155]. A fuzzy subset (A) is defined in terms of the membership values between [0-1].

One of the most widely used mapping function to do fuzzification for converting a digital image to corresponding fuzzy subset A is the standard S type or Π type function.

The Standard S-type function is given by

$$\begin{aligned} S(x; a, b, c) &= 0 && x \leq a \\ &= 2 \times \{\tfrac{(x-a)}{(c-a)}\}^2 && a \leq x \leq b \\ &= 1 - 2 \times \{\tfrac{(x-c)}{(c-a)}\}^2 && b \leq x \leq c \\ &= 1 && x \geq c \end{aligned} \tag{2.1}$$

where, b = (a+c)/2

Fig. 2.1 shows its graphical representation, where the parameter b is the cross over point, i.e., S(b;a,b,c)= 0.5. Similarly c is defined as the shoulder point at which $S(c;a,b,c)$=1.0 and a is the feet point i.e. S(a;a,b,c)= 0.0.

The Standard Π function is given by

$$\begin{aligned} \Pi(x;b,c) &= S(x;c-b,c-b/2,c) \qquad x \leq c \\ &= 1 - S(x;c,c+b/2,c+b) \quad x \geq c \end{aligned} \tag{2.2}$$

It should be mentioned that the assignment of membership is subjective in nature and reflects the context at which the problem is viewed. Other form of membership function which is exponential in nature as given in [155] is expressed as follows :

$$\mu(x) = \exp\left[-d(X,C)\right] \tag{2.3}$$

where d(X,C) is the distance of an element with feature vector $X = [x_1, x_2, \cdots, x_n]$ from a prototype $C = [c_1, c_2, \cdots, c_n]$ of a class and $d(X,C) \geq 0$. The availability of the mathematical function depends upon the nature of the prototype. The form of exponential type function is shown in Fig. 2.2(a), (b).

fuzzy alpha cut : A fuzzy subset can be divided by suitable thresholding of membership values around the range of interest. The fuzzy alpha-cut, s_α comprises all elements of X, whose degree of membership in S is greater or equal to α where

$$s_\alpha = \{x \in X : \mu_A(x) \geq \alpha\} \tag{2.4}$$

where $0 \leq \alpha \leq 1.0$

Fuzzy geometrical properties

Rosenfield [173] proposed certain Geometrical properties based on fuzzy set theoretic approach for characterizing regions of an image. Of them the concept of connectedness has been used in our experiment.

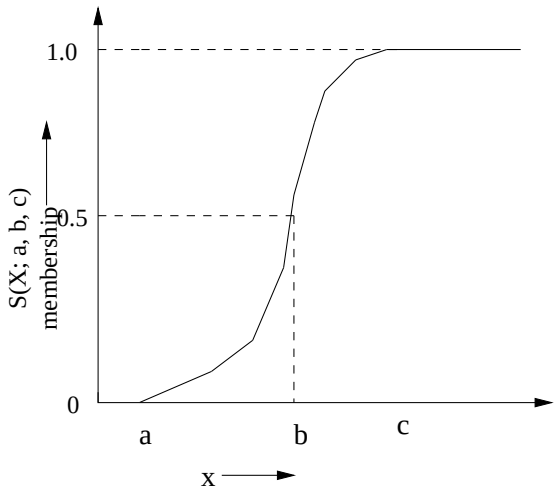
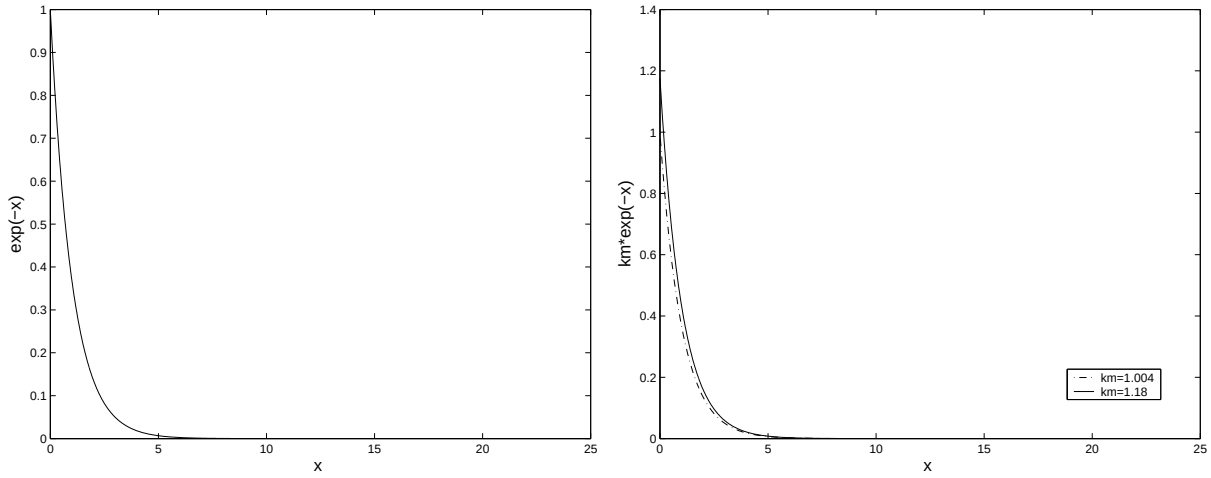


Figure 2.1: Standard S-type function.



(a) (b)

Figure 2.2: (a) Exponential function. (b) Transformed exponential function.

Degree of Connectedness : let μ be a mapping from S into [0,1] i.e., a fuzzy subset of S. Let $\rho : P = p_0, p_1, ... \cdots p_n = Q$ be any path between two points (P,Q) of S. The path strength $s_\mu(\rho)$ of ρ with respect to μ is defined as

$$s_\mu(\rho) = \min_{0 \leq i \leq n} \mu_\rho(p_i) \tag{2.5}$$

The Degree of connectedness of P and Q with respect to μ is defined as

$$C_\mu(P, Q) = \max_\rho(s_\mu(\rho)) \tag{2.6}$$

where max is taken over all paths from P to Q.

Proposition : If μ is into [0,1] and $\{P \mid P \in S \text{and} \mu(p) = 1\}$ then $\mu(\rho) = 1$ consists entirely of points of S and $C_\mu(P, Q) = 1$ if there exists a path from P to Q all of whose points are mapped to 1 by some property in μ. This condition generalizes the non fuzzy concept of connectedness [155].

Fuzzy Compactness :

For a $M \times N$, μ_{mn} array the fuzzy compactness comp(μ) is defined as [156].

$$comp(\mu) = \frac{a(\mu)}{p^2\mu} \tag{2.7}$$

where $a(\mu)$ and $p(\mu)$ are the fuzzy area and perimeter of (μ). Where $a(\mu)$ and $p(\mu)$ are given by eqn.2.8 and 2.9

$$a(\mu) = \sum_m \sum_n \mu_{mn} \tag{2.8}$$

$$p(\mu) = \sum_{m=1}^{M} \sum_{n=1}^{N-1} |\mu_{mn} - \mu_{m,n+1}| + \sum_{n=1}^{N} \sum_{m=1}^{M-1} |\mu_{mn} - \mu_{m+1,n}| \tag{2.9}$$

For example we consider a 4x4 μ_{mn} array.

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \alpha & \beta & 0 \\ 0 & 0 & \beta & \gamma \\ 0 & \delta & 0 & 0 \end{bmatrix}$$

here $a(\mu) = \alpha + 2\beta + \gamma + \delta$ and $p(\mu) = [\alpha + |\beta - \alpha| + \beta + \beta + |\gamma - \beta| + \delta + \delta] + [\alpha + \alpha + \delta + \beta + 0 + \beta + \gamma + \gamma]$ where $\alpha, \beta, \gamma, \delta$ are the membership values with $1 \geq \alpha, \beta, \gamma, \delta \geq 0$

2.3 Proposed technique for CBIR

The proposed feature extraction technique could be explained using the building block shown in Fig. 2.3. The algorithmic steps involve extraction of possible edge candidates, which are obtained by employing the properties of Plateaus. The extracted candidates are assigned membership values, $\mu_m(P)$ and $\mu_d(P)$. The memberships $\mu_m(P)$ and $\mu_d(P)$ are used to estimate the edge strength and connectivity respectively, which in turn are used to handle the uncertainties in determining the strength and location of the edge point.

The formulations and definitions used in the feature extraction process are described in the following subsections.

2.3.1 Identification of extremas

Plateau Top and Plateau Bottom :

Edges in an image are the transitions between two uniform intensity surfaces. The uniform intensity surfaces are defined in terms of **Plateaus**. Plateaus are defined as connected subsets of uniform intensity pixels. This definition can be taken as the special case of **fuzzy plateaus** introduced by Rosenfield [173].

Plateau : Let S_1 denote the set of all pixels in an image. The pixels P,Q $\in S_1$. By a

plateau in S_1, is meant a maximum connected subset S_p on which the intensity (I) has a constant value. In other words $S_p \in S$ is a Plateau if

(i) S_p is connected. (ii)$I(P) = I(Q)$ for all P,Q $\in S_p$ (iii) $I(P) \neq I(Q)$ for all pair of neighboring points $P \in S_p$ and $Q \notin S_p$

Clearly $P \in S$ belongs to one Plateau.

We call the Plateau S_p a Top if its gray value is a local maximum i.e. $I(P) \geq I(Q)$ for all pairs of neighboring point $P \in S_p$ and $Q \notin S_p$. Similarly we call S_p a Bottom if its gray value is a local minimum. The gray values of all elements of **Plateau Top** is greater than the gray value of any element of the border and the gray values of any element of a **Plateau Bottom** is less than gray values of the border. Let S_{pt} be a Plateau Top and S_{pb} a Plateau Bottom. The pixels in border region B(S_{pt}, S_{pb}) between a Plateau Top and Plateau Bottom can be defined as the points which are eight neighbors of at least one element of S_{pt},S_{pb}. Considering a 3×3 neighborhood around a pixel, Uma Shankar et.al proposed a method for characterizing extreme components of an image [49]. The pixels are labeled as pels of a (a) **Plateau Top** for which the pixel is greater or equal to maximum in its 3×3 neighborhood (b) **Plateau Bottom** for which pixels are less or equal to minimum in its 3×3 neighborhood.(c) The remaining pels are **Candidate pels** for forming fuzzy gradient map.

Employing the mentioned properties, the candidate edge pixels are extracted.

$$\begin{bmatrix} 2 & 4 & 8 \\ 3 & \mathbf{8} & 7 \\ 6 & 6 & 6 \end{bmatrix}$$ Plateau Top

$$\begin{bmatrix} 4 & 2 & 3 \\ 6 & \mathbf{2} & 3 \\ 5 & 4 & 2 \end{bmatrix}$$ Plateau Bottom

$$\begin{bmatrix} 2 & 4 & 8 \\ 4 & \mathbf{8} & 9 \\ 7 & 7 & 9 \end{bmatrix}$$ Candidate pel

2.3.2 Membership formulation

Membership function $\mu_m(P)$ for determining fuzzy gradient map :

The modified form of membership function of eqn. 2.3 has been used to deal with this problem. The border region of the plateaus of the blurred (Gaussian) image is taken as input for finding the fuzzy gradient map and the total dynamic range of 0 to 256 is considered.

Around each candidate pixel, the neighboring pixels are situated in eight distinct directions at a step of 45 degrees. The line joining two opposite pixels which makes four distinct lines are considered. The gradient membership function for an edge point is given in equation below.

$$\mu_m(P) = k_m \exp(-x) \tag{2.10}$$

Considering Fig. 2.4, if d_{ac} is the gray level difference between the opposite neighbors about the candidate pixel in a direction say, (0,4) and d_{al} is the gray level difference in the perpendicular direction say (2,6) then $x = \frac{1+d_{al}}{1+d_{ac}}$. We compute the ratios say (x_1, x_2, x_3, x_4) in four direction pairs which are explained below.

In a window of eight neighborhood, an edge pixel will have maximum gray level difference in a direction, perpendicular to its true edge direction. The edge direction should point along the minimum difference direction [75].

Let I_{mn} be the gray level of the point P(m,n). If $d_{al} = |I_{m,n-1} - I_{m,n+1}|$ then $d_{ac} = |I_{m-1,n} - I_{m+1,n}|$ vice-versa as shown in Fig. 2.4. The values of (x_1, x_2, x_3, x_4) may be obtained from the local neighborhood of 3×3 window as follows : x_1=$\frac{1+|I_{m,n-1}-I_{m,n+1}|}{1+|I_{m-1,n}-I_{m+1,n}|}$, x_2=$\frac{1+|I_{m-1,n}-I_{m+1,n}|}{1+|I_{m,n-1}-I_{m,n+1}|}$, x_3=$\frac{1+|I_{m-1,n-1}-I_{m+1,n+1}|}{1+|I_{m+1,n-1}-I_{m-1,n+1}|}$, x_4=$\frac{1+|I_{m+1,n-1}-I_{m-1,n+1}|}{1+|I_{m-1,n-1}-I_{m+1,n+1}|}$

The maximum value computed from the four direction pair is used to compute gradient membership of the candidate pixel.

$$x = max[x_1, x_2, x_3, x_4] \tag{2.11}$$

The value of k_m can be estimated from the following condition. Since $\mu_m(P)$ is ranging between [0,1] therefore the minimum value of $d_{al} = 0$ and maximum value of d_{ac} say (L) should map the membership value to 1. Different values of L should give rise to different value of k_m hence different crossover point i.e., ($\mu_m(P) = 0.5$). The strongest edge pixel at any point (m,n) is supposed to possess gray level difference $\simeq$256 in one direction and nearly zero in its perpendicular direction and the edge should points along the minimum difference direction. Putting $d_{ac} = 255$ and $d_{al} = 0$ in eqn.2.10 value of k_m obtained is $\simeq$ 1.0039. Higher membership values characterize strong edge pixels while lower membership values will characterize weak edge pixels. This has been depicted in the multilevel plot as shown in Fig. 2.6. Changing the value k_m by a factor is in effect changing the pixel contrast [155] i.e., transforming the membership value.

It has been mentioned in section 2.2.1 that the choice of membership function is problem dependent. Here, a monotonic type Exponential function has been chosen for suitable representation of the ambiguities of the set, computed from gray level contrast. With such selection, the nature of the curve is not very steep in the range of interest, such that small variation in gray level difference, may not result in large variations of membership values.

Membership function $\mu_d(P)$ for determining connectedness of the edge points :

The membership for connectedness is also represented by an exponential function. We consider a $M \times N$, $\mu_m(P)$ array. In the neighborhood of a pixel p(m,n) if there exits a pixel $p(m', n')$ such that,

$$|\mu_{mn} - \mu_{m'n'}| \leq T \tag{2.12}$$

are considered similar in magnitude. Here, T is the non-negative threshold. Again if,

$$|ang(m,n) - ang(m',n')| \leq A \tag{2.13}$$

are considered similar in direction [75]. Here, A is the angle threshold. These two properties are combined to link similar pixels and the degree of linking is expressed by a new membership function as expressed in eqn. 2.14.

The number of pixels possessing $\mu_m(P)$ above some threshold value (≥ 0.5) and ≤ 1.0 and simultaneously having directions which didn't differ more than 45 degree in a neighborhood of 3×3 or 5×5 window size are counted. The membership function is of the form.

$$\mu_d(P) = k_d \exp(-x) \tag{2.14}$$

where $x = \frac{1}{n1+n2+n3}$

n1 denotes is the number of edge pixels possessing same angle to that of the candidate pel, n2 is the number of edge pixels having +45 degree difference with the candidate pixel, n3 is the number of edge pixels having angle -45 degree difference with the candidate pixel. The value of k_d is estimated by the maximum number of connected pixels say K, along any of the four directions within the defined window. With $x = \frac{1}{K}$ maps the membership to 1. Changing the value of k_d will in effect change the number of connected pixels over a defined window.

Similarity metric

Euclidean distance metric is used to compute the dissimilarity value between two feature vectors. If $X = [x_1, x_2, \cdots, x_i]$ and $Y = [y_1, y_2, \cdots, y_i]$ are two feature vectors then the Euclidean distance metric between X,Y is given by,

$$E_d = \sqrt{\sum (x_i - y_i)^2} \tag{2.15}$$

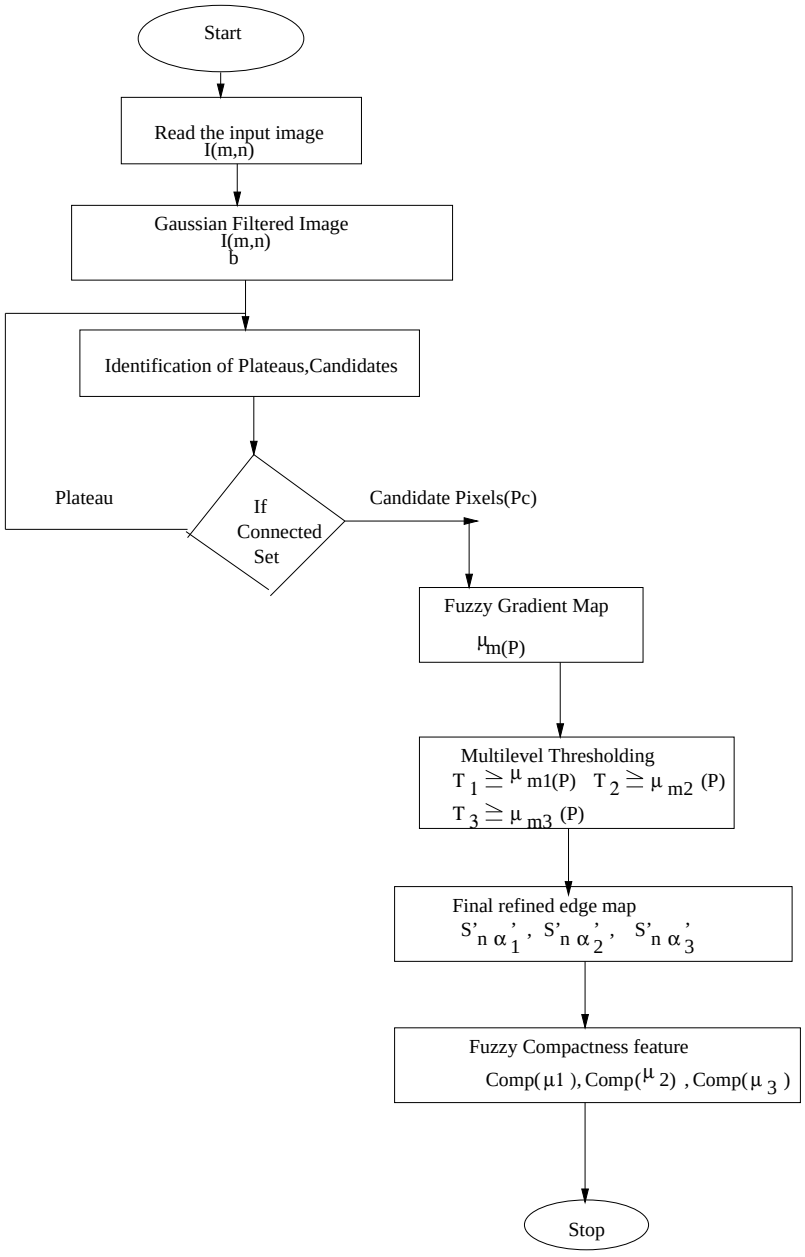


Figure 2.3: Flow chart for proposed feature extraction.

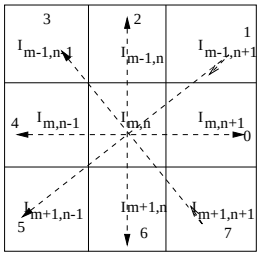


Figure 2.4: 3×3 neighborhood of the candidate pixel.

Performance evaluation

To evaluate the performance of retrieval the standard retrieval benchmarks such as **Recall rate** and **Precision rate** [30] are considered. Let n_1 be the number of images retrieved in top 20 positions that are close to the query. Let n_2 represent the number of images in the data base similar to the query. Evaluation standards **Recall rate** , **Precision rate** are defined as follows :

$$Recall \quad rate = \frac{n_1}{n_2} \times 100\% \tag{2.16}$$

$$Precision \quad rate = \frac{n_1}{20} \times 100\% \tag{2.17}$$

2.3.3 Building block

Different steps involved in generating the features are explained in the following paragraphs.

The intensity image matrix $I(m,n)$ is convolved with Gaussian function and resulted into to a blurred image matrix $I_b(m,n)$.

$$I_b(m,n) = I(m,n) * G(x,y) \tag{2.18}$$

where $G(x,y) = \frac{1}{\sqrt{2}\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$ where σ effectively determines the window size of the

filter which in turn determines the degree of smoothing. The blurred image preserves the original shape. By Gaussian filtering, a part of the intensity is distributed to the neighbors while preserving the edge information. The border region which characterizes the transition between two uniform Plateaus can be determined from the blurred image.

The basic steps include, finding the transition between the Plateau regions consisting of the possible edge candidates P_c which acts as an input for finding the fuzzy gradient map. The points of this region are characterized by $\mu_m(P)$ representing the edge strength in the interval [0,1] shown in eqn. (2.19).

$$s_n = \{(\mu_m, P_c)\} \tag{2.19}$$

The gradient map $\mu_m(P)$ is thresholded at different levels using ($\alpha - cut$) [155] with membership values greater from 0.5 onwards. Let the candidates of the multilevel gradient map be represented by the equation given below.

$$s_{n\alpha} = \{(P \in s_n : \mu_m(P) \geq \alpha\} \tag{2.20}$$

where $0.5 \leq \alpha \leq 1$.

Now $s_{n\alpha}$ is taken as input to find out the refined edge map using the notion of highest connectivity which resulted in a subset $s\prime_{n\alpha}$ as shown in eqn.2.21. Points which are preserving highest connectivity within a defined window are extracted using the membership function $\mu_d(P)$ shown in eqn. 2.14.

$$s\prime_{n\alpha} = \{(x \in s_{n\alpha} : \mu_d(P) \geq \alpha\prime\} \tag{2.21}$$

where, α' is $\simeq 1$ to change the degree of connectivity between the adjacent pixels to the highest value.(proposition of eqn.2.6).

Fuzzy compactness feature comp(μ) is computed on $s\prime_{n\alpha}$ using eqn.2.7.

2.4 Algorithm and Implementation details

Algorithm

Step1 : Gaussian filtering is performed on the image.

Step2 : Each pixel is labeled as (a) Plateau Top or maxima or (b) Plateau Bottom or minima or (c) Candidate pel

Step3 : If the labeled pixel is a candidate pel, its gradient membership value $\mu_m(P)$ and its edge direction is estimated. The points thresholded above different membership values ($\mu_m(P) \geq 0.5$) are selected.

Step4 : Candidate pixels similar in edge magnitude and direction are linked.

Step5 : Fuzzy compactness feature vector is computed from the edge map.

Step6 : Euclidean distance is computed between the feature vector elements of the query image and images in the database and ranked according to distance. The results of fuzzy edge map extraction are shown between Figs. 2.9- Figs. 2.12.

Feature vector elements comprising of $[comp(\mu_1), comp(\mu_2), comp(\mu_3)]$ are computed using eqn.2.7.

Performance using different set of features, has been studied by transforming the values of (k_m) and varying the window size, to estimate k_d, as discussed in subsection 2.3.1.

Features

The schematic diagram for feature extraction is shown in Fig. 2.3. The features are computed as follows.

Feature set (A) (for k_m = 1.0039 and 3×3 window size is used to estimate k_d.)

$comp(\mu_1)$ is computed from the edge map by thresholding at $\mu_m(P) \geq 0.6$.

$comp(\mu_2)$ is computed from the edge map obtained by thresholding at $\mu_m(P) \geq 0.7$.

$comp(\mu_3)$ is computed from the edge map obtained by thresholding at $\mu_m(P) \geq 0.8$.

This set of features are computed using the value of $k_m \simeq 1.0039$ as shown in eqn. 2.10 and value of k_d is determined from eqn. 2.14 over 3×3 window size.

using feature set A, the gradient maps of Fig. 2.5(a), (b) are shown in Fig. 2.6(a), (b). The refined edge maps of Fig. 2.5(a), (b) are shown in Fig. 2.7. The refined edge maps of

the deer and bird image, of Fig.2.8 (a), (b) are shown in Figs. 2.9(a), (b).

Feature set (B) : This set of features are computed as follows :

$comp(\mu_1)$ is computed from the enhanced map thresholded at $\mu_m(P) \geq 0.8$

$comp(\mu_2)$ is computed from the enhanced map thresholded at $\mu_m(P) \geq 0.9$.

Using the value of $k_m \simeq 1.18$ and value of k_d as determined from eqn.2.14 over 5×5 window size, the feature elements $comp(\mu_1)$ and $comp(\mu_2)$ are computed.

The edge map of Fig. 2.8(a) as obtained using feature set (A) is shown in Fig. 2.12(a). That using feature set (B) for k_d estimated over 3×3 window is shown in Fig. 2.12(b) and that for k_d estimated over 5×5 window is shown in Fig. 2.12(c).

$comp(\mu_3)$ is computed from the edge map obtained by thresholding at $\mu_m(P) \geq 0.7$ with $k_m \simeq 1.18$ and k_d is determined in 3×3 window. In this case, only those pixels are selected as candidates which are having other pixels in the direct neighborhood with gradient membership $\mu_m(P)$ greater than 0.7. The edge candidates of Figs. 2.8(b) and (c) are shown in Fig. 2.13. The candidate edge points as obtained for Fig. 2.5(a), are shown in Fig. 2.14.

Although, three thresholding levels of membership have been selected, but variation in selection of thresholding levels will not change the result drastically. Low, medium, high types of edge strength are grouped in three distinct categories. Value of k_m and k_d can be tuned to select different candidates based on their membership values.

The retrieval performance was checked by altering the value of k_m and k_d, resulting into feature sets (A) and (B). The motivation behind changing the values of k_m is to alter the property of the membership plane. This is required to emphasize the importance of the different type of categorized edge points (strong, medium and weak). Increasing the value of k_m in effect, transforms the membership values as shown in Fig. 2.2(b), where the values of the strong or medium edge pixels are increased and well separated. The weak edge pixels are less increased in membership values. However, the upper limit of values of $\mu_m(P)$ are always kept less than 1.0. Similarly the values k_d varies with different window size and results in different edge set at the same threshold of $\mu_m(P)$. The effect

on performance has been analyzed in section 2.5.

Computational complexity : The analysis of the computational complexity (worst case) involved in different operations for an image of size $M \times M$ and with window neighborhood $N \times N$, is explained as follows: 1. Identification of border regions requires M^2N^2 operations. 2. Assignment of $\mu_c(P)$ involves $\lambda\, M^2N^2$ operations. 3. Assignment of $\mu_d(P)$ involves λM^2N^2 operations. 4. Thresholding involves λM^2 operations. 5. The total operation is of the order of O(M^2N^2).

2.5 Experimental results

Extensive experiments have been performed using two databases of which the first database consists of 120 gray level images (256×256) of different objects down loaded mainly from Google engine (www.google.com), OlivettiFace database
($http\ :\ //ftp.uk.research.att.com\ :\ pub/data/att\_faces.tar.Z$) and some synthetically generated images. The images in this database are mainly of single object both with textured and non textured background. This collection consists of dissimilar group of images, of which each group is having some similarity within itself. Such database consisting of both visually distant set of images as well as closely related images are useful in studying the system performance properly [147].

The second database consists of the trade mark and logo images of United States Patent and Trademarks office (USPTO) consisting of 1053 binary images. The images are converted to gray level and reduced to (256×256) for the experiment. The size of the Gaussian smoothing filter is fixed to 3x3 pixels and value of σ to 1.5, used to limit the degree of smoothing across the edge. The average cputime time required for computing the feature vector (elements comprising of the fuzzy compactness value at three different thresholds) is 95.60 secs.The experiment is performed in (SUN microsystems Ultra 60) system using MATLAB.

The experimental results are shown from Figs. 2.6 to Figs. 2.22. Fig. 2.5(a) is a

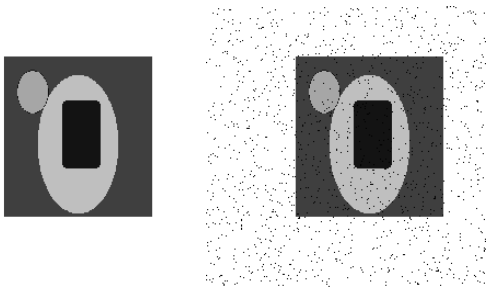
synthetically generated image and its corresponding multilevel gradient map as obtained from eqn. 2.10 is shown in Fig. 2.6(a). The noisy version of the image is shown in Fig. 2.5(b). The corresponding gradient map is shown in Fig. 2.6(b). Figs. 2.7(a-c) show the refined edge map of Fig. 2.5(a) at different threshold with ($\mu_m(P) \geq 0.6, 0.7, 0.8$). Fig. 2.7(c) is the refined edge map of the noisy image of Fig. 2.5(b). From Figs. 2.6(b) and Fig. 2.7(c) it is seen that the proposed technique has considerably good noise tolerance. The aim of this work, is to select significant candidates from the multilevel edge map, image containing meaningful shape and gray level information. The refined edge maps of the bird and deer image at $\mu_m(P) \geq 0.6$ is shown in Fig. 2.9 and for other thresholds in Fig. 2.10 and Fig. 2.11. As seen from Fig. 2.9(a), that beside the edge points of the object, lot of edge points are generated from the background, which may result in a misleading characterization of the object by edge map.

To reduce the effect of the background to some extent, a membership transformation scheme is used with varying the value of k_m of eqn. 2.10. This makes nonlinear intensification of the edge membership value depending upon the value of k_m. Two set of experiments have been performed to calculate feature values, feature set (A), feature set (B) described in Sec. 2.4, by varying value of k_m and size of support window for determining k_d. Fig. 2.12(a) is obtained at $\mu_m(P) \geq 0.8$ using eqn. 2.10 and k_d evaluated from eqn. 2.14 with feature set (A). Fig. 2.12(b) is obtained by thresholding at $\mu_m(P) \geq 0.8$ by choosing $k_m \simeq 1.18$ and k_d evaluated from eqn. 2.14. Fig. 2.12(c) is obtained by thresholding at $\mu_m(P) \geq 0.8$ by choosing $k_m \simeq 1.18$ and k_d evaluated from eqn. 2.14 over 5×5 window size with feature set (B). It is evident from Fig. 2.12(a), that at higher threshold much of the significant edge points are lost. Better result is obtained from Fig. 2.12(b) which can be explained from the following considerations. By such intensification of the pixels, higher membership values are more separated resulting into better edge information while rejecting some weaker edge pixels resulted from weak textured background. This result is further improved when larger size of the support window is used as shown in Fig. 2.12(c). From the proposed algorithm, it is observed from Fig. 2.6(a)

that points of the curved regions are of lower membership value compared to the pixels belonging to the straight line. This property is used to extract points from curved regions of the edge map. In order to extract significant edge points from high curvature as candidates, those pixels which are connected to stronger pixels i.e., pixels of higher gradient membership $\mu_m(P)$ value over 3×3 window, are selected. Such significant points of the bird and elephant image of Fig. 2.8 are also shown in Fig. 2.13. Such selected pixels of Fig. 2.5(a) are shown in Fig. 2.14(a) and Fig.2.14(b). In case of Fig.2.14(b) some more pixels from sharp curvature are selected, by pushing down the threshold $\mu_d(P)$ to 0.94, than in Fig. 2.14(a). Fig. 2.15(a) shows the original image (bull image from trade mark database). The edge map of the rotated bull image using feature set (B) is shown in Fig. 2.15(b), (c) respectively.

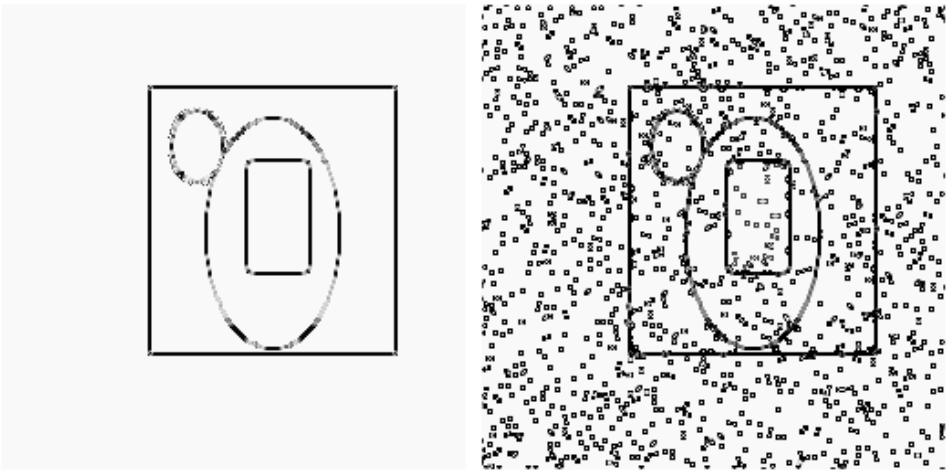
Feature values

The feature values (fuzzy compactness) of Fig. 2.16 are computed at different thresholds to study the invariance property of the feature which is used for characterizing an the image, in the proposed method. The values are shown in Table 2.1. It is observed that the compactness values obtained for the translated and rotated version (90, 180, 270 degrees) give satisfactory value. However, the value differs by some amount for angle rotated at (30 degree). This can be explained from the fact that the popular image rotation algorithms suffer from digitization error due to interpolation, thereby influencing the feature values as seen from Fig. 2.15(c). The feature value of scaled version is also changed by some amount due the reason mentioned above.



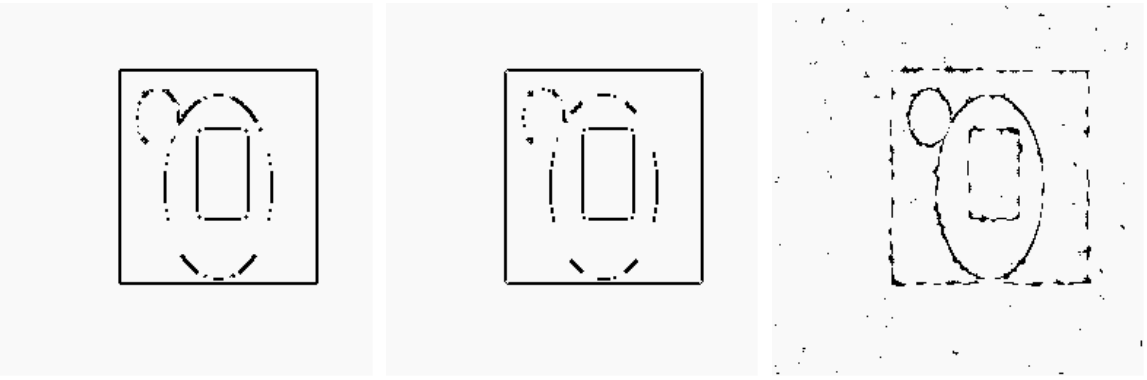
(a) (b)

Figure 2.5: (a) Original image. (b) Noise introduced.



(a) (b)

Figure 2.6: Multilevel gradient map (without linking), (a) of Fig. 2.5(a) (b) for noisy image of Fig. 2.5(b)

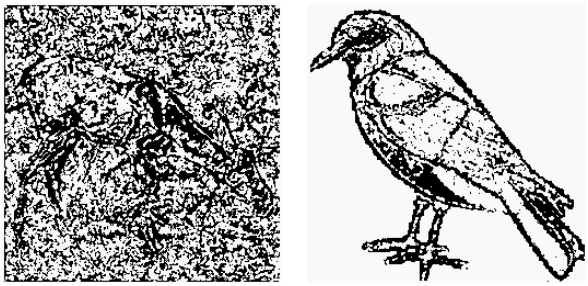


(a) (b) (c)

Figure 2.7: Edge map obtained using set (A) at $\mu_m(P) \geq$(a) 0.7 (b) 0.8 (c) 0.6 (noisy image) and $\mu_d(P) \simeq 1.0$ for all the three cases. Candidates plotted as crisp edge points.

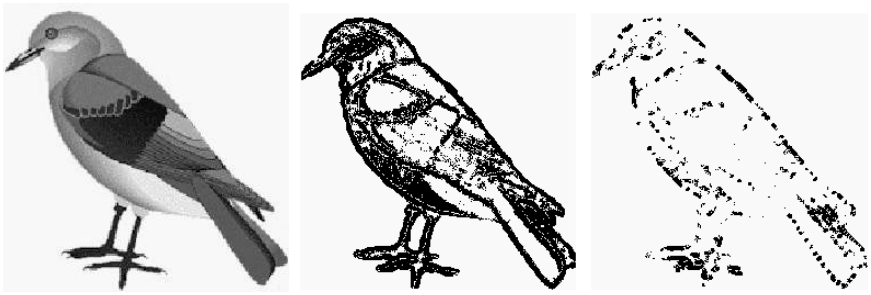
(a) (b) (c)

Figure 2.8: Original images. (a) deer image in textured background. (b) bird image (c) elephant image.



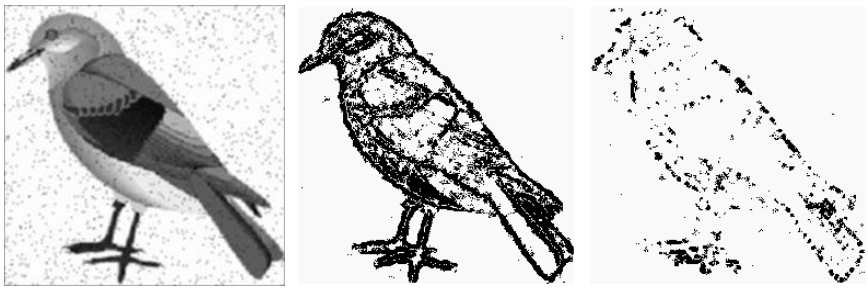
(a) (b)

Figure 2.9: Edge map, candidates with $\mu_m(P) \geq 0.6$ and $\mu_d(P) \simeq 1.0$ are plotted as crisp edge points. (a) deer image (b) bird image. Value of k_m= 1.0039 and k_d estimated over 3×3 window, set (A).



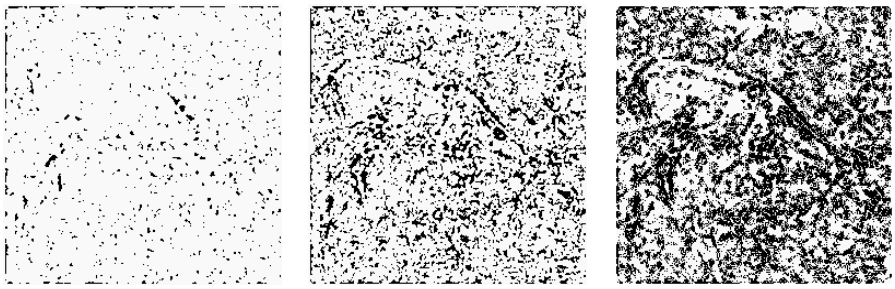
(a) (b) (c)

Figure 2.10: (a) Original image. Edge map obtained at $\mu_m(P) \geq$ (b) 0.7 (c) 0.9 and $\mu_d(P) \simeq 1.0$. Value of k_m= 1.18 and k_d estimated over 5×5 window, set (B). Candidates plotted as crisp edge point.



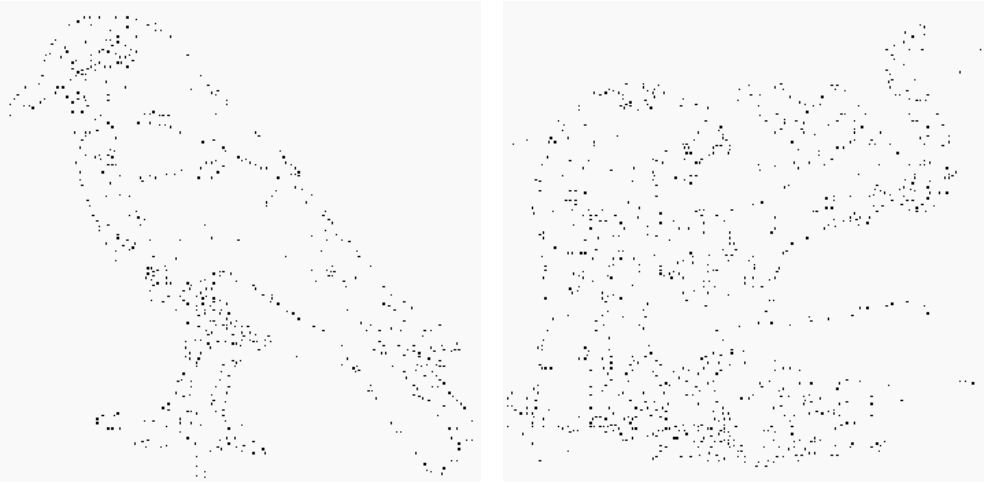
(a) (b) (c)

Figure 2.11: (a) Original image (noisy background). Edge map obtained at $\mu_m(P) \geq$ (b) 0.7 (c) 0.9 and $\mu_d(P) \simeq 1.0$. Value of k_m= 1.18 and k_d estimated over 5×5 window, set(B). Candidates plotted as crisp edge point.



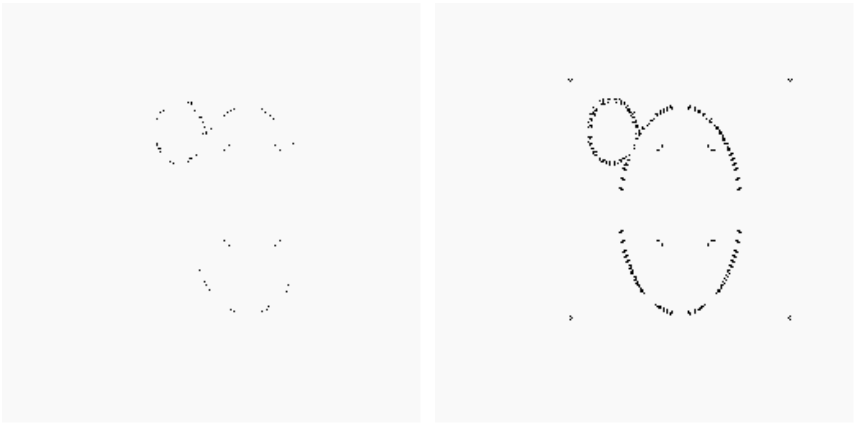
(a) (b) (c)

Figure 2.12: (a) Edge map, candidates with $\mu_m(P) \geq 0.8$ and $\mu_d(P) \simeq 1.0$ plotted as crisp edge points (a) with value of k_d estimated over 3×3 window, set (A). (b) Value of $k_m = 1.18$ and k_d estimated over 3×3 window. (c) $k_m = 1.18$ and k_d estimated over 5×5 window, set (B).



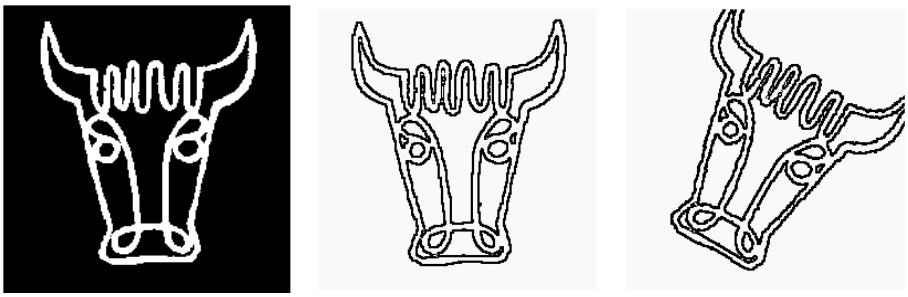
(a) (b)

Figure 2.13: Edge candidates with $\mu_m(P) \geq 0.7$ connected to stronger pixels in 3×3 window with $k_m = 1.18$ and k_d evaluated over 3×3 window. (a) bird image (b) elephant image. Candidates with $\mu_m(P) \geq 0.7$ and $\mu_d(P) \simeq 1.0$ are plotted as crisp edge points.



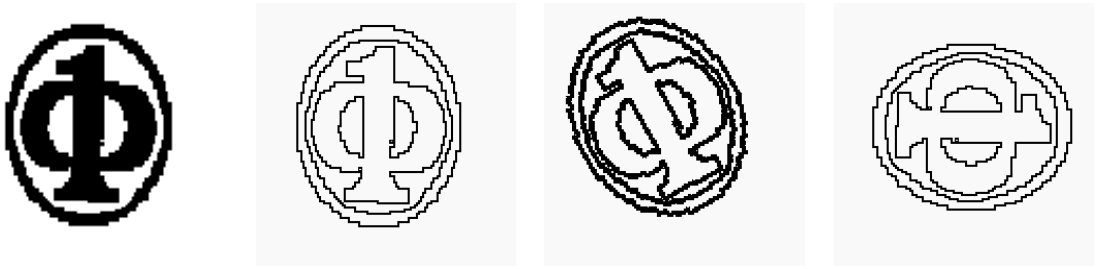
(a) (b)

Figure 2.14: Edge candidates of Fig. 2.5(a) with $\mu_m(P) \geq 0.7$ connected to stronger pixels in 3×3 window with $k_m = 1.18$ and k_d evaluated over 3×3 window. (a) $\mu_m(P) \geq 0.7$ and $\mu_d(P) \geq 0.98$ (b) $\mu_m(P) \geq 0.7$ and $\mu_d(P) \geq 0.94$. Candidates plotted as crisp edge points.



(a) (b) (c)

Figure 2.15: (a) Original image. (b) edge map (c) edge map of rotated image. Edge candidates with $\mu_m(P) \geq 0.7$ with $k_m = 1.18$ and k_d evaluated over 3×3 window. Candidates with $\mu_m(P) \geq 0.7$ and $\mu_d(P) \simeq 1.0$ are plotted as crisp edge points.



(a) (b) (c) (d)

Figure 2.16: (a) Original image (b) edge map (c) edge map of (30deg) rotated image (d) edge map of (90deg) rotated image. Edge candidates with $\mu_m(P) \geq 0.7$, with $k_m = 1.18$ and k_d evaluated over 3×3 window. Candidates with $\mu_m(P) \geq 0.7$ and $\mu_d(P) \simeq 1.0$ are plotted as crisp edge points.

Table 2.1: Feature values

Image	$comp(\mu_1)$	$comp(\mu_2)$	$comp(\mu_3)$
Fig.2.16 (0 deg)	0.000184	0.000184	0.000129
Fig.2.16(translated)	0.000184	0.000184	0.000129
Fig.2.16 (scaled)	0.000140	0.000140	0.000078
Fig.2.16(90 deg)	0.000184	0.000184	0.000131
Fig.2.16(180 deg)	0.000184	0.000184	0.000131
Fig.2.16(30 deg)	0.000150	0.000134	0.000134

2.5.1 Retrieval results

The image retrieval experiment is performed by computing features separately using feature set (A) and feature set (B). A particular image is selected as query and based on the Euclidean distance computed using eqn. 2.15 the images are ranked according to the increasing order of distance. Top twenty images are selected. These results are shown in Figs. 2.17 to Figs. 2.22. Here, the images are ranked left to right, starting from the top in increasing distance.

Retrieval of gray level images (database 1)

The experimental results are shown in Figs. 2.17, 2.18 and 2.20. The usefulness of the algorithm is demonstrated, in the cases of finding images those appear visually similar to a query. Several examples are demonstrated for finding similar scenes, similar faces or objects from the database.

Fig. 2.17(a) and Fig. 2.17(b) show the experimental results when queried with the girl image. It should be noted that the system presented here works well for faces using the same representation parameters which has been used for other cases. Better precision is obtained with feature set (B) in which the strong edge pixels are more separated than the weak pixels by nonlinear intensification of the pixels.

Fig. 2.18(a) and Fig. 2.18(b) show the experimental results when queried with the bird image. The rotated, the noisy images, are retrieved at a higher rank and some background variation is tolerated also. Better precision is obtained with feature set (B) in this case.

Figs. 2.20(a) shows the results of the first twenty retrieved images with the query image on the top left corner. It is seen from Fig. 2.20(a) that the rotated and scaled version of the image, the noisy version of the image are the candidates retrieved at a higher rank.

Fig. 2.20(b) shows the result when queried with a tiger in a textured background.The global similarity measure (fuzzy compactness) works well here. The object(tiger) in this case constitutes a significant portion of the image. By non linear intensification of the candidate pixels the effect of weak textured background is minimized to some extent. A good precision is obtained in this case. Fig. 2.19(a) and (b) present the retrieval result for the same query of bird and girl image using Fourier descriptor of Canny [40] edge detector. It is seen from the results, that the proposed method perform better in both the cases. It is seen from Figs. 2.17, 2.18 and 2.20, that there are few other dissimilar images retrieved at an higher rank. Such results can be explained considering the fact that a global measure (fuzzy compactness) has been used for computing similarity. As, fuzzy compactness is computed from the edge map irrespective whether the pixel belongs to the object or background, the resultant compactness value can have contribution both from the object and background. In case, when the background is textured, contribution from the background is more in comparison to the case where the background is uniform. This results in a higher degree of error in computing the compactness of the object of interest.

Retrieval of trademark images (database2)

Fig. 2.21 and Fig. 2.22 show the performance of our algorithm on the trade mark images. It is difficult to judge the relevance in case of trade mark images because it is difficult to judge visually similar trade marks. In case of trade mark images, color or gray level contrast information does not play a useful role in distinguishing between various marks. The searching of the trade mark image is based mostly on the decision of the

shape information present in the binary images.

Fig. 2.21(a) shows the retrieval performance when queried with the image on the top left corner. Most of the retrieved images have almost the same structure. The first five successive images depicts the rotated version of the query image. It is seen that the shape characteristics of the image is adequately captured by our approach in this case.

Fig. 2.21(b) shows the performance with the query image on the top left corner. The translated and the rotated versions of the image are retrieved at a rank within the first twenty retrieved images.

Fig. 2.21(c) shows the performance with the query image of a bull's head. The images rotated at different angles are retrieved within the first twenty ranked images. However, there are few other images retrieved at an higher rank.

Fig. 2.22 shows the performance with the query image on the top left corner. The image rotated at different angles are retrieved at a rank within the first forty retrieved images. However, certain images differing in shape with the query image but having approximately close feature value are also retrieved at a higher rank.

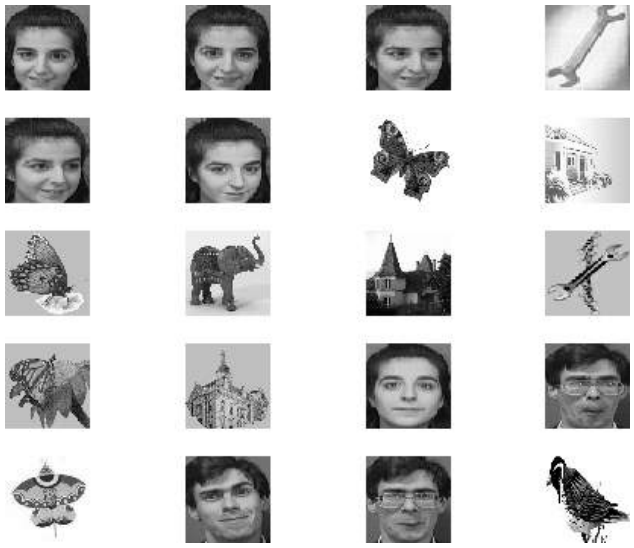
It is observed from our experimental result of Fig. 2.21 and Fig. 2.22 that some dissimilar images attain higher rank and come between similar images. This can be explained from the fact that the refined edge map, which is obtained by taking the degree of linking, is based on the turning edge angles of the object boundary. Such shape features gives accurate results with objects having one closed boundary. Practical images have many fine details and turning angle features. This may reduce the precision to some extent.

2.5.2 Performance measure

The performance measure of retrieval is judged from the recall rate and precision rate obtained from the first twenty successively retrieved images. Precision can be made high by retrieving only few images and recall can be made equal to 1 simply by retrieving all images [147]. The Recall and Precision rates as computed using eqns. 2.16 and 2.17 for

different queries are shown in Table. 2.2

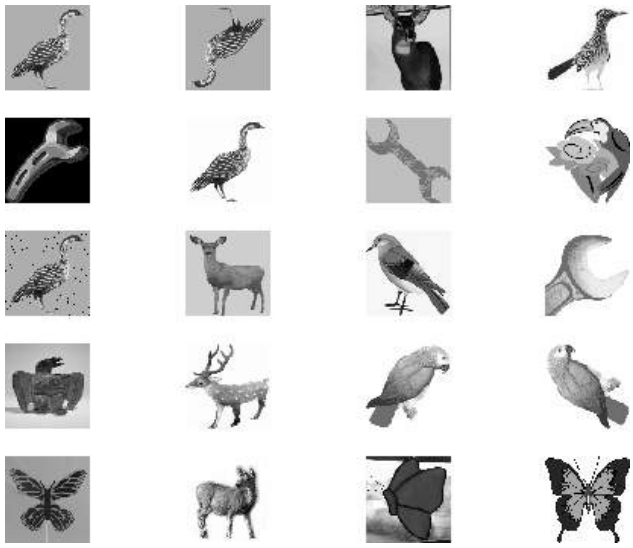
We compare the proposed method with the retrieval results of the Fourier descriptor method. Fourier descriptor is one of the commonly used boundary based shape representative. Fourier descriptors are chosen, for benchmarking the results because Fourier transformed edge features are generally invariant to geometric transformation and also noise resilient. The recall and precision rates for both the methods are shown in Fig.2.23. In the case of computing Fourier descriptors, Canny edge detection [40] is performed as a preprocessing step for computation of the features. It is seen the proposed method offers better results. The results after querying with a bird and girl are shown in Fig. 2.19. While making a comparison between the methods, it may be mentioned that, not all the selected edge points bear structurally significant information about different objects within the image. It can be claimed that, by categorizing the different type of edge points as (strong, medium and weak) and subsequent thresholding, it has been possible to generate better description of the image at multiple levels. As local contrast information is embedded in the feature values, the scheme provides some distinguishability in the features which enhances the retrieval accuracy.



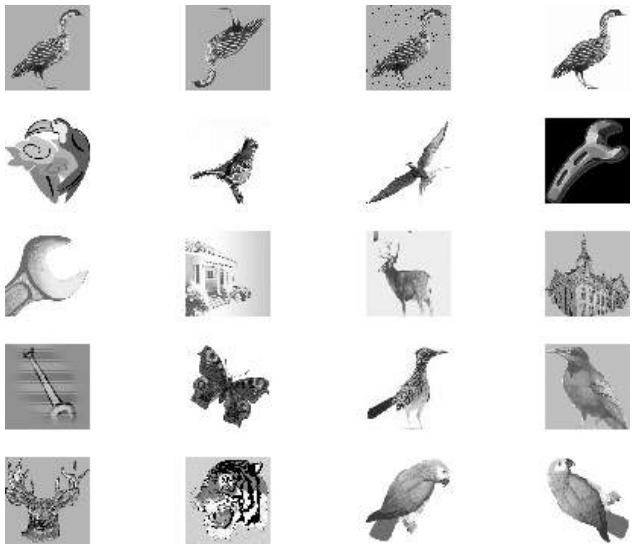
(a)

(b)

Figure 2.17: Retrieved results, with top left image as the query image. (a) using feature set (A) (b) using feature set (B).

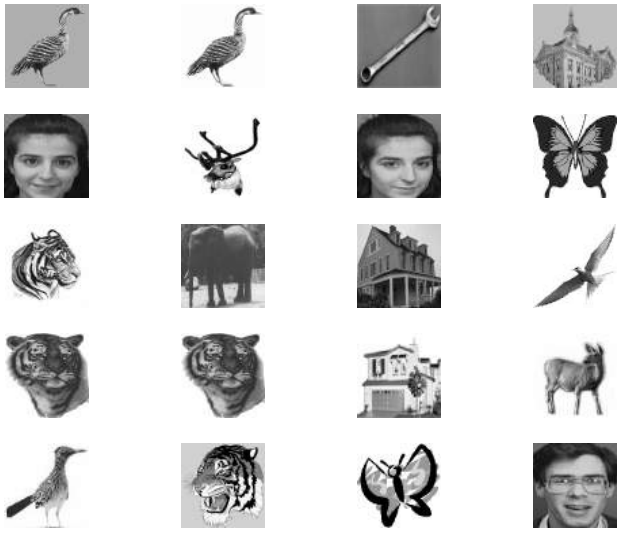


(a)

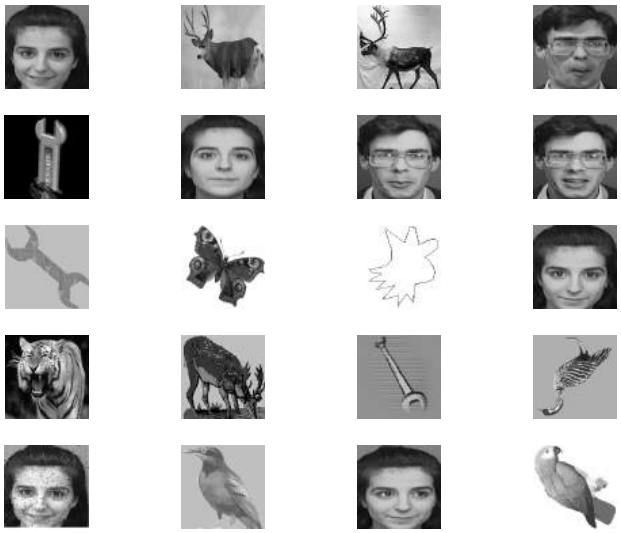


(b)

Figure 2.18: Retrieved results, with top left image as the query image. (a) using feature set (A) (b) using feature set (B).

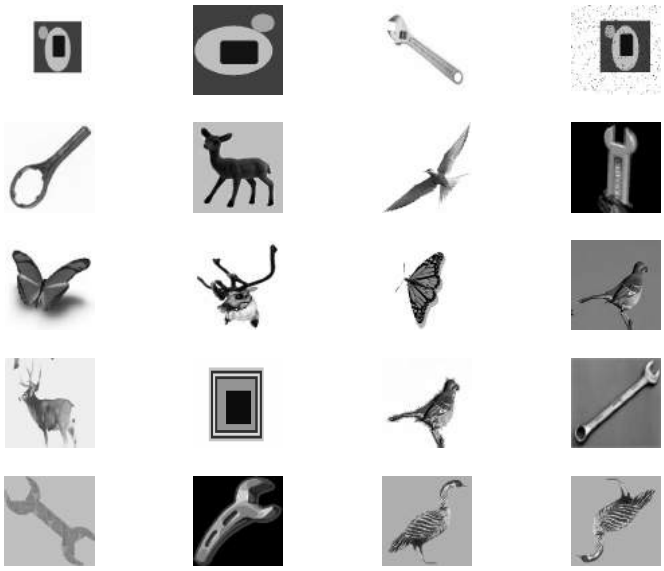


(a)



(b)

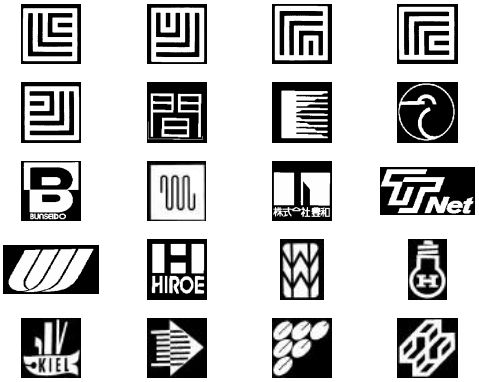
Figure 2.19: Retrieved results, using Fourier descriptor (a) bird image (b) girl image.



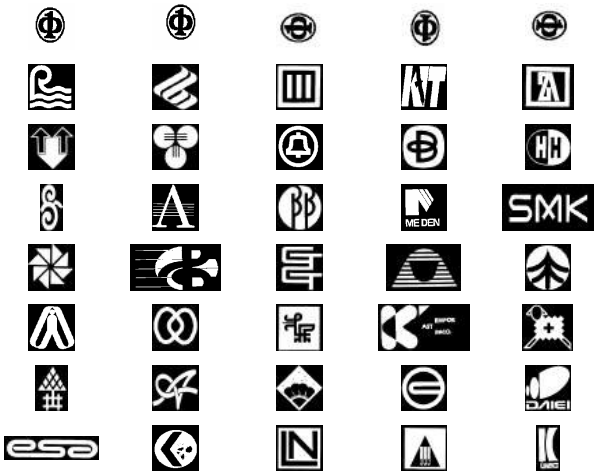
(a)

(b)

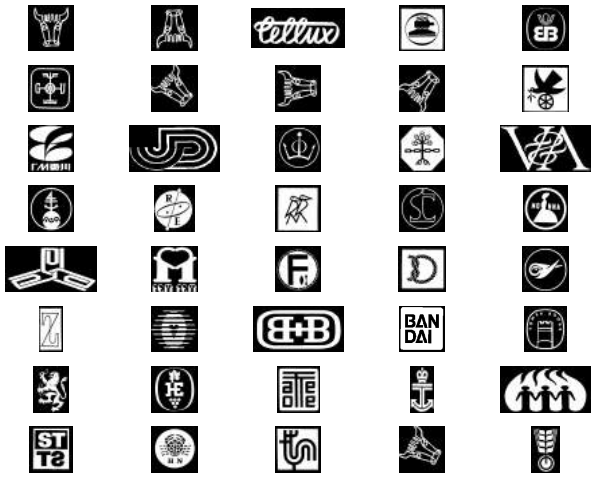
Figure 2.20: Retrieved results, with top left image as the query image. (a) using feature set (A) (b) using feature set (B).



(a)



(b)



(c)

Figure 2.21: Retrieved results, using feature set (B), from (USPTO) database. The top left image is the query image.

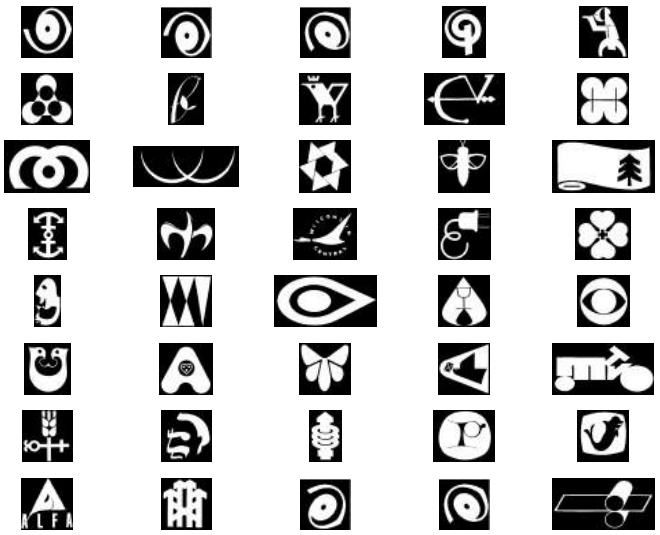
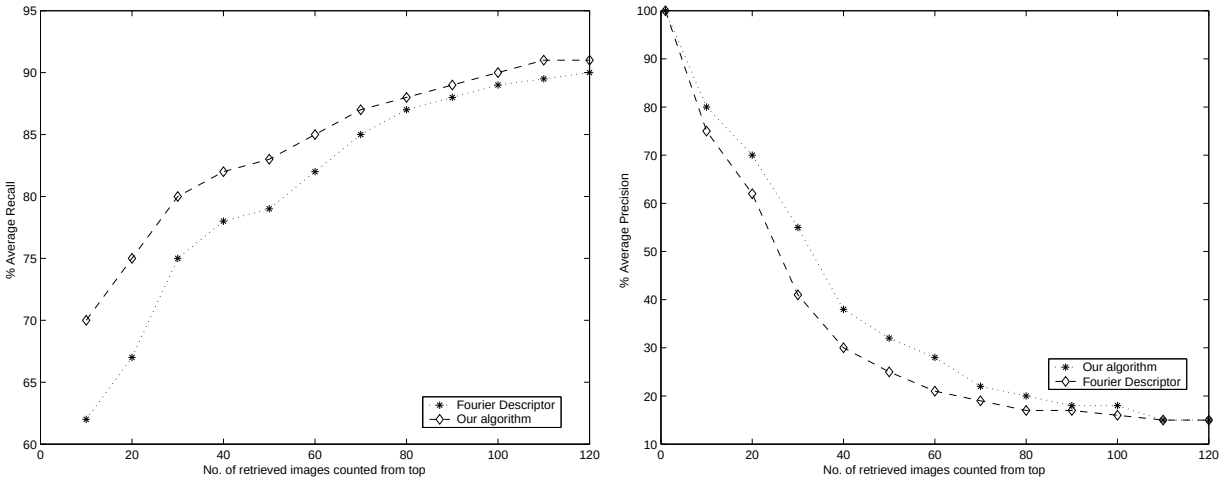


Figure 2.22: Retrieved results, using feature set (B) from (USPTO) database. The top left image is the query image.

Table 2.2: Performance evaluation %

Image	$n_1(A)$	$n_1(B)$	n_2	Rec.rate(A)	Pr.rate(A)	Rec.rate(B)	Pr.rate(B)
Fig.2.17	6	9	10	60	30	90	45
Fig.2.18	8	11	11	72	40	100	55
Fig.2.20(b)		9	12			75	45
Fig.2.20(a)	4		4	100	20		
Fig.2.21(a)		9	12			73	45
Fig.2.21(b)		8	10			80	40
Fig.2.21(c)		7	10			70	35
Fig.2.22		8	13			56	40



(a) (b)

Figure 2.23: Comparative results, between proposed method using feature set (B) and the Fourier descriptor method.

2.6 Conclusion

In this chapter, an algorithm for extraction of fuzzy edge map has been proposed. Thereafter, using the edge maps, an image retrieval system has been designed. The extracted edge points are categorized as strong, medium and weak type of points, based on the uncertainties in gradient strength, mapped by the gradient membership function $\mu_m(P)$. These points have significant advantage of embedding both gradient direction and local contrast information. The Topological feature (fuzzy compactness) computed from the edge maps obtained at different threshold, constitute the feature vector which is invariant to rotation, translation and scale by definition. Extensive experiments have been performed on both logo and general purpose images. It is evident from the experimental results, that fuzzy compactness provides an fairly good tool for indexing and can be able to retrieve similar images from a heterogeneous database whose classes are not known a priori. Such an approach, can be taken as an alternative to segment an image into coherent regions, for subsequent image to image matching applications.

The average cputime time required for computing the feature vector, (elements compris-

ing of the fuzzy compactness value at three different thresholds) is 95.60 secs. However, this time is dependent on nature of the image type. The complexity involved in computation of features is of O(M^2N^2). This includes the processing time, used to generate the refined edge map. The future directions in modifying the algorithm could be to reduce the computational time and improve the accuracy. As seen from the results, that the performance of the proposed method is best achieved when queried with single objects and having non textured background. The method could be further improved if some other features are used in association with the present feature. This may lead to better application in object recognition.

The present chapter discusses to what extent the information in an image is captured in the edges. However, the points where image edges have their maximum curvature are very important in characterizing the overall structure, with reduced number of points. These points are commonly defined as corner points and play a significant role in perceptual grouping, which leads to recognition of shape of an object. Corner detection in binary images is a simple process [148]. But in the case of gray level images locating corner position is quite uncertain due to spatial and gray level ambiguities. This arises due to imperfection of imaging systems as well as lighting conditions. So, a robust corner detection mechanism which can handle this sort of uncertainties is necessary. To achieve such goal, the problem of corner detection, using a fuzzy set theoretic approach has been addressed in the next chapter.

Chapter 3

Extraction of Gray Level corners : Perceptually Significant Points of Interest

3.1 Introduction

Visual similarity between images can be achieved through correspondence between the extracted features like color, texture, shape, etc. [123]. Similarity between two images can be evaluated by comparing shapes of different objects within the image. In such cases, high curvature points play a key role in matching between two images. In this chapter, the problem of extracting perceptually significant features, which are mainly the high curvature points including the corner points has been addressed. These extracted points may be considered as an important source of bearing structural information of different objects within an image. The high curvature points may be considered as useful candidates for characterizing an image. With the extracted fuzzy edge map, as explained in Chapter2, the process of perceptual grouping has been used to extract significant portions (high curvature points) from the spatial structure.

In the previous chapter, it has been shown that edge based features seem to be an

important source of data for use in image retrieval. Although, an image is quite well represented in terms of overall edge information and combination of local features, but the robustness of the extracted features will depend upon the degree, which it is close to human perceptual grouping. The extracted features should have some adaptability towards variations like, changes in imaging conditions resulting from blurring, varying illumination, etc. Inspired from these ideas, a fuzzy set theoretic approach has been used, to extract the significant points.

The human visual system has a highly developed capability for detecting many classes of patterns including visually significant arrangements of image elements. From the psychovisual aspect, points having high curvature are one of the dominant classes of patterns, that play important role in almost all real life image analysis applications [123]. These points represent significant amount of shape information. Corners are generally formed at the junction of different edge segments which may be the meeting (or criss crossing) of two edges. Cornerness of an edge segment depends solely on the curvature formed at the meeting point of two line segments. Since corners depict significant structural information in an image, reliable detection of corners remain an important research topic in image analysis.

Real life image data are always imprecise due to inherent uncertainties that may arise from the imaging process (such as, defocusing, noise, wide variations of illuminations, etc). As a result, precise localization and detection of corners became difficult under such imperfect situations. Fuzzy set theoretic approach, [216] which can be used for possible precise measurement with imprecise data [157] has been adapted for developing the corner detection algorithm.

Dominant curvature points can be represented as entities of a fuzzy set where exact corner point is also a member. These points may be considered as interesting points for representing an image. These points will be referred as fuzzy corner points in this contribution.

A great deal of work has been done on corner detection, in the area of image processing during the past many years. This issue is discussed in brief.

Corner detection on gray level images can be classified into two main approaches. In the first approach, the gray level image is first converted into its binary version for extraction of boundaries using some thresholding techniques. After extraction of boundaries, the corners or the high curvature points are detected using directional codes or other polygonal approximation techniques [65].

In the second approach, the gray level image is taken as input directly for corner detection. This discussion will be restricted to the second approach only. Most of the general purpose gray level corner detectors [98], [222], [170], [199], [174], [13], [1] use either topology based or auto-correlation based approach. Topology based corner detectors, mainly use gradients and surface curvature to define the measure of cornerness. Points are marked as corners, if the value of cornerness exceeds some predefined threshold condition. Alternatively a measure of curvature can be obtained using auto-correlation.

Moravec [146] developed a corner detector, in which corners are defined when there are large intensity variations in every direction. This idea is modified in Harris corner detector [84] considering the local intensity change in terms of directional derivative of surface features. It is also considered as one of the best corner detectors on gray level images. Smith and Brady proposed [190] a new method called " Smallest Univalue Segment Assimilating Nucleus (SUSAN)", which utilize simple masking operations in place of gradient convolution. They used a circular window, to obtain USAN area, and centroid is extracted. This method has better noise robustness and localization of corners compared to other methods.

Corner is an important feature used in various image analysis applications [121], [61]. Detection of corners, using classification based approach [17], [112] and fuzzy reasoning based method [115], [111] both for color images [210] and gray level images is an important research issue.

In this study, an algorithm to extract significant gray level corner points based on fuzzy set theoretic approach has been proposed. The high curvature points located at the discontinuities between different uniform intensity surfaces, constitute the fuzzy corner set. The

measure of cornerness varies with fuzzy edge strength and gradient direction. Different set of fuzzy corners are obtained using different values of threshold on the fuzzy edge map. The uncertainties in locating the corners points which may arise, due to discretization, noise and other imaging defects, are handled with fuzzy model. The robustness of the proposed algorithm is experimentally verified using both simulated image data and natural images, to justify the suitability of the algorithm. The performance has also been compared with the popular conventional detectors of [84] and [190] to illustrate the efficiency of the algorithm.

Considering the fuzzy edge map as input, a study has also been made, using a classification based approach, with Support vector machines(SVM) to extract the possible corners. The objective of which is to incorporate some learning mechanism in order to increase the robustness of the the feature extraction algorithm.

The remaining chapter is organized as follows : Section 3.2 briefly describes the mathematical model used in this work. Section 3.3 describes the fuzzy corner extraction process. Section 3.4 describes the experimental results. Section 3.5.1 explains the corner detection algorithm using Support Vector Machines (SVM). Section 3.6 provides a conclusion.

3.2 Mathematical modeling of gray level corners

The basic notion behind the proposed algorithm is that, a digital image can be thought of as 2-D plane, where there are ridges or valleys [207], [82] which are similar to Plateaus. This is true, when there are simply connected sequence of pixels having gray tone intensity values significantly higher (lower) in the sequence than the neighboring pixels. Desired features can therefore be obtained by extracting and assembling topographic characteristics of intensity surfaces, and has already been discussed in Chapter2.

Corner points are high curvature points and should lie on gray level edges. It should have significant change in edge direction with linear arm support of considerable length on

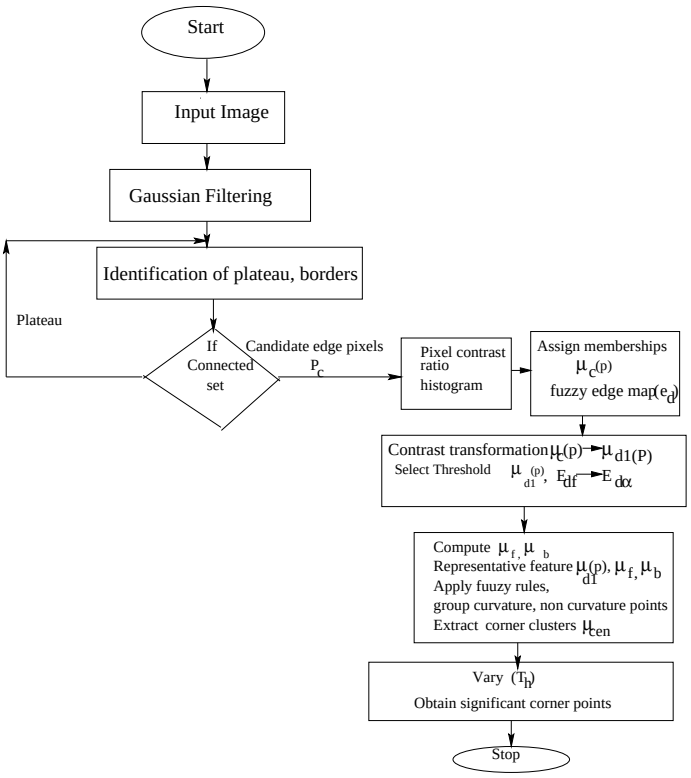


Figure 3.1: Block diagram of the proposed algorithm.

both sides. The building block of the proposed methodology is explained in Fig. 3.1

The proposed methodology for extraction of fuzzy edge map has already been explained in chapter 2. The fuzzy edge map as obtained, using an exponential function has been explained with different examples, as shown in Figs. 2.7- 2.16. However, the fuzzy edge map obtained using an exponential function, required further post processing to generate a refined edge map and make it noise tolerant. This has been done by using the properties of connectedness as shown in Figs. 2.7- 2.14.

To reduce the computation involved in generating the refined edge map, also at the same time extract significant candidate edge points for computation of corners, results using other membership functions have also been studied. Different experiments have been performed, which are summarized and shown in Table. 3.1. A Standard S-type function as shown in Fig. 2.1 has been chosen to represent the uncertainties of edge strength and the

location of true edge point. No further post-processing had been made to obtain a refined edge map.

In the present case, possible candidate edge pixels (P_c) are extracted from the border regions between the uniform intensity surfaces, as explained in the chapter2, which are defined in terms of Plateau Top and Bottom. The edge candidates (P_c), which belong to the border regions are assigned gradient membership, using the standard S-type function as shown in Fig. 2.2 and designated as $\mu_c(P)$ [19]. The value of $\mu_c(P)$ is based on their respective gradient strength. Similar to chapter2, Gaussian filtering, has also been chosen to perform effective smoothing of small distortions caused by noise and to obtain blur boundaries. The size of the Gaussian smoothing filter is fixed to 3x3 pixels and value of σ to 1.5.

A fuzzy edge set (e_d) comprising of gradient membership $\mu_c(P)$, of the edge points $P \in P_c$ is formed, as defined in (3.1)

In the next step, two membership functions ($\mu_f(P)$ and $\mu_b(P)$) are computed to estimate the fuzzy connectivity strength along a path, in the forward and backward direction with respect to the candidate pixel. The detailed implementation of the steps are described in the following subsections.

$$e_d = \{(\mu_c(P), P_c)\} \tag{3.1}$$

3.2.1 Estimation of gradient strength $\mu_c(P)$

For every edge pixel $P(p_i)$ where p_i is the gray value of pixel (P), a 3×3 window is considered as shown in Fig. 3.2. Although the method of computing the local contrast ratios is similar to that already described in Chapter2, but the parameters of the S-type functions are assigned to cover the total dynamic range, instead of using maximum ratio i.e., putting $d_{ac} = 255$ and $d_{al} = 0$ as in eqn.2.10.

In Fig. 3.2, the symbols represent the gray values at different neighborhoods of P similar to Fig. 2.4. The difference between (a_1, a_2), (c_1, c_2), (b_1, b_2), (d_1, d_2) are taken as

gray level differences in four different directions. The ratio of gray label changes (X_r) are computed from two mutually perpendicular set of pixel pairs within the neighborhood. Considering the mutually perpendicular pair ($a_1\ p_i\ a_2$), ($c_1\ p_i\ c_2$), the computed ratios are, $\frac{1+|a_1-a_2|}{1+|c_1-c_2|}$, $\frac{1+|c_1-c_2|}{1+|a_1-a_2|}$ [19]. Similarly, ($b_1\ p_i\ b_2$), ($d_1\ p_i\ d_2$) are considered. The four values of pixel contrast ratios (X_r) as obtained from the neighborhood of each candidate edge pixel are shown in (3.2).

$$X_r = [\frac{1+|a_1-a_2|}{1+|c_1-c_2|}, \frac{1+|c_1-c_2|}{1+|a_1-a_2|}, \frac{1+|b_1-b_2|}{1+|d_1-d_2|}, \frac{1+|d_1-d_2|}{1+|b_1-b_2|}] \tag{3.2}$$

As mentioned in chapter2, an edge pixel will have maximum gray level difference in a direction, perpendicular to its true edge direction (ϕ) [75]. The edge direction (ϕ) should point along the minimum difference direction [75]. The maximum pixel contrast ratio (X_mr),

$$X_mr = max\{X_r\} \tag{3.3}$$

is the parameter (x) used for computing the gradient membership $\mu_c(P)$ with a S type function, as shown equation (2.1). $\mu_c(P)$ is used to represent the uncertainties of edge strength.

$$\mu_c(P) = S(x; a_1, b_1, c_1), \tag{3.4}$$

where, $b_1 = (a_1 + c_1)/2$

The feet and the shoulder points of Fig. 2.1 have been computed using min(X_mr) and max(X_mr) values of the contrast ratios (X_mr), over which the membership $\mu_c(P)$ is mapped.

$$\begin{aligned} c_1 &= max(X_{mr}) \\ a_1 &= min(X_{mr}) \end{aligned} \tag{3.5}$$

The histogram plots of pixel contrast ratios are shown in Figs. 3.8(a) and 3.8(b) for the images of Fig. 3.7(a) and Fig. 3.7(b) respectively.

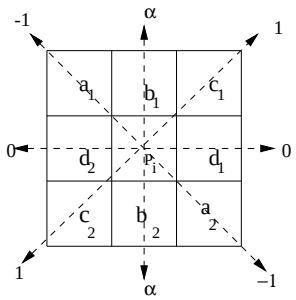


Figure 3.2: Computation of pixel contrast ratio from 3×3 neighborhood of a pixel.

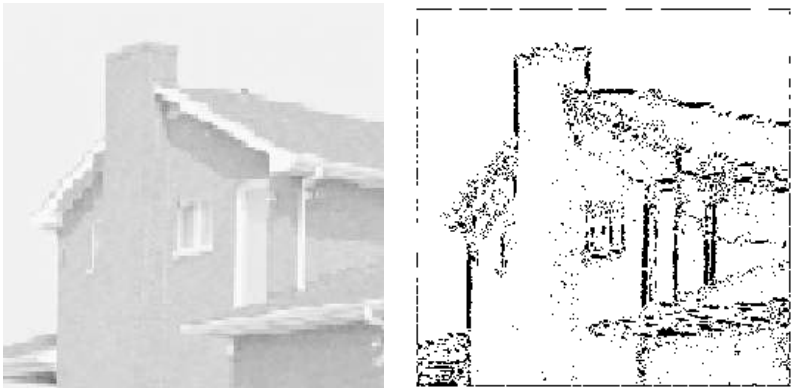
Table 3.1: Selection of Membership for generating candidates to test cornerness properties

d_{ac}	d_{al}	x	function used	generated edge map
255	0	$x = \frac{1+d_{al}}{1+d_{ac}}$	$\mu_m(P) = k_m \exp(-x)$	Figs. 3.4(b), 3.5(b)
maximum contrast	0	$x = \frac{1+d_{al}}{1+d_{ac}}$	$\mu_m(P) = k_m \exp(-x)$	Figs. 3.4(c), 3.5(a)
		$x = max(X_r)$	S(x;a,b,c)	Figs. 3.3(b), 3.5(c)

The value of $\mu_c(P)$ determines the edge strength. Higher values of gradient memberships, i.e. $\mu_c(P) \geq 0.5$ correspond to medium and strong edge points. Lower values of $\mu_c(P)$ correspond to weak or noisy edge points.

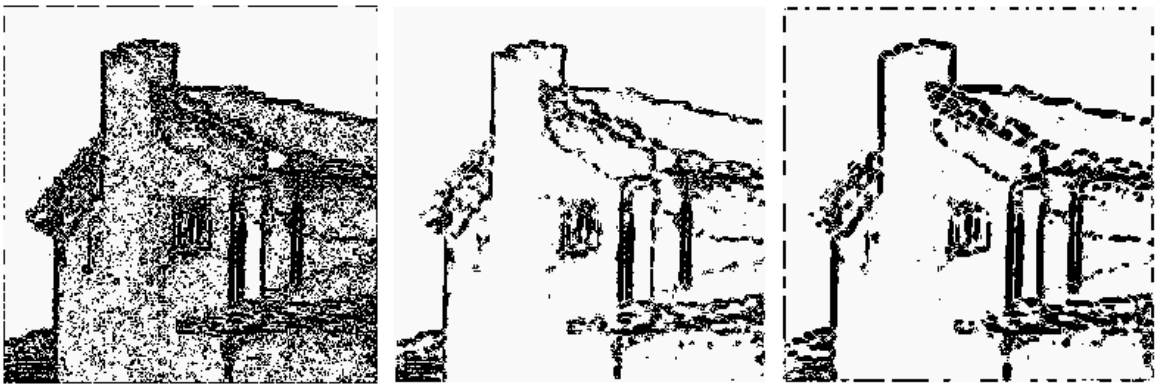
The fuzzy gradient map (e_d) as shown in (3.1) is obtained.

Different edge maps as obtained using different membership functions are shown from Figs. 3.3- 3.5. The edge maps obtained using eqn.2.10 as explained in Chapter2, are shown in Figs. 3.4(b), 3.5(b). The edge maps as shown in Figs. 3.4(c), 3.5(a) are also obtained using eqn.2.10 however the value of d_{ac} has been taken equal to the maximum local contrast value, as obtained for an image instead of 256. Figs. 3.3(b), 3.5(c) have been generated using a standard S-type function.



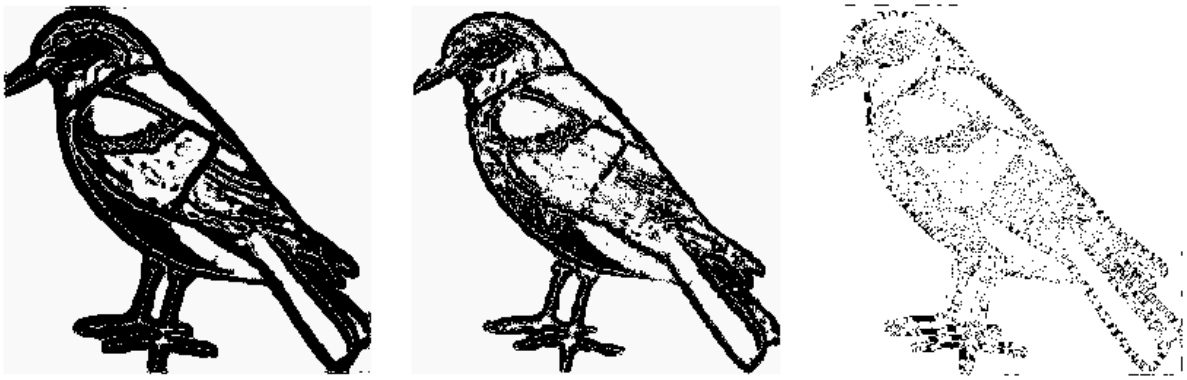
(a) (b)

Figure 3.3: (a) Original image (b) edge map with S-type function at $\mu_m(P) \geq .0.7$



(a) (b) (c)

Figure 3.4: Fuzzy edge maps (a) gradient map without linking, thresholded at $\mu_m(P) \geq 0.8$ (b) refined edge map with exponential function, thresholded at $\mu_m(P) \geq 0.8$, set (B) (c) using exponential function with d_{ac}= maximum local contrast, as obtained within an image.



(a) (b) (c)

Figure 3.5: Refined edge map, thresholded at $\mu_m(P) \geq 0.7$ for all three cases. (a) with exponential function, d_{ac}= maximum local contrast, as obtained within an image (b) with exponential function, with d_{ac}=255 (c) edge map with S-type function.

3.2.2 Estimation of connectivity strength $\mu_f(P)$ and $\mu_b(P)$

Two membership values ($\mu_f(P)$ and $\mu_b(P)$) are computed on a selected subset of (e_d) shown in (3.1) obtained by thresholding (e_d). The memberships $\mu_f(P)$ and $\mu_b(P)$ are computed from the difference in edge directions between the connected pixels within a fixed window. The actual computation of $\mu_f(P)$ and $\mu_b(P)$ are made as follows :

Let ϕ =$\{\phi_1, \phi_2, ...\phi_n\}$ represent the edge direction of a sequence of pixels on an edge segment. The present approach, deals with the changes in edge directions. To compute relative change in direction, the angle subtended between two successive pixels are considered in a 3×3 window as shown in Fig. 3.2.

The directions (ϕ) along the horizontal line i.e. (0 and 180 deg.) are labeled as (0), similarly along the vertical lines as (∞) and along the diagonal lines as (+ 1, -1) as shown in Fig. 3.2. As a result, the edges along different directions may be labelled as shown in (3.6).

$$A_{df} \in \{0, 1, \infty, -1\} \tag{3.6}$$

The change of directions with respect to (ϕ) between the successive edge pixels may have values ($\phi + \pi/4$), ($\phi - \pi/4$), ($\phi + \pi/2$), ($\phi - \pi/2$) in an eight neighborhood. However due to blurring of the images, the sharp changes like ($\phi + \pi/2$), ($\phi - \pi/2$) between the successive pixels are converted to gentle changes having values less than $\pi/2$. As a result, the changes at a step of 45 degrees are considered.

If the direction of the candidate pixel P is ϕ, then ϕ_f =ϕ + $\pi/4$ is considered as relative forward direction and ϕ_b =ϕ - $\pi/4$ is considered as the relative backward direction with respect to ϕ. A $m \times m$ window is centered around the selected candidate edge pixels and the number of simply connected edge pixels of (e_d) which have directions ϕ_f and ϕ_b are counted. If the label of ϕ is (0) then, the labels (1, -1) represents the counts n_f and n_b respectively. Similarly if the label of ϕ is (1), the labels (∞,0) represent the counts n_f and n_b respectively and so on.

This count is expected to vary with the sharpness of the curvature type. The values ($\mu_f(P)$, $\mu_b(P)$) are represented with the form of membership function,

$$\mu_f(P) = K * \exp(-x) \tag{3.7}$$

where $x = \frac{1}{n_f}$,

Similarly μ_b is defined by

$$\mu_b(P) = K * \exp(-x) \tag{3.8}$$

where $x = \frac{1}{n_b}$, K is a constant multiplier. It is so selected that the value of $\mu_f(P)$ or $\mu_b(P)$ should lie in between 0 to 1.0 from the finite counts of n_f and n_b of the image. K is determined from the maximum count of n_f or n_b

Each candidate edge pixel (P) selected for cornerness testing, is thus represented by a 3-dimensional feature vector F_i where,

F_i = [$\mu_c(P)$, $\mu_f(P)$, $\mu_b(P)$]. Extraction of possible fuzzy corners from the input edge map (e_d) will be discussed in the next section.

3.3 Multilevel fuzzy corner extraction

The fuzzy edge map (e_d) is represented as set of points $\{(\mu_c(P), P_c)\}$. In the initial stage, a suitable threshold value of gradient membership, has to be decided to select a subset $E_{d\alpha}$ of e_d, and only those points are used for computation of, $\mu_f(P)$, $\mu_b(P)$ for detection of fuzzy corners.

3.3.1 Membership transformation

Any natural image consists of different homogeneous regions, where the shape of each region is characterized by its bounding lines. But in many practical situations the boundaries are so faint that it becomes difficult to distinguish between two regions. Moreover due to noise and non uniform illumination, spurious edges may also appear. It is also difficult to discriminate between spurious edges and weak edges. Under such situation, the gradient information (both edge strength, and direction information) may be required to cut of where $\mu_c(P)$ is very small. To locate points from significant portions on the image, a contrast transformation may be used as a preprocessing step. The extraction of probable edge candidates, is achieved by thresholding through non-linear transformation of membership values $\mu_c(P)$ such that, the points having values greater than 0.5 are stretched and those below 0.5 are squeezed.

$$E_{df} = T'(e_d) \tag{3.9}$$

A pixel contrast transformation operation [155] is represented in (3.10)

$$\begin{aligned} \mu_{d1}(P) &= 2 \times \mu_c(P)^2 && ,0 \leq \mu_c(P) \leq 0.5 \\ &= 1 - 2 \times (1 - \mu_c(P))^2 && ,0.5 \leq \mu_c(P) \leq 1.0 \end{aligned} \tag{3.10}$$

The results before and after transformation of membership values for Figs. 3.7(a) are shown in Fig. 3.9(a), Fig. 3.9(b) and (c). As seen from Figs. 3.9(a), (b) the number of insignificant candidate points, considered for cornerness testing are reduced at the same

threshold value. Thresholding the transformed edge map (E_{df}=$\{\mu_{d1}(P), P\}$) above different membership values may be obtained by using proper ($\alpha - cuts$) [155] as mentioned in section 3.2. As a result, we obtain the edge maps $E_{d\alpha}$ at different levels from E_{df}, as shown in (3.11).

$$E_{d\alpha} = \{P \in E_{df} : \mu_{d1}(P) \geq \alpha\} \tag{3.11}$$

where $0 \leq \alpha \leq 1.0$. The candidates of $E_{d\alpha}$ can be represented by the local features F_{if} = [$\mu_{d1}(P)$,$\mu_f(P)$, $\mu_b(P)$]

By such thresholding of E_{df}, multilevel fuzzy edge maps may be generated, where the pixels may be segregated as (strong, medium, weak) edge pixels based on their gradient membership values $\mu_{d1}(P)$ as shown in Figs. 3.12(b), (c), (d) for Fig. 3.12(a).

If the local contrast of a region is very poor, then $\mu_{d1}(P)$ values of different edge points are very close to each other. Ambiguity in locating curvature points in these regions may increase due to close proximity of values of different points, as seen in the bottom rectangle of Fig. 3.12(b). On the other hand, the membership values of different points are widely separated above the cross over points ($\mu_{d1}(P) \geq 0.5$), where the local contrast is better resulting in less ambiguity.

In the transformed set, points having ($\mu_{d1}(P) \geq 0.5$) will include edge points with higher and medium strength. Whereas those having values ($\mu_{d1}(P)) \leq 0.5$) will select lot of spurious edge points along with high and medium type of curvature points.

Thus a proper choice of threshold ($\mu_{d1}(P)$) selection is necessary, below which the variations are considered to be noise.

3.3.2 Selection of threshold on membership value

The gradient membership value $\mu_{d1}(P)$ used for thresholding the edge map, is decided from the pixel contrast ratio histogram.

The histogram of contrast ratio gives an estimate of global description of the appearance of an image.

In general, the choice of threshold is made as follows:

A higher threshold value, typically $\mu_{d1}(P) \geq 0.8$ is chosen, to reduce the false acceptance rate, if the nature of the contrast histogram is as follows : (i) The contrast histogram occupies most of the histogram levels, which are in contiguous locations. (ii) The number of occurrences for each ($X_m r$) value is quite close and covers the majority of the total dynamic range. This is seen from the histogram plots of Figs. 3.7(a), 3.10(a), (b).

As we are concerned with the dynamic range, and not the absolute gray scale values, such thresholding can be applied for almost all natural images, even undergone varying imaging conditions like overexposed, underexposed, blurred, etc. The contrast histogram plot for the Figs. 3.10(a), (b) are shown in Figs. 3.11(a), (b).

On the other hand, a lower threshold value of $\mu_{d1}(P)$ typically $\mu_{d1}(P) > 0.0$ is chosen, if the histogram has the following properties. (i) Sparsely distributed contrast levels. (ii) Having widely different occurrences for different ($X_m r$) values (iii) Does not cover majority of the dynamic range. Such cases may arise for nearly binary images as seen for, Figs. 3.7(b), Fig. 3.21. In such cases transformation of $\mu_c(P)$ to $\mu_{d1}(P)$, does not affect the results much, as the candidate weak edges are less in number.

This has been tested over number of images and the strategy described is found to be satisfactory.

3.3.3 Estimation of local shape parameters

Once the suitable threshold value of $\mu_{d1}(P)$ is chosen, the next task is to categorize the edge pixels based on the local properties estimated from $\mu_f(P)$ and $\mu_b(P)$. The selected edge candidates constitute the points of $E_{d\alpha}$ for which the membership values $\mu_f(P)$, $\mu_b(P)$ are computed. The properties of $\mu_f(P)$, $\mu_b(P)$ are used to examine local shape parameters, which are defined as straightness and cornerness. Properties of $\mu_f(P)$ and $\mu_b(P)$ for any of the selected points (P) on the edge map is shown in Table. 3.2.

Straightness : This property is determined by comparing pixels translated along the direction of edge. It is expected that a pixel translated in the direction of straight edge will be connected to pixels of same direction. Hence $\mu_f(P)$ and $\mu_b(P) \simeq 0.0$.

Table 3.2: Fuzzy cornerness measure

$\mu_f(P)$	$\mu_b(P)$	cornerness	Straightness	Location
high	low	high	low	forward arm
high	high	high	low	near curvature junction
low	high	high	low	backward arm
low	low	low	high	straight edge

Cornerness : This property is determined from comparing pixels having reflexive symmetry. The pixels are expected to be reflected from one arm to the other on both sides of the curvature junction within the region of evaluation, as shown in Fig. 3.6.
The points of $E_{d\alpha}$ as shown in equation (3.11) having both $\mu_f(P)$ and $\mu_b(P)$ equal to zero can be filtered out as the non corner pixels. As a result, the interesting regions constituting a group of curvature points of the fuzzy edge image can be separated. We attempt to approximate this region with a quantitative measure by exploiting the properties of $\mu_f(P)$ and $\mu_b(P)$.

The pixels in the proximity of the curvature junction as shown in Fig. 3.6 can be categorized from the following rules. (i) The points with $\mu_f(P)$ high and $\mu_b(P)$ low constitute the points on the left side of the junction point. We designate these points as P_{ijf}, (as shown in Fig. 3.6) on forward arm and assign membership $\mu_f(P)$ -$\mu_b(P)$. This difference is expected to vary with the sharpness of curvature. The points of P_{ijf} represent a fuzzy subset as μ_{farm}.

$$\mu_{fram}(P) = \mu_f(P) - \mu_b(P) \tag{3.12}$$

(ii) The points with $\mu_b(P)$ high and $\mu_f(P)$ low constitute the points on the right side of the junction point. We designate these points as P_{ijb}, (as shown in Fig. 3.6) on backward arm and assign membership μ_b -μ_f. The points of P_{ijb} represent a fuzzy set μ_{barm}.

$$\mu_{bram}(P) = \mu_b(P) - \mu_f(P) \tag{3.13}$$

(iii) The points very near to the junction is expected to have high or medium values of $\mu_f(P)$ and $\mu_b(P)$.

Having obtained the two fuzzy sets, μ_{fram} and μ_{bram}, a third fuzzy subset μ_{cen} which is surrounded by both μ_{fram} and μ_{bram} lie approximately on the axis of symmetry. This region constitute the ambiguous corners. The cluster of such points (*) represented as μ_{cen} are shown in Figs. 3.13(a), 3.14(a), 3.15(a), 3.16(a) respectively for image 3.7(a). The points belonging to μ_{cen} are those points, having other points with $\mu_{fram}(P) > 0$ and $\mu_{bram}(P) > 0$ in the neighborhood of fixed window size.

The extracted curvature points may be of different sharpness type (sharp, medium, weak). The characteristics of sharp curvature points will be confined within a small region but for that of medium and weak type the region will be larger. In view of the above facts, we use a measure T_h that controls the shape and size of the extracted μ_{cen}.

In order to define a quantitative measure of the region, constituting of points of μ_{cen}, we compute the sum total of differences for all the pairs of μ_{fram} and μ_{bram} which fall within the region of evaluation, a nxn window. This value is subtracted from a large value y_{max} (kept fixed at 2.0, found experimentally better) to make it increase with sharpness.

$$T_h = y_{max} - \sum_{i=-n/2}^{j=n/2} \sum_{i=-n/2}^{j=n/2} \mu_{farm} - \mu_{barm} \tag{3.14}$$

The representative points i.e, the cluster center of each localized region (μ_{cen}) is represented by C_{ij} whose coordinate is equal to the average value of the co-ordinates of the n points of each cluster as shown in Figs. 3.13(b), 3.14(b), 3.15(b), 3.16(b). C_{ij} = $[\sum x_j/n, \sum y_j/n]$

At at fixed value of $\mu_{d1}(P)$ the value of T_h is experimentally varied from (0.1 to 0.3) to generate corners of different sharpness.

Computational complexity : The analysis of the computational complexity (worst

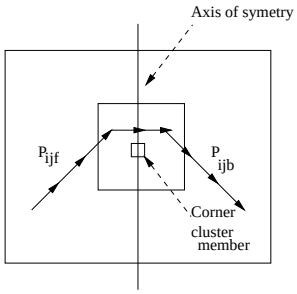
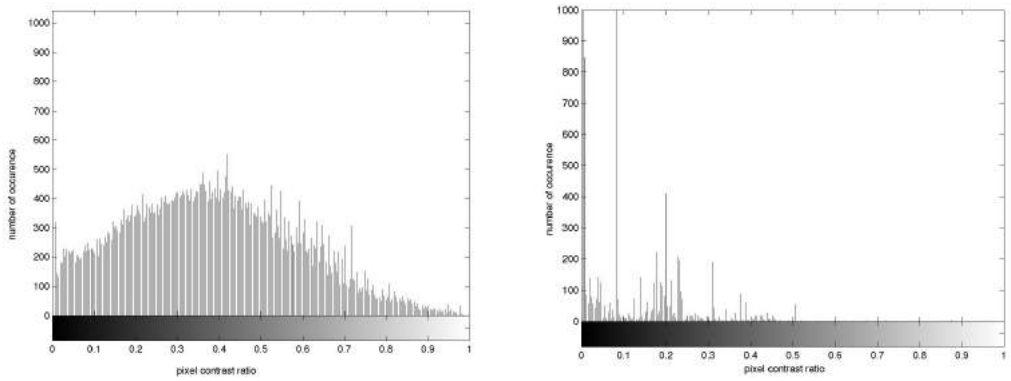


Figure 3.6: Determination of cornerness.

(a) (b)

Figure 3.7: (a) Original image of house. (b) Image having prominent curvature junctions.

case) involved in different operations for an image of size $M \times M$ and with window neighborhood $N \times N$, is explained as follows: 1. Identification of border regions requires M^2N^2 operations. 2. Computation of pixel contrast ratio involves λM^2N^2 operations (where $0 < \lambda < 1.0$). 3. Assignment of $\mu_c(P)$ involves $\lambda\ M^2$ operations. 4. Membership transformation involves λM^2 operations. 5. Thresholding involves λM^2 operations. 6. computation of μ_f and μ_b involves λM^2N^2 operations. 7.Grouping of pixels based on fuzzy rules involves λM^2 operations.8. Computation of T_h involves λM^2N^2 operations. The total operation is of the order of O(M^2N^2).



(a) (b)

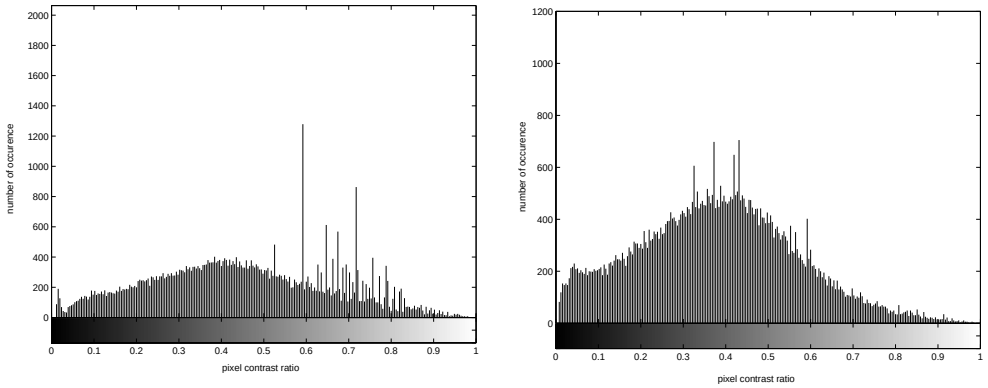
Figure 3.8: (a) Pixel contrast histogram of Fig. 3.7(a) (b) pixel contrast histogram of Fig. 3.7(b).

(a) (b) (c)

Figure 3.9: Fuzzy edge map (a) ($\mu_c(P) \geq 0.4$) (b) ($\mu_{d1}(P) \geq 0.4$) after membership transformation (c) ($\mu_{d1}(P) \geq 0.9$)

(a) (b)

Figure 3.10: Image of house (a) underexposed (b) overexposed

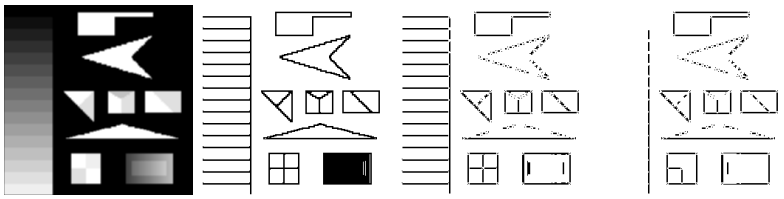


(a) (b)

Figure 3.11: Pixel contrast histogram plots for image of house (a) underexposed (b) overexposed

3.4 Experimental Results

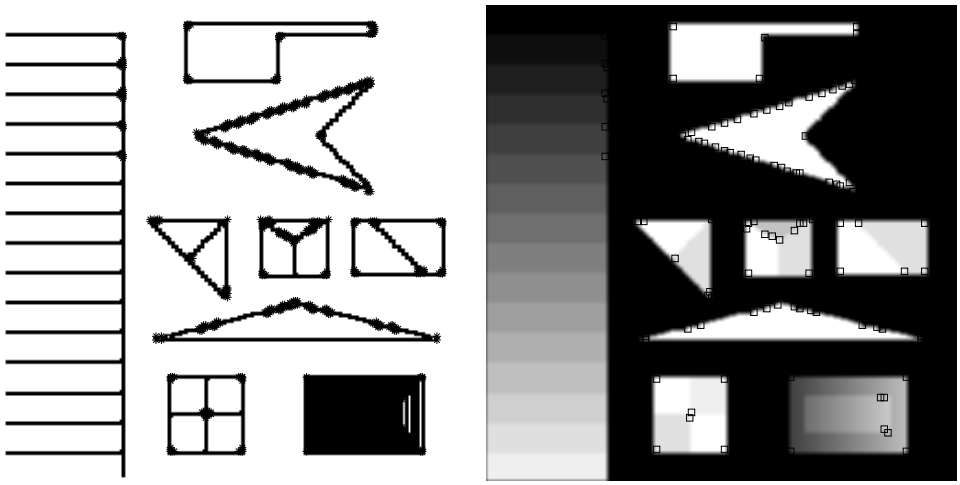
The performance of proposed detector has been examined on various types of images including images which have undergone image alterations like blurring, illumination change, noise, etc. Image as shown in Fig. 3.12(a) contains objects of different shapes with varying illumination. The procedure involves extraction of edge map. The edge map of Fig. 3.12(a) are thresholded above different membership values as shown in Figs. 3.12(b), (c), (d) for ($\mu_{d1}(P) \geq 0.0$), ($\mu_{d1}(P) \geq 0.6$), ($\mu_{d1}(P) \geq 0.9$) respectively. The points above the threshold values are represented as dark edge pixels. It is to be noted that the internal structure of the rectangle (represented as dark region) shown in Fig. 3.12(b) could not be extracted properly due to poor contrast between the two regions. In such cases, the extraction of the structure could not be done properly as the gradient value of the edge pixels are very low and the threshold is ($\mu_{d1}(P) \geq 0.0$). At a higher threshold value of $\mu_{d1}(P)$, stronger edge points mainly representing the boundary points can easily be separated as shown in Fig. 3.12(d). The curvature points of various regions are depicted by symbol '*' as shown in Figs. 3.13(a), 3.14(a), 3.15(a), 3.16(a). The representative point from each cluster is shown in Figs. 3.13(b), 3.14(b), 3.15(b), 3.16(b). The different type of curvature points are obtained by varying T_h and $\mu_{d1}(P)$ and shown in Figs. 3.13(b), 3.14(b), 3.15(b). The results of our algorithm are comparable to that of most popularly used corner detectors like Harris and SUSAN detector and are shown in Figs.3.17 - 3.18. The performance of different corner detectors varies with the type of the image and to obtain the best results, several parameters need to be adjusted for almost all detectors. We have tried to compare our results with the best results obtained from each detector with the parameter values as suggested by authors. In our algorithm better results are obtained by keeping the threshold of $\mu_{d1}(P)$ at a lower value when there are lesser number of gray level variations e.g., Fig. 3.7(b). On the other hand when there are large variations of distinct gray values as that of Fig. 3.7(a), higher threshold value of $\mu_{d1}(P)$ is chosen to reduce the number weak and noisy edge points. Such results are shown in Figs. 3.18 - Figs.3.20. It is seen from Fig. 3.17(a), (b), (c) that the corner points



(a) (b) (c) (d)

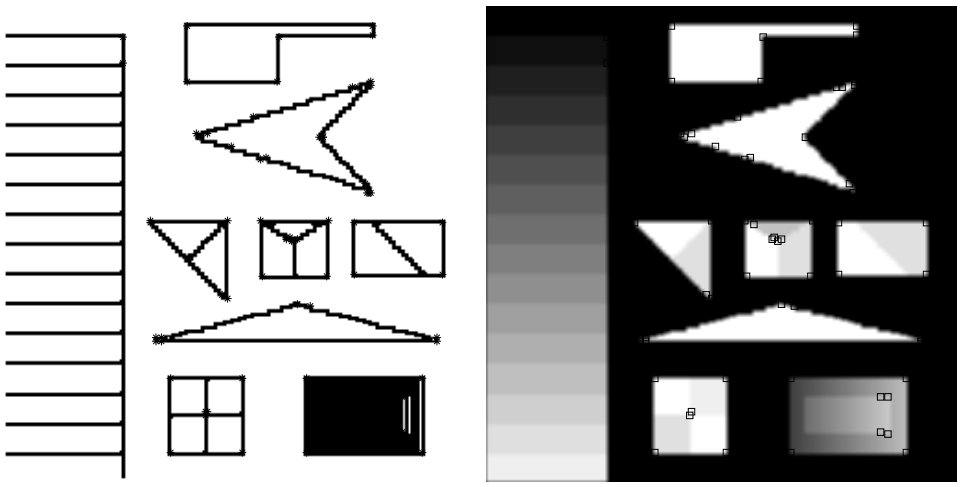
Figure 3.12: (a) Original image, same as Fig. 3.7(a) (b) edge image for ($\mu_{d1}(P) > 0.0$) (c) ($\mu_{d1}(P) \geq 0.6$) (d) ($\mu_d(p) \geq 0.9$). Points above threshold plotted as crisp edge points.

obtained by our method Fig. 3.17(a) is quite comparable to that of Harris shown in Fig. 3.17(b) and SUSAN detector as shown in Fig. 3.17(c). However, SUSAN is able to extract corners from very low contrast area. The results on the house image with threshold value ($\mu_{d1}(P) \geq 0.9$ and T_h= 0.2) for our algorithm, Harris and SUSAN are shown in Figs. 3.18. It is seen from Fig. 3.18 that the corner points obtained by our method Fig. 3.18(a) is comparable to that of Harris as seen in Fig. 3.18(b) and SUSAN detector in Fig. 3.18(c). Our result is closer to that of SUSAN with some more details of curvature information that exists in different regions of the house image. The results obtained under different imaging conditions are shown from Figs. 3.19 - 3.20. It is to be noted that our proposed detector is able to extract most of significant structural corner points under varying imaging conditions. This is due to the fact that the slope of the fuzzy property plane is determined from the dynamic range. Although the gray level contrast information is reduced in the overexposed case in Fig. 3.20, but due to additional contrast intensification, significant edge pixels are selected above threshold for cornerness detection. Even for nearly binary images our algorithm works satisfactorily as seen from Figs. 3.21(a), (b) and (c).



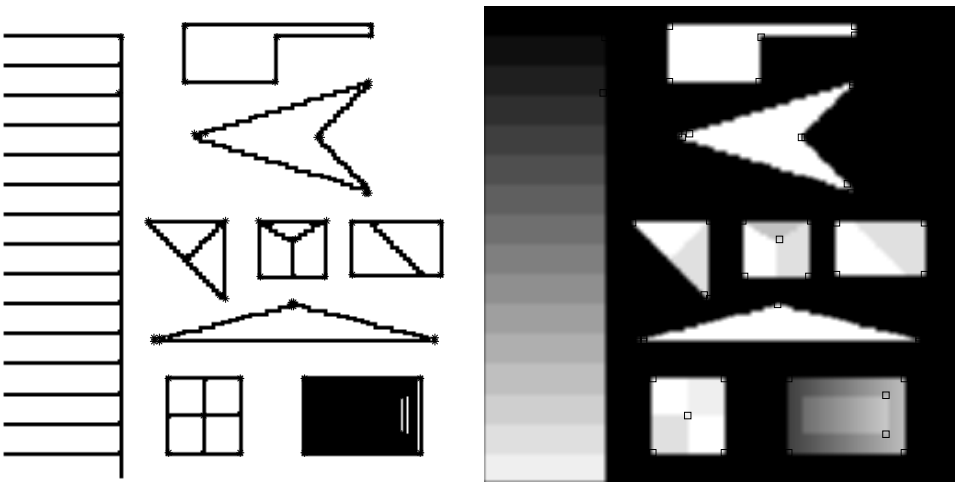
(a) (b)

Figure 3.13: (a) Curvature points (*), ($\mu_{d1}(P) > 0.0$ and T_h =0.1) (b) representative point of each cluster.



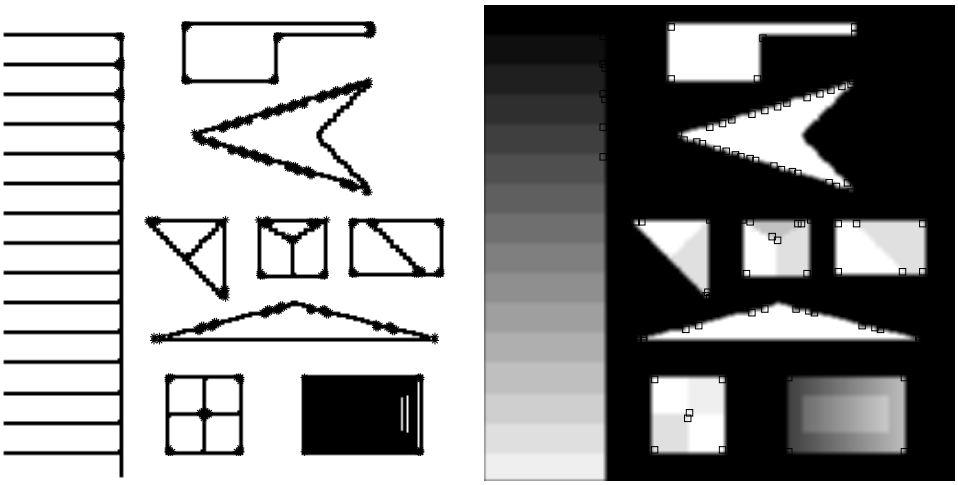
(a) (b)

Figure 3.14: (a) Curvature points (*), ($\mu_{d1}(P) > 0.0$ and T_h=0.2) (b) representative point of each cluster.



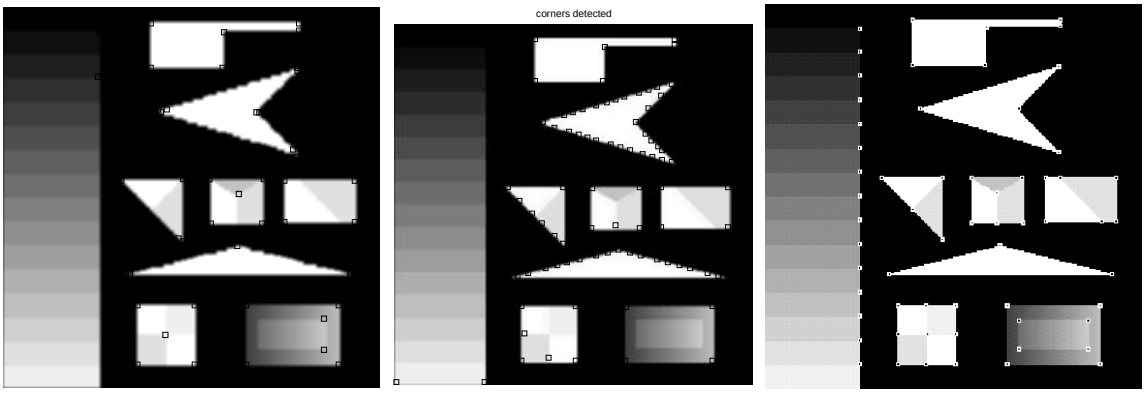
(a) (b)

Figure 3.15: (a) Curvature points (*), ($\mu_{d1}(P) > 0.0$ T_h=0.3) (b) representative point of each cluster.



(a) (b)

Figure 3.16: (a) Curvature points (*), (a) ($\mu_{d1}(P) > 0.6$ and T_h =0.1) (b) representative point of each cluster.



(a) (b) (c)

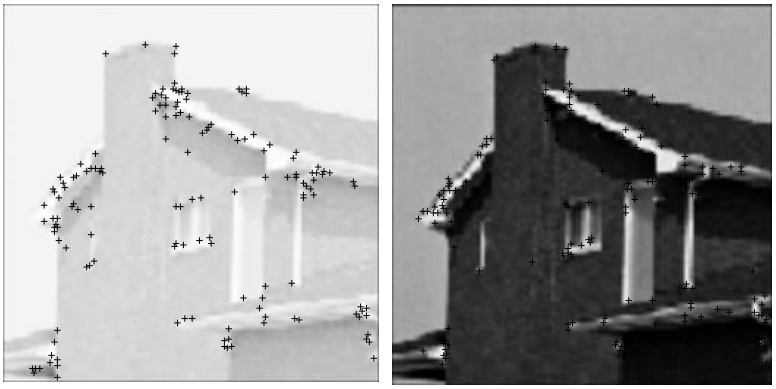
Figure 3.17: Corner points (a) Our detector ($\mu_{d1}(P) > 0.0$ and T_h=0.3) (b) Harris detector (c) SUSAN

(a) (b) (c)

Figure 3.18: Corner points (a) Our detector ($\mu_{d1}(P) \geq 0.9$ and T_h=0.2) (b) Harris detector (c) SUSAN

(a) (b)

Figure 3.19: Corner points from our detector (a) blurred image($\mu_{d1}(P) \geq 0.9$ and T_h=0.2 (b) house image with noisy regions.



(a) (b)

Figure 3.20: Corner points under illumination change (a) overexposed ($\mu_{d1}(P) \geq 0.9$ and T_h=0.3) (b) underexposed case ($\mu_{d1}(P) \geq 0.9$ and T_h=0.3)

(a) (b) (c)

Figure 3.21: Corner points (a) Our detector($\mu_{d1}(P) > 0.0$ and T_h=0.2) (b) Harris detector (c) SUSAN

3.5 Corner detection approach using Support Vector Machines

As seen from the experimental results, that the corner detection algorithm generates quite satisfactory results on gray level images. However, the performance of the detector based on fuzzy rule based approach is dependent on the tuning parameters $\mu_{d1}(P)$ and T_h. Although an ad-hoc threshold selection scheme based on pixel contrast histogram is proposed, but the complexity involved in total operations is of the order of O(M^2N^2). To enhance the performance of the detector in terms of reducing parameter tuning, an alternative classification based approach using SVM, has also been studied, which incorporates implicit learning capabilities. Using the labeled edge directions of the fuzzy input edge points, a classification based approach using SVM is experimented.

Support vector machines are a general class of learning architecture inspired from statistical learning theory that performs *structural risk minimization* on a nested set structure of separating hyperplanes [202]. Given a training data, the SVM training algorithm obtains the optimal separating hyperplane in terms of generalization error. We describe below the SVM design algorithm for a two class problem. Multiclass extensions can be

done by designing a number of one-against-all on one-against-one two class SVMs.

Suppose we are given a set of examples (x_1, y_1), ...(x_l, y_l), $x \in R^N$, $y_i \in \{-1, +1\}$, considering functions of the form *sgn* $((w.x) + b)$ and imposing the condition ,

$$inf|(w.x_i) + b| = 1, i = 1.., l \tag{3.15}$$

We would like to find a decision function $f_{w,b}$ with the properties $f_{w,b}(x_i) = y_i$, i= 1,2,...l. If this function exists, condition 3.15 implies that,

$$y_i((w.x_i) + b) \geq 1, i = 1, 2...l \tag{3.16}$$

In many practical situations a separating hyper plane does not exist. To allow for possibilities of violating eqn.3.16, slack variables are introduced. like

$$\xi_i \geq 0, 1 = 1, 1...ltoget \tag{3.17}$$

$$y_i((w.x_i) + b) \geq 1 - \xi_i, i = 1, 2...l \tag{3.18}$$

The support vector approach for minimizing the generalization error consists of the following. Minimize,

$$\phi(w, \xi) = (w.w) + C \sum_{i=1}^{l} \xi_i \tag{3.19}$$

subject to the constraints in eqns.3.17 and 3.18. It can be shown that minimizing the first term amounts in eqn. 3.19 to minimizing the VC dimension, and minimizing the second term corresponds to minimizing the misclassification error, with constraints C. The above minimization problem can be posed as a constrained quadratic programming (QP) problem [39]. The solution gives rise to a decision function of the form

$$f(x) = sgn[\sum_{i=1}^{l} y_i \alpha_i (x.x_i) + b \tag{3.20}$$

Only a small fraction of α_i co-efficients are non zero. The corresponding pairs of x_i entries are known as support vectors and they fully define the decision functions. The support vectors are geometrically the points lying near the class boundaries.

The aforesaid two class SVM can easily be extended for multiclass classification by designing a number of one-against-all two class SVMs, e.g., a k-class problem is handled with k two class SVMs.

3.5.1 The Proposed Approach

The proposed approach consists of two phases. In the first phase, with the fuzzy edge input image, a four dimensional feature vector describing an edge pixel is constructed. A local window about an edge point is considered, and the count of (other) edge pixels in that window having different labels is used to constitute a four dimensional feature vector for the edge pixel in the center of the window. Then a support vector machine is designed on a labeled set of four dimensional feature vectors. The support vectors generated in the design process correspond to corner points.

A four dimensional feature vector describing an edge pixel is constructed in the following way. Considering a $m \times m$ window about an edge pixel p_i, the number of other edge pixels within this $m \times m$ window which have directions 0, +45, 90, and -45 are counted. Let this counts be n_0, n_{+45}, n_{90}, and n_{-45} respectively. Let the direction for pixel p_i itself be $d_i \in \{0, +45, 90, -45\}$ which is the majority of the counts. The pixel is assigned a class denoting its own direction of maximum gradient, thus obtaining a labeled data set having four classes. Thus the feature vector for pixel p_i is $[n_0, n_{+45}, n_{90}, n_{-45}]$ and class label is d_i. It may be noted that for a non-corner edge pixel, all other edge points in the $m \times m$ window about it has the same direction as itself. As a consequence only one of the 0, $+1$, n_{90}, -1 is non-zero and the other three are zero. While for corner points multiple of these counts are non-zero which in shown in Fig. 3.22. A support vector machine (SVM) is then designed on this data. It is evident on analysis that the feature vectors for non-corner points will be interior points and those for corner points

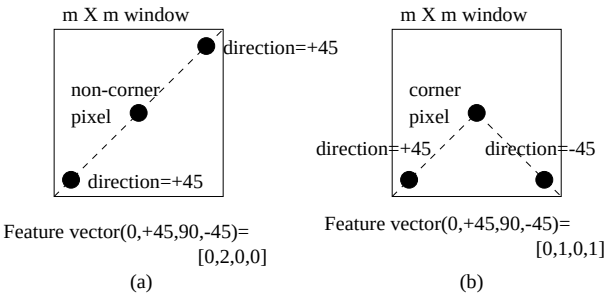


Figure 3.22: Computation of gray-level changes.

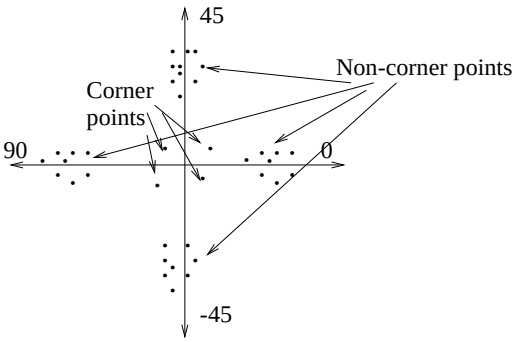


Figure 3.23: Visualization of the distribution of corner and non-corner points.

will be points near the class boundaries and correspond to the support vectors. Thus the support vector points generates the corner pixels. Thus if one considers the scatter plot of these four dimensional feature vectors, the non-corner points lie on the axes and are interior points of the corresponding class, while the corner points are off-axis points and constitute the boundary region of the class as shown in Fig. 3.23.

The corners obtained by the proposed algorithm are visually compared with the popular Harris corner detector for two graylevel images.

3.6 Conclusion

A fuzzy set theoretic approach for detection of corners is proposed in this chapter. The proposed algorithm does not require computation of chain codes or complex differential

Figure 3.24: Corners obtained for the house image using (a) Harris and (b) SVM based corner detectors

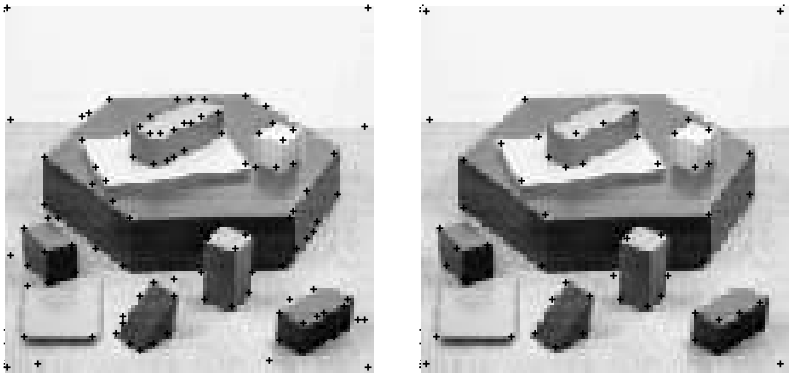


Figure 3.25: Corners obtained for the blocks image using (a) Harris and (b) SVM based corner detectors

geometric operators. Experiments have been performed on various types of images to illustrate the efficiency of the proposed algorithm. The algorithm performs reasonably well under different imaging conditions like changes in illumination, blurring, etc.
Corner detection has been considered an important problem due to its wide applications in the area of image processing and and computer vision. There are many corner detector algorithms reported in the literature. In most of the cases, the performance on the detector are dependent on several tuning parameters. Selection of right parameters, suitable for particular image is extremely important. In the present case, a new method of selection of tuning parameters is proposed, based on the pixel contrast histogram plot. However, the algorithm may be further refined, so that the parameters may be selected adaptively for thresholding and increase its suitability. To cope up with the problems like, parameter tuning and computational complexity, an alternative classification based approach using SVM, has also been studied, which incorporates implicit learning capabilities. The algorithm performs corner detection by learning on a given image. The SVM based method is faster than the fuzzy approach. The results obtained are quite satisfactory. Extensive experiments may result in good performance for a wide range of images and increase suitability and robustness of the proposed methodology. This study may be extended as a future scope.

The corner points and the points near about to it, holds a great potential in generating perceptually significant information. Significant features computed from these high curvature points can be used directly for characterizing an image for image retrieval application. As high curvature points depict perceptually significant information of an image with limited number of pixels, local color information of these points may be incorporated also with spatial information. However by combining some global(i.e., combining properties from all regions) one can have a CBIR system by representing images as a combination of significant local as well as overall semantic description of an image. An image retrieval system using these points, may be a cost effective proposition. This problem has been addressed in the next chapter.

Chapter 4

Retrieval of Color Images using Significant Features and their Evaluation with Fuzzy Model

4.1 Introduction

It is very common to see that most of the natural images contain edges and corners which are locally defined by their positions. These features are visually significant and encode a great proportion of information contained in the image, with limited number of pixels. Human visual system is highly efficient in selecting such significant features from an image, which may be used to evaluate similarity between images [53], [215]. This may lead to a comparatively fast and cheaper mechanism for image retrieval.

In this chapter, a color image retrieval scheme has been proposed, which incorporates significant information from the high curvature points, for characterizing an image.

In general, the notion of similarity between images is dependent both on local (object level representation) and global (image in total) view. For this reason, a large number of CBIR approaches employ detailed region representation via different segmentation techniques [218], [51], [138] for image retrieval task. As already mentioned in chapter 2,

that although segmentation is important for effective characterization of an image, but it is a difficult task without a prior knowledge about the objects present in an image.
Instead of detailed region representation, the information in the locality of a corner may be considered as a meaningful local feature descriptor. Information obtained from these significant points may be combined with some global measurements to generate better representation of an image.
The main focus of research in the previous problem was to generate efficient feature descriptors, rich in visually significant information and robust to different transformations like blurring, illumination changes, etc. With the extracted visually significant points designated as points of interest (poi), a CBIR problem as applicable in color images is dealt in this chapter.

Although extensive research on CBIR has been performed over decades, it has not been possible to achieve desired accuracy from a fully automated CBIR system because, of semantic gap. A detailed review has been made in chapter 1. In an image, each feature tries to capture only one aspect of the image property such as (color, texture, shape, etc.). In conventional CBIR approaches [107], [144], [43], [70], [103], [36], [50], [147] an image is usually represented by a set of features, where the feature vector is a point in a multidimensional feature space. It is therefore required to select the optimal set of features suitable for a particular type of query, in which the component features may also be varied according to its importance based on user's feedback for generating better results.

The motivation behind developing the proposed CBIR system could be explained in the following way. Feature extraction is one of the most important part in designing a CBIR system. The features should be well separated in the feature space. This requirement leads to selection of non redundant optimal set of features, based on some information criteria. For a query, the importance of the component features of the selected feature set, may further be varied by a suitable weighting factor based on the user's relevance feedback to improve the retrieval accuracy.
The proposed technique is based on the assumption that two visually similar images will

have similar visual characteristics.

The proposed methodology includes three parts. In the first part, the clusters extracted around different types of curvature points (sharp, medium, weak) as explained in the previous chapter, serve as the candidate points for computation of features. Invariant moments of the extracted point sets are used for similarity evaluation. Beside these features, some global measurements are also considered to improve the results further, as the notion of similarity is dependent on both local (significant curvature points and around) and global (image in total) view.

In the next part, a feature ranking of the extracted features has been made to select, a set of relevant and non-redundant features based on the mutual information criterion. From which it is shown, that not only the features from the significant locations are sufficient, but also the spatial and color information of the total image is necessary to generate subtle discrimination between images. This set of features are used to retrieve images in the first pass.

In the third part, a fuzzy entropy based feature evaluation mechanism [153] is proposed to enhance the accuracy of the system based on user's feedback. The user marks the relevant and the irrelevant set of images within the retrieved set. The individual feature weights are updated using a measure known as fuzzy feature evaluation index (FEI) [153]. This value is computed from the 'intraset ambiguity' and the 'interset ambiguity' as obtained from the relevant and irrelevant set of images. The results of the proposed methodology are compared with that of some well known techniques namely, integrated region based approach [206], [51] and color histogram method [176]. The results have also been compared in the framework of MPEG-7 visual feature standards.

The organization of the chapter is as follows: A brief discussion on image retrieval methods, using visually significant points and relevance feedback mechanism in provided in 4.2. In Section 4.3, the proposed methodology is described in details, while the experimental steps and corresponding extensive results are demonstrated in Section 4.4. The concluding remarks are given in Section 4.5.

4.2 Review of relevance feedback and image characterization methods

This section provides a brief discussion on some of the feature extraction techniques and relevance feedback mechanism used in CBIR applications.

4.2.1 Image characterization methods

Image characterization used in CBIR, are mainly based on (a) global feature based representation (b) region feature based representation. In the first approach, global histogram (color, texture orientation, entropy, etc.) are used to represent an image [183], [62], [139]. Similarity between images are computed based on some specified distance measures [30]. Global representation gives a gross measurement. Images can be more precisely described by the local properties. In the region based approach, an image is segmented into significant regions and each region may be treated separately for extraction of local features [51], [43]. The accuracy of the segmentation based techniques greatly depends upon the method used and the a priori knowledge about the total number of classes present in an image. One way of tackling uncertainty arising out of segmentation errors may be, to avoid a prior rigid segmentation and try to estimate local features from visually significant points.

From the psycho visual studies [123] it has been observed that, high curvature points (corners) play a significant role in characterizing an object. Characterizing images with visually significant points (sharp curvature) have been successfully applied in image matching and retrieval applications [121], [141]. A recent work on optimal use of color points of interest for CBIR is proposed in [76]. Detection of significant points based on the analysis of statistics of color derivatives has been recently proposed in [209]. Mikolajcvzyk proposed an interest point detector invariant to affine changes [141] where matching between two images are made from determining point to point correspondence. In [122], Scale Invariant Feature Transform (SIFT), invariant features have been proposed, which have

good applications for object recognition. An image retrieval technique in which local features (color, texture, etc.) are computed on a fixed sized window of regular geometrical shape, surrounding the corner points is proposed by Loupias et al., in [121]. Local image representations using pruned salient points for applications to CBIR has been proposed in [220]. In this approach, the features for salient points represent local characteristics of that point so that the similarity between features indicate similarity between the salient points.

In case of estimating local features around significant curvature points, proper selection of the window size and shape is necessary. A big window may incorporate lot of insignificant information(noise) along with significant information, whereas for small ones lot of important information may be left out. So the approach of fixed window selection may not always guarantee the best result. The curvature points may be of different types (sharp, medium, weak). The characteristics of sharp curvature points will be confined within a small region, but for that of medium and weak type the region will be larger. This fact indicates that, selection of set of points of interest (poi) based on the curvature type could be a better approach.

4.2.2 Relevance feedback mechanism

Effective relevance feedback mechanism has also been identified as an important [178], [224], [80], [45], [214], [192], [117] problem, used to bridge the semantic gap. Majority of the relevance feedback methods employ two approaches [177], [178] namely, query vector moving technique and feature re-weighting technique to improve retrieval results. In the first approach, the query is reformulated by moving the vector towards positive / relevant examples and away from the negative examples, assuming that all positive examples will cluster in the feature space. Feature re-weighting method is used to enhance the importance of those components of a feature vector, that help in retrieving relevant images, while reducing the importance of the features that does not help. However in such

cases, the selection of positive and negative examples, from a small number of samples having large number of features, still remain as a problem.

In MARS, [177] a feature weight is determined by examining the variance of the features, across the set of retrieved images marked as relevant by the user. Relevance feedback techniques in CBIR, have mostly utilized information of the relevant images but have not made use of the information from irrelevant images. Zin et al., [94] have proposed a feature re-weighting technique by using both the relevant and the irrelevant information, to obtain more effective results.

Recently, relevance feedback is considered as a learning and classification process, using classifiers like Bayesian classifiers [203], neural network [165], etc. However trained classifiers become less effective when the training samples are insufficient in number. To overcome such problems, active learning methods have been used in [85]. Our proposed method uses the concept of combining information from both relevant and irrelevant images, returned to the user. The aim of relevance feedback approaches is to develop new methods for iterative improvement of the results, based on a particular set of image features. However, generation of satisfactory results even from the first set of retrieved images is also desirable. This necessitates development of suitable low cost feature extraction methods, for better representation of images.

4.2.3 Definitions and formulations used

The image properties used in the feature extraction and feature evaluation methods are explained in this section.

fuzzy entropy : The entropy of a fuzzy set having n points is defined as [155],

$$H(A) = (\frac{1}{n \ln 2}) \sum_i S_n(\mu(x_i)); i = 1, 2...n \tag{4.1}$$

where the Shannon's function ,$(S_n\mu(x_i))$=-$\mu(x_i)$ln$\mu(x_i)$-{1-$\mu(x_i)$}ln{1-$\mu(x_i)$}. Entropy is dependent on the absolute values of membership (μ) and satisfies the properties, $H_{min} =$

0 for μ=0 or 1, $H_{max} = 1$ for μ=0.5

Invariant moments :

Moments are a global descriptor of a shape with invariance properties [75]. The moments m_{pq} of order p and q of a function $f(x, y)$ for discrete images, are usually approximated as,

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \tag{4.2}$$

The centralized moments are expressed as,

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})(y - \bar{y}) f(x, y) \tag{4.3}$$

where $\bar{x}$=$\frac{m_{10}}{m_{00}}$, $\bar{y}$=$\frac{m_{01}}{m_{00}}$

The centralized moments upto order 2 are expressed as,

$$\begin{aligned} \mu_{00} &= m_{00} \\ \mu_{20} &= m_{20} - \frac{m_{10}^2}{m_{00}} \\ \mu_{02} &= m_{02} - \frac{m_{01}^2}{m_{00}} \\ \mu_{11} &= m_{11} - \frac{m_{10} * m_{01}}{m_{00}} \end{aligned} \tag{4.4}$$

The normalized central moments are computed as,

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{4.5}$$

where $\gamma = \frac{p+q}{2} + 1$ for $p + q$ = 2, 3,.... A set of seven moments invariant to translation, rotation and scale can be computed from η_{pq}. Of which the invariant moments (ϕ^1) is considered for the proposed feature extraction represented as (4.6).

$$\phi^1 = \eta_{20} + \eta_{02} \tag{4.6}$$

4.3 The proposed methodology

The feature extraction and feature evaluation algorithms used in the proposed technique may be explained as follows. We obtain a number of clusters of significant curvature points (signature), whose centroid almost represent the stable corner point. This set of points can be considered as the reduced representative set of the original image. The detailed methodology for extraction of significant points has been explained in Chapter 3. A set of relevant and non redundant features is selected using the mutual information based minimum redundancy-maximum relevance criterion. The importance of the individual feature component is evaluated from fuzzy feature evaluation index (FEI) and the weights are updated based on the (FEI) values.

In the present work, thresholding e_d, of equation 3.1, at values ($\mu_{d1}(P) \leq 0.5$) are not considered because, this may select lot of spurious curvature points along with high and medium type and reduce the accuracy of the techniques. The points extracted as high curvature points from the edge map $E_{d\alpha}$ with $\mu_{d1}(P) \geq 0.6$ are chosen for computation of features, which characterize the perceptually significant regions. Experimentally, better results are found with values $\mu_{d1}(P) \geq 0.5$, typically greater than (.6, .7, .8). The image is represented by a signature consisting of a set of clusters, $\{c_k = m_k, w_{mk}\}$ where m_k represents the spatial locations of the fraction w_{mk} of pixels that belong to each cluster. k ranges from one to the number of clusters extracted from the edge map.

The extracted signature for some images are shown in Figs. 4.1, 4.2, 4.3. The procedure is implemented for color images by converting RGB plane to HSI and considering only the intensity component for detecting the significant points.

4.3.1 Computation of color moments at selective points

Color moments have been very successfully used in many retrieval systems like QBIC [62], [70], [44]. Color image can be specified by tristimulus values (RGB, LUV, YCrCb etc.) where each image can be split into independent image planes corresponding to

different attributes of an image namely hue, intensity, etc. The first order(mean), the second (variance) have been proved effective for representing color distributions of images. Among the different color models reported in literature, normalized rgb representation, illumination and viewing geometry invariant representations, which mostly belong to HSI family of color model are popular. In addition to these traditional color spaces, new invariant color models (c_1,c_2,c_3) have been proposed in [69] which discounts the effects of shading and shadows also. We have chosen the (c_1,c_2,c_3) invariant feature model which are defined in the following (4.7). This model is able to denote the difference between two colors, based on their perceptual difference. Invariant color representation are popularly used in CBIR. However, these models have shortcomings under certain situations, due to loss of discrimination power among images. The RGB plane is converted to (c_1,c_2,c_3)

$$\begin{aligned} c_1 &= arctan(R/max(G,B)) \\ c_2 &= arctan(G/max(R,B)) \\ c_3 &= arctan(B/max(R,G)) \end{aligned} \tag{4.7}$$

The image is characterized in the following manner. The invariant moment (ϕ^1) as shown in section 4.2.3, computed from the locations (poi), will help to identify color similarity of the perceptually significant regions. Invariance of moments ensures better results for separate objects, than the description used for the image. However, the features from (poi) may be used to produce efficient matching between images, due to inherent structural information. Natural images (consisting of different objects), are more of semantic significance. Considering these facts, additional features are computed, taking all the points over the component planes. The combined set of features, would be able to provide better semantic representation. The proposed feature vector has six components as represented in (4.8).

$$f_l = \{\phi_l^1(S_c)\}_{l=1}^{6} \tag{4.8}$$

S_c represents the component planes. ϕ_l^1 are the invariant moments computed using equations, (4.2) to (4.6). For l ={1,2,3}, $f(x,y)_{S_c}$ = $\{c_1, c_2, c_3\}$ values respectively, of

all points $(x, y) \in M \times N$. For l ={4,5,6}, $f(x, y)_{S_c}$ = {c_1,c_2,c_3} value respectively, for $(x, y) \in w_m k$ (clusters) and default white (1) for other locations, as shown in Fig. 4.1(b). The feature vector is represented as, F= [$f_1, f_2, f_3, f_4, f_5, f_6$]. The similarity between images are evaluated, from measuring Euclidean distance between the feature vectors.

4.3.2 Feature selection method

For natural images (consisting of different objects) the representation obtained from corner signature, although important but may not be sufficient for discriminating them from other categories. To show this, the extracted features have been ranked using an optimal classifier. The minimum redundancy-maximum relevance based feature selection method, proposed by Battiti [25], is used to extract a set of nonredundant and relevant features. The relevance of a feature with respect to the class and the redundancy between two features are calculated using the mutual information. To calculate the mutual information, the values of each feature is thresholded using μ (mean) and σ (standard deviation): any value larger than $\mu + \sigma/2$ is transformed to state 1; any value between $\mu - \sigma/2$ and $\mu + \sigma/2$ is transformed to state 0; any value smaller than $\mu - \sigma/2$ is transformed to state -1. To evaluate the effectiveness of the selected features, the leave-one-out cross validation is performed using the C4.5 based decision tree algorithm [166]. The C4.5 is a popular decision tree-based classification algorithm. It is used for evaluating the effectiveness of reduced feature set for classification. The selected feature set are fed to C4.5 for building classification models. The C4.5 is used here because it performs feature selection in the process of training and the classification models it builds are represented in the form of decision trees, which can be further examined.

This approch has been popularly used for efficient image characterization. The features $[f_1, f_2, f_3 \ldots, f_6]$ are evaluated on all images of the (SIMPLIcity) database which consist of images from 10 different semantic categories. As seen from the classification results, of Table 4.1 and Table 4.2, that addition of the components f_4, f_5, f_6 with the components of f_1, f_2, f_3 improves the recognition score. This implies that combination of local prop-

erties (information from corner signature) and global moments (considering all points) are important for effective characterization of an image. However the feature f_5, f_6 alone donot take significant part in improving the recognition score.

The features [f_1 ... f_6] are invariant to rotation, translation and scaling by definition as explained in section. 4.2.3. To test the robustness of the features against noise, we inject noise at different levels with SNR rations (50dB, 20dB, 0dB and -10dB) on the query examples of a particular type. We add this class of images into the database images to form a separate noise injected class. We label the class as level 10. The error rate from such class is also shown in Table 4.1 and table 4.2. This indicates that the proposed feature is not noise invariant.

It has been mentioned that invariant representations can cover a wide range of perceptual variations, but at the same time lacks discrimination power among images. With the features f_1... f_6 computed directly from the RGB plane, i.e., without using equation (4.7) better recognition has been obtained for some classes, even for the noise injected case, as shown in Table 4.3 and Table 4.4. We have designated the feature set $[f_1, f_2, f_3..., f_6]$ as computed from (c_1, c_2, c_3) model using (4.7) as set (A1), and those computed directly from RGB plane as set (B1). The recognition score as obtained from both the cases are shown in Tables 4.1 - 4.4.

4.3.3 Estimation of relative importance of different features from relevance feedback

An effective relevance feedback system should provide solutions about how to learn and accumulate knowledge from small set of feedback images. Each type of visual feature tends to capture only one aspect of the image property. It is hard to specify how different aspects are to be combined to form a query. In the proposed approach, we combine discriminant information from the relevant and irrelevant set of retrieved images, to evaluate

Table 4.1: set (A1), classified as

feature:: 1 rank ::3

feature:: 2 rank ::2

feature:: 3 rank ::1

feature:: 4 rank ::-1

feature:: 5 rank ::-1

feature:: 6 rank ::-1

Errors :: 236(23.2 percent)

Image	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	
class 0	77	2	3	2	12	3	1			1		(a)
class 1	1	75	4	1	3	5			11			(b)
class 2	7	1	57	8	5	7			13	2		(c)
class 3	1	3		88			5		2			(d)
class 4	5	1			94							(e)
class 5	9	2	4	2	3	74	1	1	1	3		(f)
class 6	2			9			53	4		32		(g)
class 7	1	1				1		97				(h)
class 8	1	16	6	5	1	3		1	66	1		(i)
class 9	3			3	2	2	4	2		84		(j)
class 10								10			9	(k)

Table 4.2: set (A1), classified as

feature:: 1 rank ::3

feature:: 2 rank ::2

feature:: 3 rank ::1

feature:: 4 rank ::4

feature:: 5 rank ::-1

feature:: 6 rank ::-1

Errors :: 160(15.7 percent)

Image	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	
class 0	93		1	2			3			1		(a)
class 1	3	86		4	1	4			2			(b)
class 2	5	7	74	5	1	2	1		3	2		(c)
class 3	1	2		91			3		2	1		(d)
class 4	2				97			1				(e)
class 5	5	2	6	2	2	79	2	1		1		(f)
class 6	1		2	4		2	79	2	1		1	(g)
class 7	1					1		97			1	(h)
class 8	2	12	3	3	1	2	1	1	75			(i)
class 9	3		2	4	2	1	10	2		76		(j)
class 10							10				9	(k)

Table 4.3: set (B1), classified as

feature:: 1 rank ::3

feature:: 2 rank ::1

feature:: 3 rank ::-1

feature:: 4 rank ::2

feature:: 5 rank ::-1

feature:: 6 rank ::-1

Errors :: 204(20.0 percent)

Image	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	
class 0	82	3	2	2	2	6	1			2		(a)
class 1	4	69	4	8		7			7	1		(b)
class 2	9	6	76	3		5	2		1	1		(c)
class 3	6	9	5	72			4		3	1	1	(d)
class 4					98	2						(e)
class 5	4	6	3	2		70		5	4	6		(f)
class 6	2	1	2	3			84	1		5	2	(g)
class 7	3		1			6		76		14		(h)
class 8	3	3	3	5	1	10	2		73			(i)
class 9	10	2		3	1	3	6	1	1	72	1	(j)
class 10											18	(k)

Table 4.4: set (B1), classified as

feature:: 1 rank ::3

feature:: 2 rank ::1

feature:: 3 rank ::4

feature:: 4 rank ::2

feature:: 5 rank ::-1

feature:: 6 rank ::-1

Errors :: 189(18.6 percent)

Image	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	
class 0	82	1	3	2	2	6				4		(a)
class 1	1	70	5	9		5			8	2		(b)
class 2	5	4	81	3		2	2		2	1		(c)
class 3	6	3	7	75			5		3	1		(d)
class 4					100							(e)
class 5	4	5	3		1	74		5	3	5		(f)
class 6	3	4		2			88			3		(g)
class 7			1			8		76	2	15		(h)
class 8	1	3	5	6	1	1	1	2	80			(i)
class 9	7	3	2	4	2	3	6	1	1	70	1	(j)
class 10			1								17	(k)

the importance of individual components for a particular query.

Let an image database S_d be composed of d distinct images, I=$\{I_1$,I_2,...,$I_d\}$ where $I \in S_d$. For a query I_{qr}, the retrieval decision can be made by measuring similarity between the feature vectors, based on some specified distance measure. The image I is represented by a set of features F =$\{f_q\}_{q=1}^{N}$, where f_q is the qth feature component in the N dimensional feature space. The commonly used decision function for measuring similarity between the query image I_{qr} and other images I, is represented as,

$$D_{is}(I, I_{qr}) = \sum_{q=1}^{N} w_q ||f_q(I) - f_q(I_{qr})|| \tag{4.9}$$

where $||f_q(I) - f_q(I_{qr})||$ is the Euclidean distance between the qth component and w_q is the weight assigned to the qth feature component. The weights should be adjusted such that, the features have small variations over the relevant images and large variation over the irrelevant images. Let k similar images I_s=$\{I_1$,I_2,...,$I_k\}$ where, $I_k \in I_s$, are returned to the user. Let I_r be the set of relevant images and I_{ir} be the set of irrelevant images as marked by the user. I_r = $\{I_j \mid I_j$ relevant, for $I_j \in I_s\}$ and I_{ir} = $\{I_j \mid I_j$ irrelevant, for $I_j \in I_s\}$. The information from I_r and I_{ir} are combined to compute the relative importance of the individual featues, from fuzzy feature evaluation index (FEI) proposed by Pal et al., [153], [154] in pattern classification problems.

Feature evaluation index

The fuzzy measure (FEI) is defined from interclass and intraclass ambiguities and explained as follows. Let C_1, C_2,.... C_j ... C_m be the m pattern classes in an N dimensional ($f_1, f_2, f_q, ...f_N$) feature space where class C_j contains, n_j number of samples. It is shown that the value of the proposed measure, H_{qj} (fuzzy entropy) [155] as explained in section 4.2.3 gives a measure of 'intraset ambiguity ' along the qth co-ordinate axis in C_j. The entropy of a fuzzy set, having n_j points in C_j is computed as,

$$H(A) = (\frac{1}{n_j \ln 2}) \sum_{i} S_n(\mu_(f_{iqj})); i = 1, 2...n_j \tag{4.10}$$

where the Shannon's function ,$(S_n\mu(f_{iqj}))$=-$\mu(f_{iqj})$ln$\mu(f_{iqj})$-{1-$\mu(f_{iqj})$}ln{1-$\mu(f_{iqj})$}

For computing H of C_j along qth component as described in section 4.2.3, a standard S-type function is considered and the parameters are set as,

$$\begin{aligned} b &= (f_{qj})av \\ c &= b + max\{|(f_{qj})av - (f_{qj})max|, |(f_{qj})av - (f_{qj})min|\} \\ a &= 2b - c \end{aligned} \tag{4.11}$$

where $(f_{qj})av$, $(f_{qj})max$, (f_{qj})min denote the mean, maximum and minimum values respectively computed along the qth co-ordinate axis over all the n_j samples in c_j. Since $\mu(b)$ =$\mu(f_{qj})av$ =0.5, the values of H are 1.0 at $b = (f_{qj})av$ and would tend to zero when moved away from b towards either c or a of the S function. Selecting b= $(f_{qj})av$ indicates that, the cross over point is near to the query feature component. Higher the value of H, more would be the number of samples having $\mu(f)$ equal to 0.5. This indicates that, more would be the tendency of the samples to cluster around the mean value, resulting in less internal scatter within the class. After combining the classes C_j and C_k the mean, maximum and minimum values (f_{qkj})av, (f_{qjk})max, (f_{qjk})min respectively of qth dimension over the samples (n_j + n_k) are computed similarly, where n_k are the samples in class C_k. The criteria of a good feature is that, it should be nearly invariant within class, while emphasizing differences between patterns of different classes [153]. The value of H would therefore decrease, after combining C_j and C_k as the goodness of the qth feature in discriminating pattern classes C_j and C_k increases. The measure denoted as H_{qjk} is called "interset ambiguity" along q^{th} dimension between classes C_j and C_k. Considering the two types of ambiguities, the proposed Feature evaluation index (FEI) for the qth feature is,

$$(FEI_q) = \frac{H_{qjk}}{H_{qj} + H_{qk}} \tag{4.12}$$

The lower the value of FEI_q, the higher is the quality of importance of the qth feature in recognizing and discriminating different classes. The precision of retrieval can be improved with these values.

In the proposed algorithm, the number of classes are two. The relevant images constitute the (intraclass) and the irrelevant images constitute the (interclass) image features. To evaluate the importance of the qth feature, the qth component of the retrieved images is considered. i.e., $I^{(q)}$ =$\{I_1^{(q)}, I_2^{(q)}, I_3^{(q)},, ...I_k^{(q)}\}$
H_{qj} is computed from $I_r^{(q)}$= $\{I_{r1}^{(q)}, I_{r2}^{(q)}, I_{r3}^{(q)},, ...I_{rk}^{(q)}\}$. Similarly H_{qk} is computed from the set of images, $I_{ir}^{(q)}$ =$\{I_{ir1}^{(q)}, I_{ir2}^{(q)}, I_{ir3}^{(q)},, ...I_{irk}^{(q)}\}$. H_{qkj} is computed combining both the sets. Images are ranked according to Euclidean distance. The user marks the relevant and irrelevant set from 20 returned images, for automatic evaluation of (FEI).

Effect of sample size on evaluating importance

We address the issues like, how to manage with limited number of returned images, to extract discriminant information. The relevant feature components are expected to be closer in the feature space than the irrelevant components. Hence bandwidth defined as, B_r=$c_{qj} - a_{qj}$ of the standard $S-type$ function for computing intraset ambiguity is small. In the case of computing H_{qk} or H_{qjk}, the membership values spread over a wider range in the feature space. This indicates that features marked as irrelevant, also have significance in evaluating (FEI). To study the effect of image sample size on the FEI values, we combine two distinct categories of images. We mark category (1) as relevant and those from category (2) as irrelevant. We increase the sample size of the relevant and irrelevant images (double, triple, half, etc). We also test with other combinations. The FEI values computed from different combinations are shown in Table. 4.5. As seen from the table, the (FEI) values indicate same order of importance for all the cases. We consider first 20 images from each iteration and select only those images as (intraset), which are visually very close to the query. The rest are considered (interset).

Adjustment of weights

The marking of relevant and irrelevant images is subjective. One may emphasize more on similarity between one feature (e.g., color) than some other features (e.g., shape),

when comparing two images. Accordingly the weights need to be adjusted. Let the query feature vector be represented as F. The individual components of relevant images are expected to vary within a smaller range say (ϵ). These images may be characterized as,

$$I_r = \{I_j \in I_s : \frac{\delta f_q}{|F|} \leq \epsilon\} \tag{4.13}$$

In the first pass, all features are considered to be equally important. Hence w_1=w_2,...=w_q=1. The feature spaces of the relevant images are therefore altered in a similar fashion after updating the components with w_q. As a result, the ranks of the relevant images are not affected much. For irrelevant images, one feature component may be very close to the query, whereas other feature component may be far away from the query feature. But the magnitude of the similarity vector may close to the relevant ones. These images may be characterized as,

$$I_{ir} = \{I_j \in I_s : \frac{\delta f_{q1}}{|F|} \gg \epsilon \; and \frac{\delta f_{q2}}{|F|} \ll \epsilon\} \tag{4.14}$$

After the features of the query and the stored images (S_d) have been updated with the FEI values, the weighted components are expected to dominate over feature space such that, the rank of the irrelevant ones are pulled down. We have tested the results, from updating the weights with w_q= ${FEI_q}^2$, $\frac{1}{{FEI_q}^2}$ and obtained better results from w_q= ${FEI_q}^2$, in majority of the cases. Intuitively this depends on how the combination of important features dominate over the others for a particular query. As seen from the section 4.3.3, the importance of the individual components is decided automatically from the decreasing order of the (FEI) values. In order to further enhance the effect of the important features, we introduce another multiplying factor t_{qf} to make the component more dominating. The weights of the individual features for successive iterations are expressed as follows ;

$$w_{qf} = t_{qf} \times (FEI_{qf})^2 \tag{4.15}$$

where the value of t_{qf} are chosen as $\{1.0, 0.1\}$. The weights for the query results are shown in Table. 4.6.

Another important issue in evaluating similarity between images, is how to manage the weighted components within the similarity function. Since different components within the feature vector may be of totally different physical quantity, the presence of very large values of some feature components may bias the similarity, in generating the first set of retrieved candidates. One solution is normalization of the features [178], to ensure that individual component receive equal emphasis. The proposed feature set almost represent the same physical quantity (moments). Therefore normalization is not required in computing Euclidean distance between the feature vectors.

4.4 Experimentation

The performance of image retrieval system is tested upon two databases (a) SIMPLIcity images (b) corel 10000 miscellaneous database down loaded from,
(http://bergman.stanford.edu/cgi-bin/$www\_wavesearch$). SIMPLIcity Database has 1000 images from 10 different categories (Africa, Beach, Buildings, Buses, Dinosaurs, Elephants, Flowers, Mountains and Food). Each category is having around 100 images along with some images undergone changes due to (rotation, translation, scaling, noise injection, illumination, etc). Our main objective is to design an efficient CBIR system using simple techniques involving low cost feature extraction mechanism. Moments generate a compact representation with fewer number of features, compared to other sophisticated features. It may at the same time yield poor results if queries are of higher semantic significance. The experiments are performed in the following manner. We started with the proposed feature set as explained in section 4.3.1, and obtained satisfactory results for almost all categories except for few cases. In such cases better results are obtained by computing invariant moments (ϕ^1) directly from the RGB component planes without using (4.7). We designate the feature set as computed from (c_1, c_2, c_3) model using (4.7) as set (A1), and those computed directly from RGB plane as set (B1). Such differences in performance as can be explained from the fact that, the RGB components are sensitive to

varying imaging conditions but have better discriminating power among images. For an unknown database, feature set(B1) becomes a good choice when there is less variation in the imaging conditions or perception. Both the set of features are combined in a hierarchical fashion in some cases, to test the results. The Illumination invariant set (A1) is used first to get the short listed candidates around (100 images) and a second set of retrieval is performed on the short listed candidates using set (B1). Each retrieved set can be further subjected to feature evaluation scheme, based on relevance feedback, for generating still better results.

4.4.1 Performance evaluation

The experimental results are shown from Figs. 4.1 to Figs. 4.18. The results are explained as follows : The signature of the image of Fig. 4.1(a) is shown in Fig. 4.1(b) at $\mu_{d1}(P) \geq 0.8$. Similarly for Fig. 4.2(a), the signatures thresholded at $\mu_{d1}(P) \geq 0.7, 0.8$ respectively, are shown in Fig. 4.2(b), (c). The edge map at $\mu_{d1}(P) \geq 0.6$ and corner signature of Fig. 4.3(a) is shown in Fig. 4.3(b), (c). The query results of SIMPLIcity database are shown in (Figs. 4.4 - Figs. 4.8). The signature is generated at a threshold value of $\mu_{d1}(P) \geq 0.8$. Images are displayed from left to right according to the Euclidean distance, with the top left image as the query image.

The query as shown in Fig. 4.4(a) is of a red flower, which is able to identify similar images undergone illumination changes. The image in Fig. 4.4(b) is from horse category. The precision obtained is high and the retrieved images have similarity in semantics. Fig. 4.5(a) shows the result, when queried with dinosaur. The precision obtained is highly satisfactory for such images having distinct objects. The feature is fairly invariant to linear transformations. We mark the images similar in shape and color from Fig. 4.5(a) as relevant and the rest as irrelevant. The further improvement in precision can be seen from Fig. 4.5(b), from its ability to retrieve blurred, noisy images (at position 4th and 6th from left) after updating the weights calculated from the FEI values. The results when queried with a flower(yellow) is shown in Fig. 4.6(a) as obtained with set(B1). Fig. 4.6(b) is

obtained from set (A1). The precision obtained depends more on user's judgment. Set (A1) has been able to retrieve images undergone illumination change but precision is low. The images of this category have objects with some regularity in shape and background. We mark only those images as relevant from Fig. 4.6(a), which are visually close to the query, both in (color and shape). The improvement after updating the feature weights is observed in Fig. 4.7(a), (b) from retrieving more flowers with its literal color properties. The query in image Fig. 4.8(a) is from category of Africa, which has multi colored and textured regions. The query involves more of semantic reasoning than color and shape. Improvement in results after feature evaluation has been observed in Fig. 4.8(b). As seen from the results of Fig. 4.9(b), the precision did not improve significantly from Fig. 4.9(a). From Fig. 4.9(a), the images having similar color combination are marked relevant. However as seen from Fig. 4.9(b), the semantically similar images have attained a higher rank. From Fig. 4.10(a) it is seen that, the query on bus has retrieved some flowers, from the proposed features. After marking the flowers as irrelevant and emphasizing more on shape signature, better results are seen from Fig. 4.10(b) and 4.10(c). From Fig. 4.11(a), we mark images close to the color of the query, as relevant and arrive at results shown in Figs. 4.11(b), Fig. 4.11(c). As seen from the results, in each iteration the number of intraset images are increased. The results almost converge after second iteration, which is a desirable property. However the initial precision, from which the results are improved in subsequent stages, depends more on the suitability of the proposed feature for a particular query.

The results obtained from database (b) 10000 miscellaneous images are shown in Fig. 4.12-4.13. The results of Fig. 4.12(a) show the retrieved candidates by combining the features in a hierarchical fashion. Improvement in results is shown in Fig. 4.12(b). As the query image of Fig. 4.12(a) is a single object, the shape signature should be more discriminant in this case. To study this, we have set w_1, w_2, w_3 = zero. The result for this observation is shown in Fig. 4.13(a). The result did not vary significantly from that in the

case of Figs. 4.12(a). The result for a query (scene) is shown in Fig. 4.13(b). The results obtained proves to be satisfactory for retrieving scenes.

To evaluate the performance of the proposed scheme, we have randomly chosen queries from each category of SIMPLIcity data set. The average precision value from each category, after retrieving (10, 20, 40) images are shown in Table 4.7. The performance is evaluated in terms of precision rate given by the fraction of correctly retrieved images within the retrieved set. It can be a better choice for an unknown database, because for evaluating the recall rate the number of relevant images of a particular type need to known, which is difficult for a large database. Considering the number of relevant images to be 100 in each category the overall the Precision Recall curve for the SIMPLIcity test database after first iteration is shown in Fig. 4.14

Complexity : We have implemented the algorithm on a SUN Blade system with a 700 MHz Processor. The CPU time for computing the features is on the average 10 secs. The computation required for matching and sorting our results is explained as follows : Let T(n) be the average time to retrieve a query. The complexity involved in computing Euclidean distance is O(ND) where D is the number of components within the feature vector. N is the number of images in the database. The next step involves sorting of (N) images according to the similarity distance. For quick sort the order is of O(NlogN). The time complexity T(n) is represented as, T(n) = O(ND)+ O(NlogN). As the proposed number of features (D=6). Hence $D \ll logN$. Therefore $T(n) \simeq O(NlogN)$.

4.4.2 Performance comparison

There are many CBIR systems reported in the literature, but in most cases the system relies on user defined parameters. In segmentation dependent techniques, the accuracy depends upon the number of classes defined by the user [51], [43]. As a result, it is difficult to judge the performance of a CBIR system unless the parameters match exactly. So a proposed scheme can be compared and evaluated best when the results are tested over the same database. In order to prove the efficiency of our the tech-

Table 4.5: Feature evaluation index

Images(Intra, Inter)	$(FEI_1)^2$	$(FEI_2)^2$	$(FEI_3)^2$	$(FEI_4)^2$	$(FEI_5)^2$	$(FEI_6)^2$
(10,10)	0.207	0.115	0.199	0.351	0.351	0.351
(20,20)	0.207	0.115	0.199	0.351	0.351	0.351
(30,30)	0.220	0.121	0.195	0.350	0.350	0.350
(5,5)	0.309	0.151	0.247	0.349	0.349	0.349
(5,10)	0.289	0.164	0.265	0.306	0.306	0.306
(10,5)	0.433	0.213	0.401	0.534	0.534	0.534

Table 4.6: Weights of the components

Img(Iter)	$t_1(FEI_1)^2$	$t_2\ (FEI_2)^2$	$t_3(FEI_3)^2$	$t_4(FEI_4)^2$	$t_5\ (FEI_5)^2$	$t_6(FEI_6)^2$
Fig.4.5(1)	1×0.64	1×0.32	1×0.38	0.1×0.2	0.1×0.2	0.1×0.2
Fig. 4.5(2)	1×0.62	1×0.52	1×0.51	0.1×0.2	0.1×0.2	0.1×0.2
Fig.4.7(1)	1×0.38	1×0.12	1×0.13	1×0.27	1×0.29	1×0.29
Fig.4.7(2)	1×0.39	1×0.29	1×0.11	1×0.34	1×0.39	1×0.39
Fig.4.8(1)	1×0.35	1×0.34	1×0.37	1×0.24	1×0.23	1×0.23
Fig.4.8(2)	1×0.48	1×0.38	1×0.57	1×0.41	1×0.30	1×0.36
Fig.4.12(1)	1×0.48	1×0.57	1×0.21	1×0.32	1×0.32	1×0.32
Fig.4.9(1)	1×0.25	1×0.41	1×0.27	1×0.51	1×0.50	1×0.50
Fig.4.10(1)	1×0.22	1×0.32	1×0.39	1×0.36	1×0.32	1×0.37
Fig.4.10(2)	1×0.22	1×0.32	1×0.39	0.1×0.36	0.1×0.32	0.1×0.37
Fig.4.11(1)	1×0.35	1×0.39	1×0.31	1×0.41	1×0.45	1×0.44
Fig.4.11(2)	1×0.27	1×0.31	1×0.24	1×0.34	1×0.33	1×0.34

Table 4.7: Average precision % from our algorithm

Category	unweighted features			weighted features (iteration 1)		
	10 images	20 img.	40 img.	10 img.	20 img.	40 img.
Africa	68.20	60.25	55.03	70.50	61.00	58.00
Beach	70.00	55.23	50.60	71.48	56.23	52.58
Building	70.26	60.50	54.46	72.40	63.67	56.00
Bus	80.00	70.59	60.67	81.45	72.77	62.67
Dinosaur	100.0	95.0	90.7	100.0	95.0	92.0
Elephant	83.5	75.5	65.8	85.0	77.0	66.0
Flower	90.0	80.5	70.6	92.0	83.0	71.2
Horses	100.0	90.0	80.5	100.0	95.0	83.0
Mountains	70.5	65.8	60.9	72.0	68.0	62.0
Food	60.8	55.8	53.40	62.0	57.0	55.20

Table 4.8: Comparative evaluation of Weighted average Precision

class	Our method	SIMPLIcity	Histogram based	FIRM
Africa	.45	.48	.30	.47
Beach	.35	.32	.30	.35
Building	.35	.35	.25	.35
Bus	.60	.36	.26	.60
Dinosaur	.95	.95	.90	.95
Elephant	.60	.38	.36	.25
Flower	.65	.42	.40	.65
Horses	.70	.72	.38	.65
Mountains	.40	.35	.25	.30
Food	.40	.38	.20	.48

nique, we benchmark our results with the well known standard image retrieval algorithms namely the FIRM [51], SIMPLIcity [206] and and color histogram matching [176] using the same data set. The online demonstration of their methods are provided at the cite (http//:www.wang.ist.psu.edu/IMAGE). The results are compared in terms of (weighted precision) [206]as defined in (4.16). The weighted average of the precision values within k_1 retrieved images are computed as,

$$\bar{p}(i) = 1/(100) \sum_{k_1=1}^{100} n_k/100 \tag{4.16}$$

k_1= 1,.....,100. n_k is the number of matches within the first k_1 retrieved images. The weighted precision as obtained from each category are shown in Table. 4.8

SIMPLIcity method has reported better than color histogram method. It is difficult to obtain satisfactory results for retrieving from all categories using the same set of features. In most of the cases the average precision for each category can be made better than SIMPLIcity and FIRM via the feature evaluation scheme. Better results are obtained from

SIMPLIcity and FIRM, when the number of class matches with the number of objects. The results are shown in Table. 4.8.

4.4.3 Comparison with MPEG-7

The proposed feature evaluation method, based on user's relevance feedback has been tested on MPEG-7 (http://www.chiariglione.org/mpeg) visual feature standard. MPEG-7 [137] is an ISO/IEC standard, which provides a collection of specific, agreed upon standard descriptors [134], [108] including low level audio or visual features like color, texture, shape and motion etc. The color and texture descriptors of MPEG-7 standard [133] have undergone extensive evaluation in image retrieval. The color descriptors consist of a number of histogram descriptors, a dominant color descriptor, a color layout descriptor. In general, the descriptors were accepted and defined based on detailed studies of their efficiency (i.e., in terms of descriptor size and retrieval accuracy). There are too many dependent dimensions in generic color histogram which include choice of color space, choice of quantization in color space.

The chosen agreed upon standard for color descriptors originating from histogram analysis are the scalable color descriptor (SCD). This descriptor is defined in hue- saturation-value (HSV) with fixed color space quantization. The color structure descriptor (CSD) histogram aims at identifying localized color distribution using a small structuring window, which is constructed in the hue-min-max-difference (HMMD). The dominant color descriptor gives the distribution of the salient colors in the image, where the specification of colors is limited by the color space quantization. The Color layout descriptors (CLD) captures the spatial layout of the dominant colors on a grid superimposed on the region of interest. These descriptors are compact and effective in fast browsing and search applications. In simialrty matching of histograms the (L1 or L2) norms is selected as it usually generates good retrieval accuracy. Textures, like colors is a powerful low-level descriptor for image search and retrieval applications. The commonly chosen texture descriptors are the "texture browsing descriptors" that characterize perceptual attributes such as direc-

tionality, regularity and coarseness of a texture. Homogeneous texture descriptor (HTD) provides a quantitative characterization of homogeneous texture regions for similarity retrieval. The edge histogram descriptor (EHD) captures spatial distribution of edges. The distribution of edges is a good texture signature which is useful in evaluating similarities between images. The similarity matching between EHD descriptors can be done by directly computing L1 norms between two edge histograms. The proposed features evaluation index has been tested using the CSD and EHD features of MPEG-7. The MPEG-7 visual descriptors have been down loaded from (http://www.chiariglione.org/mpeg). The results after relevance feedback has been shown Figs. 4.15 - 4.16. The performance of feature evaluation has been shown on each set of features separately namely CSD and EHD. L2 distance is used for similarity evaluation in the case of CSD and L1 distance is used in the case of EHD. The results of Figs. 4.15 - 4.16 could be explained as follows. The results are shown with an example of a bus from SIMPLIcity database. The bus images generally have distinct object of almost regular shape and in diverse background. The results using CSD of MPEG-7 is shown in Fig. 4.15(a). The result is quite satisfactory from the first pass. Improvement of result after feature evaluation as obtained from our proposed relevance feedback mechanism is shown in Fig. 4.15(b). The result obtained using the same query, that with EHD of MPEG-7 is shown in Fig. 4.16(a). The improvement of results are also obtained after updating the weights based on feature evaluation index as shown in Fig. 4.16(b). In both of the cases the images of busses as obtained from the first pass are considered as intraset images and the rest as interset images.

In order to test the robustness of the proposed CBIR system, we compare the overall performance of our system with MPEG-7 experimentation Model (XM) software [136] which is the frame work of all the reference codes currently being used in several CBIR applications. The software SchemaXM uses MPEG-7 visual descriptors available in XM software, which can serve as a test bed for evaluation and comparison of features in CBIR context. The features of XM commonly used as standard visual content descriptors for still images are listed in Table. 4.9.

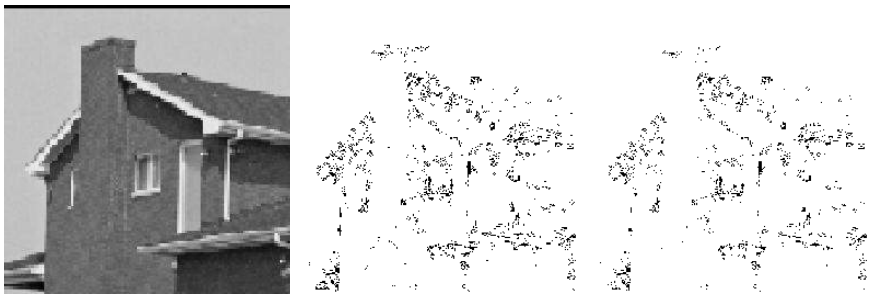
Table 4.9: Standard Visual content descriptors of MPEG-7

Color Descriptors	Texture Descriptors	Shape Descriptors based
Dominant Colors	Edge Histogram	Region Based Shape
Scalable Color	Homogeneous Texture	Contour Based Shape
Color Layout		
Color Structure		

These features have been rigorously tested in the standardization process. Schema provides results on COREL dataset having a good coverage of images from different classes like (beaches, buildings, horses, cars, flowers, etc.) with hundred images from each class. We have seen the query results for all the classes available in Schema which utilize the MPEG-7 visual descriptors. The features used in SchemaXM are promising in modeling low level visual similarity as well as semantic aspects. We provide comparative results on some of the common categories like horses, roses, beaches and vehicles. The performance is compared for the case of evaluating overall similarity between images. The comparison is made against average precision P_r=$(1/n_1)\sum r/n$. For a query, system retrieves r images that belongs to the same class $C1$ from n retrieved images. n_1 is the number images queried from category $C1$. While computing precision, the images from other categories are taken as outliers. The results are shown in Figs. 4.17 and 4.18. Our results can be fairly compared with MPEG-7 visual descriptors as seen from Figs. 4.17 and 4.18. Although the initial precision is found better for SchemaXM as shown in Fig. 4.17(a) and Fig. 4.18(b) but our algorithm fairly catches the results.

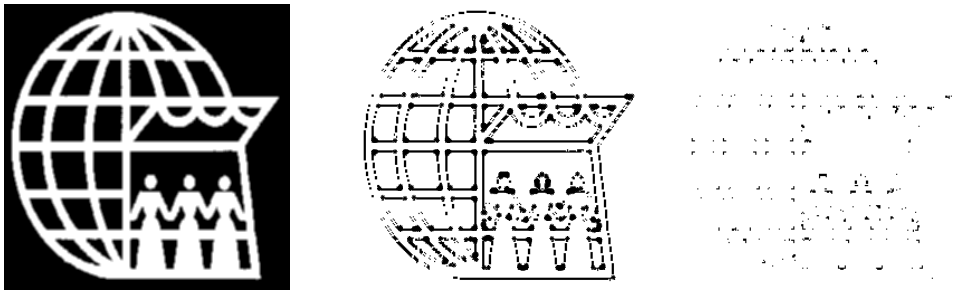
(a) (b)

Figure 4.1: (a) Flower image (b) fuzzy corner signature at, $\mu_{d1}(P)$ $\geq$0.8



(a) (b) (c)

Figure 4.2: (a) House image (b) fuzzy corner signature at, $\mu_{d1}(P)$ $\geq$0.7 (c) $\mu_{d1}(P)$ $\geq$0.8

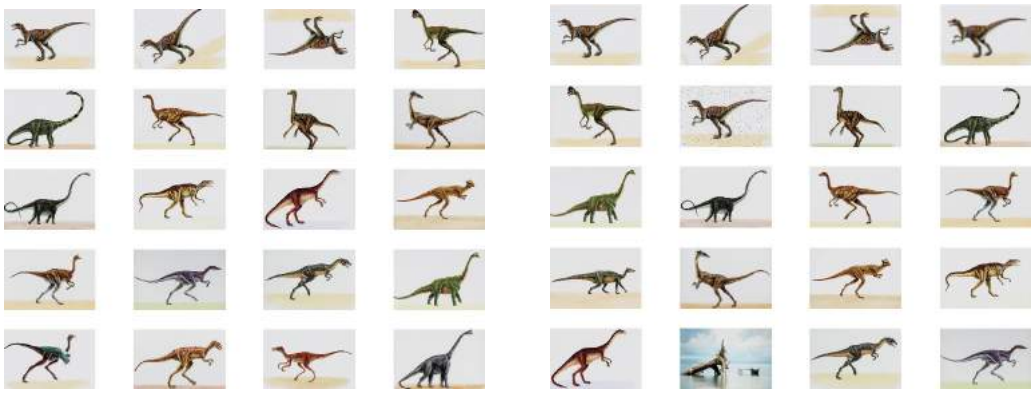


(a) (b) (c)

Figure 4.3: (a) Original image (b) edge map at $\mu_{d1}(P)$ $\geq$0.6, on which the high curvature region marked as (*) (c) corner signature.

(a) (b)

Figure 4.4: Retrieved results using set (A1) (a) flower image (Test for illumination invariance.) (b) horse image, using set (A1). The top left image is the query image.



(a) (b)

Figure 4.5: Retrieved results using set (B1). (a) Test for invariance properties. (b) After feature evaluation with (FEI), the noisy and blurred images are retrieved (iteration 1). The top left image is the query image.

(a) (b)

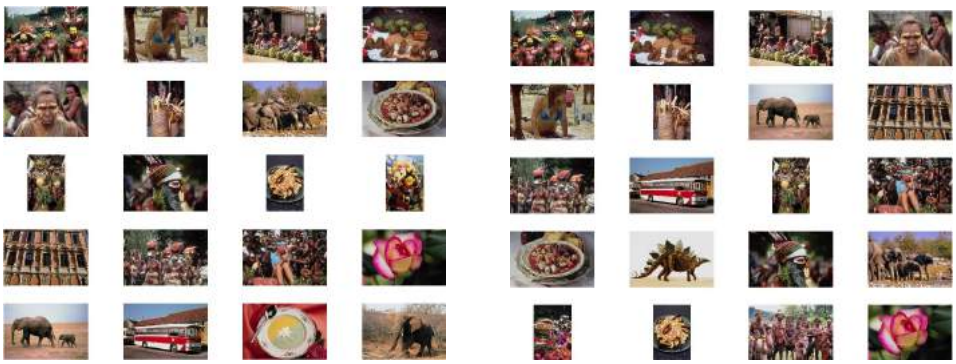
Figure 4.6: Retrieved results (a) first set of candidates, using set (B1) (b) using set (A1). The top left image is the query image.

(a) (b)

Figure 4.7: Retrieved results using set (B1). (a) After feature evaluation with (FEI), (iteration 1) (b) (iteration 2). The top left image is the query image.

(a) (b)

Figure 4.8: Retrieved results combining both set (A1) and (B1). (a) First set of candidates (b) results after feature evaluation with (FEI) on set (B1) (iteration 1). The top left image is the query image.



(a) (b)

Figure 4.9: Retrieved results using set (A1). (a) First set of candidates. (b) after feature evaluation with (FEI), (iteration 1). The top left image is the query image.

(a)

(b)

(c)

Figure 4.10: Retrieved results using set (A1). (a) First set of candidates (b) after feature evaluation with (FEI), (iteration 1) (c) (iteration 2). The top left image is the query image.

(a)

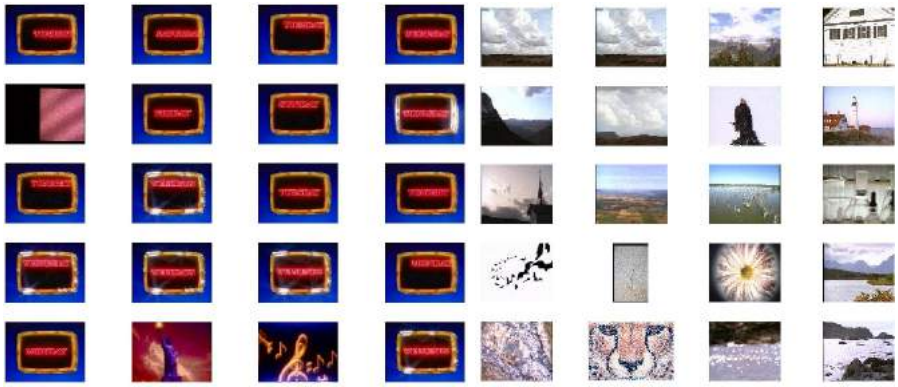
(b)

(c)

Figure 4.11: Retrieved results using set (A1). (a) First set of candidates (b) after feature evaluation with (FEI), (iteration 1) (c) (iteration 2). The top left image is the query image.

(a) (b)

Figure 4.12: Retrieved results from miscellaneous database, combining both set (A1) and (B1) (a) First set of candidates. (b) results after feature evaluation with (FEI) on set (B1) (iteration 1). The top left image is the query image.



(a) (b)

Figure 4.13: Retrieved results from miscellaneous database, combining both set (A1) and (B1). (a) Taking contribution from shape signature only (i.e., w_1=w_2=w_3=0.0) (b) query results on a natural scene. The top left image is the query image.

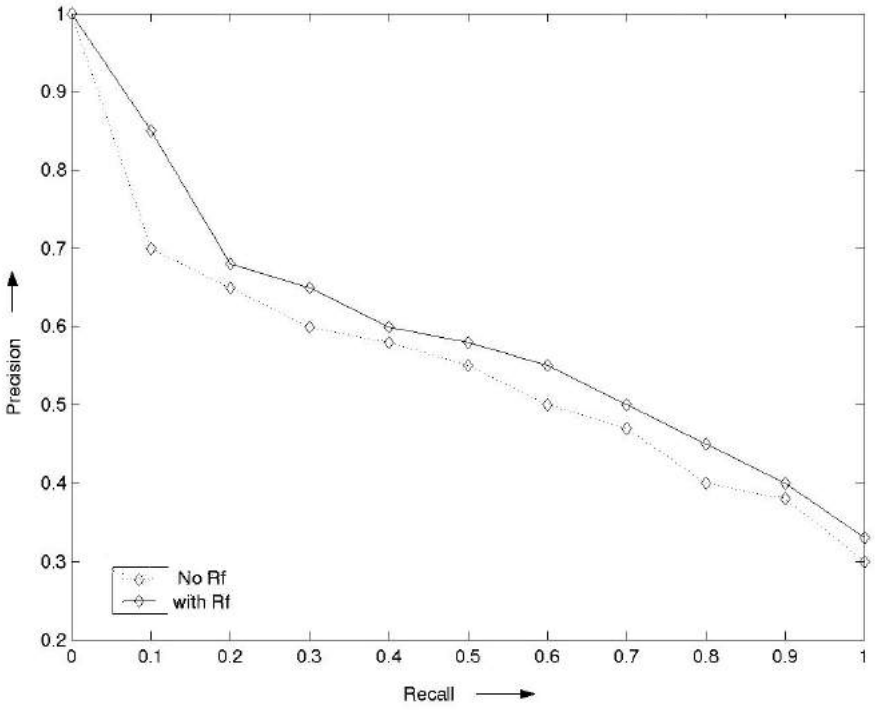
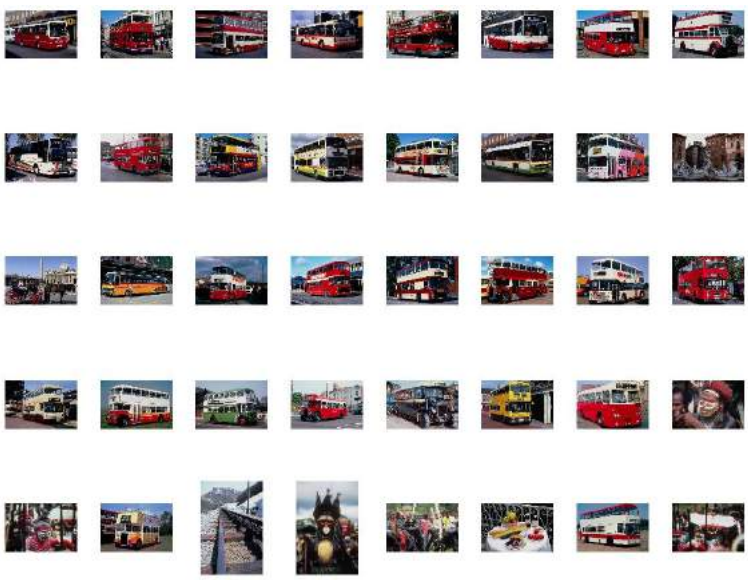
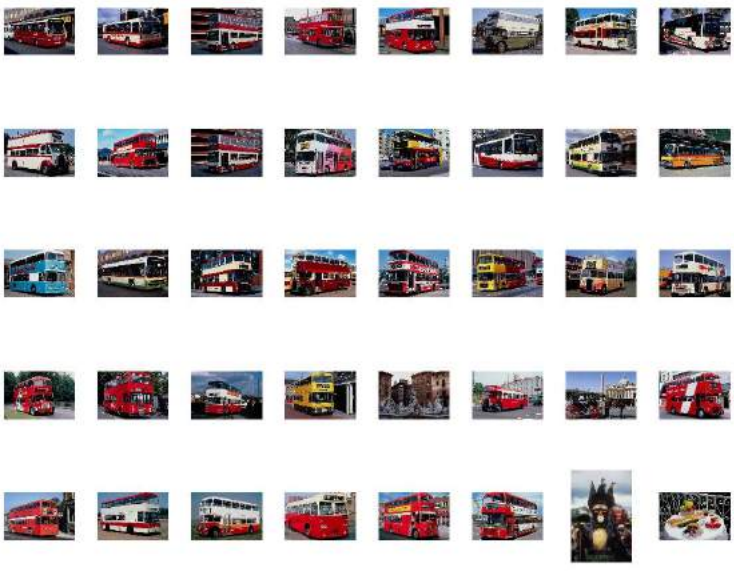


Figure 4.14: Precision, Recall curve of the SIMPLicity database.

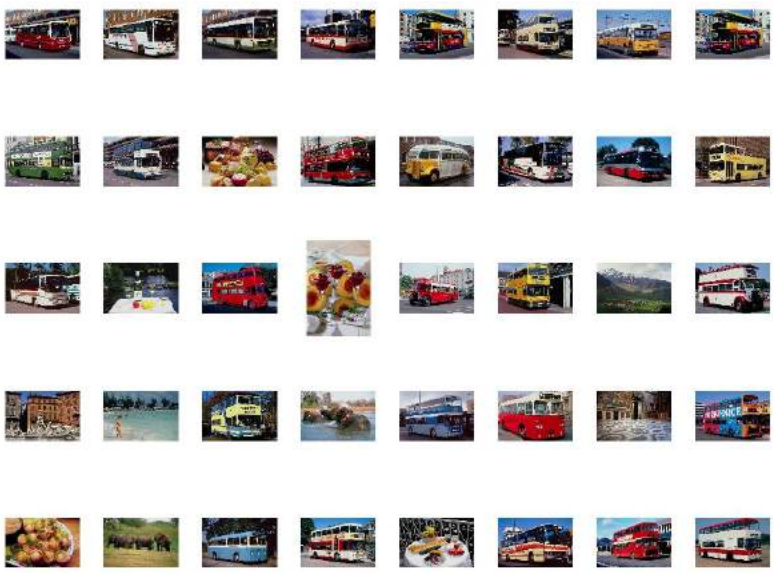


(a)



(b)

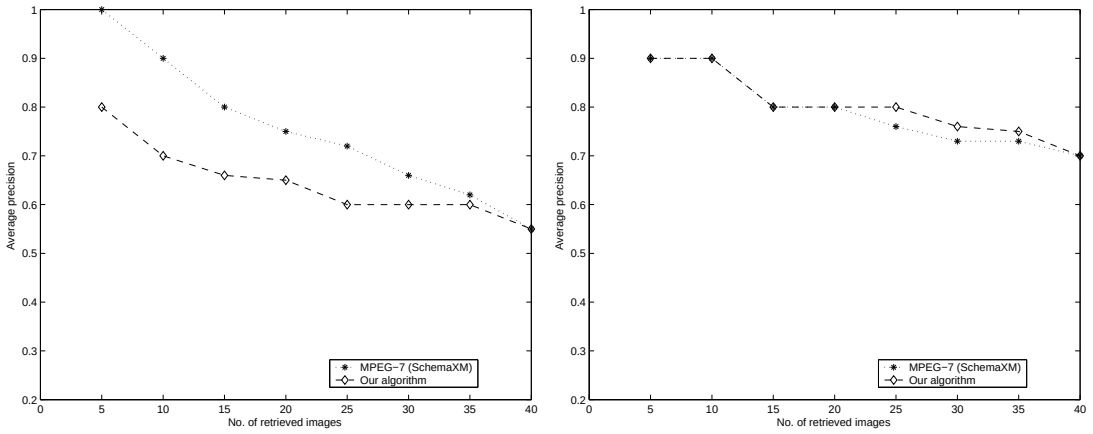
Figure 4.15: The proposed relevance feedback scheme on MPEG-7 visual descriptors (a) CSD descriptors (b) results after feature evaluation (iteration1)



(a)

(b)

Figure 4.16: The proposed relevance feedback scheme on MPEG-7 visual descriptors (a) EHD descriptors (b) results after feature evaluation (iteration1)

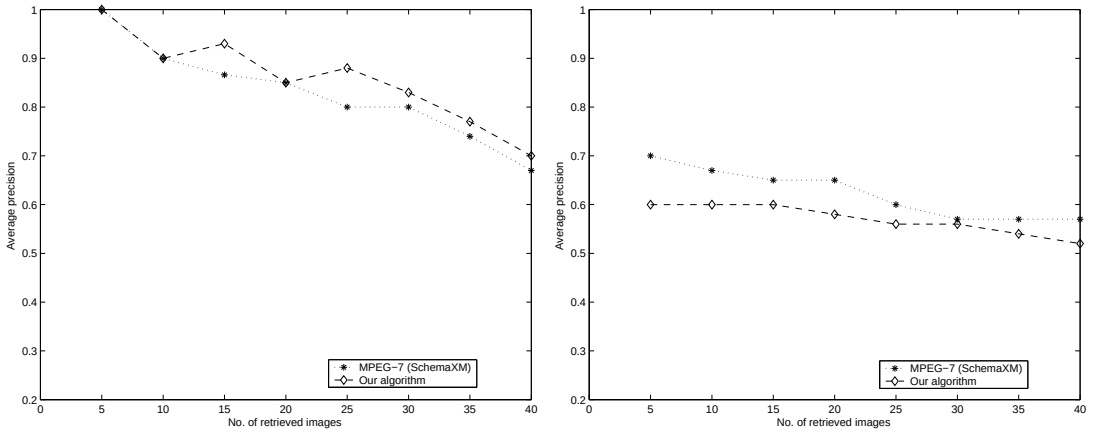


(a) (b)

Figure 4.17: Comparison with MPEG-7 visual descriptors (a) Category beaches (b) Flowers.

4.5 Conclusion

The studies in this chapter, mainly emphasize on the feature extraction, feature ranking and feature evaluation aspects of an image retrieval system. The suitability of a proposed feature set can be best understood from its ability to discriminate images of different types. This particular aspect has been studied from, selecting relevant and non redundant features using the mutual information based minimum redundancy-maximum relevance criterion. Such a ranking scheme on the features of a smaller database consisting of images of different classes of images, can be useful for handling large database for improving the accuracy. From different query examples and classification results using different training set on a standard database, it is shown that a single set of features may not be suitable for handling all types of queries. The set of features which may be able to retrieve images having a range of illumination variations, may fail to retrieve noisy images. To overcome such limitations, the features are combined and the weights are updated to obtain a trade-off between feature selection and accuracy, in the case of retrieving from an unknown database. Using the selected features, a feature evaluation scheme has been proposed using user's relevance feedback mechanism. The relative importance of the fea-



(a) (b)

Figure 4.18: Comparison with MPEG-7 visual descriptors (a) Category horse (b) vehicles.

tures is evaluated from both the relevant and irrelevant images as indicated by the users, and used for iterative improvement of results. The proposed method is quite effective, in evaluating the features from a small number of feedback images.

We intend to test the effectiveness of the system, using simple features involving low computational cost. Invariant moments computed from the cluster of visually significant points are used to represent an image with minimum number of points. The proposed system can identify relevant images undergone transformations, like translation, rotation, scaling, illumination change, etc. To reduce the computation, a S-type function has been used which reduces the number of insignificant edge candidates for corner testing used in the experiment. With such a choice some of the insignificant noisy points have been reduced, but it is not robust to noise corruption. Noise removal could be achieved through a post processing and generating an refined edge map, as it was in the case of using the exponential function described in Chapter2. The algorithm has been implemented on a SUN Blade system with a 700 MHz Processor. The CPU time for computing the features is on the average 10 secs., using Matlab. The computation required for matching and sorting the results is of the order of O(NlogN), N is the number of images in the database. Such matching is independent to the number of classes, in contrast to most of segmentation

based techniques, which depend on number of classes [51]. The algorithm may be further refined to make is noise tolerant and at the same time reduce the computation time. The suitability of the proposed features can be increased if the application may be extended to specific images like logos or facial images, other than general purpose images only. We also try to introduce other sophisticated feature and test its suitability for MPEG-7 in order to extend the algorithm for video domain to achieve a better goal.

The aim of an ideal CBIR system is to mimic the human visual system. However, the success achieved is far from the goal. To enhance the retrieval accuracy, better characterization of images is necessary. One such way could be to include more important features like texture. From semantic point of view, if we consider the spatial organization among the objects, in a multi object scene, then the description would be more meaningful and expected to generate better results for CBIR. The properties of the coherent regions and the spatial relations between the regions may depict, much significant information contained within an image. These, require suitable segmentation algorithms to extract properties like, color texture, shape information of individual regions. In the next chapter, we address the problem of segmenting images into homogeneous colored textured regions, with which spatial relationships are combined to generate better characterization of an image.

Chapter 5

Image Retrieval Using Wavelet Based Segmented Regions and their Fuzzy Spatial Relations

5.1 Introduction

This chapter deals with retrieval of color images, based on regional properties. An algorithm is proposed to segment an image into fuzzy regions based on the coefficients of multiscale wavelet packet transform. Fuzzy Spatial relations computed between the segmented regions are combined with the color and texture properties of the segmented regions, to provide overall characterization of an image. Natural images constitute mosaics of homogeneous regions, where the features like color, texture and shape play a significant role in depicting the salient aspects of an image. Although a wide variety of approaches of image retrieval have been suggested based on these features, but the success achieved is limited using these features only. Serious efforts are being made by the researchers, for developing suitable image characterization methods which are quite capable of capturing semantic details, derived from image features [206], [51], [203], [117]. A general way to cope up with this problem is to incorporate spatial relations between the

regions. The properties of individual regions if coupled with spatial relationships between them, may generate more realistic matching between two images.

In such an approach, segmenting images into homogeneous regions considering the combination of color and textural properties is essential. Segmentation based techniques can better handle, querying on particular objects i.e., finding coherent image regions which roughly correspond to objects. For example, such method has been popular and used in the Bloworld system [43]. Querying in Blobworld, is based on the attributes of one or two regions of interest, rather than a description of the entire image. However, in segmentation based methods it is difficult to decide the number of classes, for different images of a database. Spatial relationships are also dependent on the segmented regions. To handle such problems in the case of overall matching between images, image retrieval scheme based on visually significant points has been explored in the previous chapter. However, the aim of CBIR research is to develop better image characterization methods which can still generate better results, close the human visual perception system. Color and texture properties of individual coherent regions are very important [88]. Also, a significant aspect of discriminating among images, depends on the spatial locations and relationship of objects or regions. Looking into these facts, we are interested to take up such a problem. The proposed methodology has two parts. In the first part, an algorithm to segment an image into fuzzy regions based on the coefficients of multiscale wavelet packet transform is proposed. In the second part, Fuzzy Topological relationships are computed between the regions. The color and texture properties as indicated by centroids and spatial relations between the segmented regions are used together to provide overall characterization of an image. The closeness between two images are estimated from these properties.

Segmentation into homogeneous regions has been made with the use of wavelet packets. Such a choice has been made because, wavelet analysis provides an effective multiresolution representation of a signal/image, in terms of its constituent energies. It is capable of capturing the salient aspects of a data, with a small set of co-efficients, which

spans both in spatial and frequency domains. In the recent times, effective characterization of region based properties, have been made with the use of wavelets and proved to be promising [144], [3], [4]. It has also become a good alternative to Gabor filters [130]. Significant texture information requires over complete decomposition [4]. With the use of wavelet packet transform, a tree structured extension of the dyadic 2-D discrete wavelet transform may be obtained. Thus a finer and adjustable resolution at high frequencies is allowed as compared to the dyadic wavelet transform. One important point is that, multiresolution representations facilitate representation of data at different levels of detail. However, not all the bases are equally important. Hence finding out computationally efficient optimal basis based on the some statistical criteria becomes important [130]. This problem has been taken care of, by selecting best bases for further decomposition, based on a entropy measure.

Many authors have also proposed different methods to quantify spatial relations between regions. Some of them are discussed in brief.

Methodologies to compute spatial relationships of the objects have been proposed in [66], [152]. However, the methods mostly deal with the geometric attributes of a region like (area, shape adjacency, surroundedness, etc.). Methods used to compute approximate spatial relations like left to, above, below, etc. have been proposed in [102]. A new histogram representation called R-Histogram, which extends histogram of angles by incorporating both angles and labeled distances is proposed in [208]. This representation considers Topological spatial relations between region like ”inside” , ”overlap” for image retrieval applications. In Visual Seek, [189], color set selection and back-projection method is used for automatic extraction of salient color regions from the images. The spatial location and size are added to the region relation and subsequently used for queries.

Symbolic descriptions of segmented regions have become popular, where comparisons are based upon 2-D strings [47], [116]. However, the methods cannot always accommodate similarity of the symbols, based on the visual features. Quad trees as proposed in [211], [167] and graph matching methods as in [51], [103] are used for efficient spatial

matching.

Fuzzy logic has also been used, to provide solutions for quantitative representations of spatial relations and similarity matching.

Fuzzy similarity measure between regions has been introduced in FIRM [51], where an image is represented by a set of segmented regions reflecting color, texture and shape properties. Properties of all regions are integrated by a family of fuzzy features. However, geometric relationships between the regions are not considered.

In FIRST, [103] a Fuzzy Attributed Relational Graphs (FARGs) has been used to represent images, where each node in the graph represents an image region (e.g., blueness, size and texturedness) and each edge represents a relation between two regions with attributes (e.g., spatial relation, adjacency, etc.). The given query is converted to a FARG, and a low-complexity fuzzy graph matching algorithm is used to compare the query graph with the FARGs in the database.

An new approach, in which fuzzy interrelations between the segmented regions are used to capture Topological spatial relations between the regions is proposed in [63]. Fuzzy spatial relations are mapped through building a matrix, representing the truth values between [0,1]. In the proposed method, the matrix is used to express the graduality between the positional relationships only.

In the next section, the segmentation algorithm and the Spatial Relation feature extraction process used in the proposed methodology is described. The Experimental results and comparative studies are made in section 5.3. The chapter is concluded in section 5.4.

5.2 The Proposed methodology

The simple technique to identify uniform color textured regions has been proposed with the use of wavelet packets. Approximate boundary regions are also extracted for detecting the differences of textures in adjacent regions. Fuzzy C- means algorithm has been used for assigning multiclass membership values to each pixel, and the membership values

are subsequently used for approximate segmentation into homogeneous regions. Wavelet packet frames have been used, to extract features for segmentation, and the best basis are obtained based on entropy measure.

The methodology comprises of the following steps, which are mentioned below. The centroids of the segmented regions depict the color and textural properties. To provide meaningful characterization between the fuzzily partitioned regions, two Topological measures namely, Index of fuzziness [155] and Shape distance [48] are computed between the fuzzy regions of the segmented image. The performance of the proposed algorithm in the case of retrieving perceptually similar images from benchmark databases is compared with other approaches like, (a) Color, texture histogram similar to [151] (b) Gabor texture features [121], to prove the efficiency of the proposed scheme.

5.2.1 Multiresolution Analysis using wavelet packets

An arbitrary function $f(x)$ can be represented with linear combination of a set of wavelets or basis functions. These basis functions are obtained from a single prototype wavelet called mother wavelet by dilations(scale) and translations(shift). Wavelet transformation can be used to analyze the space and frequency content of an image. Mallat, has shown that a signal at any resolution, can be decomposed into a signal at a coarser resolution, plus the associated wavelet components. The multiresolution characteristics of wavelet transform is very attractive to different image processing and vision applications [130].

The discrete normalized scaling and wavelet basis functions for 2D-DWT are defined as [130],

$$\begin{aligned} \phi_{i,k}(l) &= 2^{i/2} h_i(2^i l - k) \\ \psi_{i,k}(l) &= 2^{i/2} g_i(2^i l - k) \end{aligned} \tag{5.1}$$

where i and k are the dilation and translation parameters and h_i and g_i are respectively the sequence of lowpass and bandpass filters of increasing width indexed by i. (i, $k \in Z$) a positive set of integers. The full discrete wavelet expansion of a signal that form an

orthonormal basis for $L^2(R^2)$ is given as,

$$x(l) = \sum_{k \in z} c_{(d_0)}(k)\phi_{d_0,k} + \sum_{i=1}^{d_0} \sum_{k \in z} d_{(i)}(k)\psi_{i,k} \tag{5.2}$$

$d_{(i)}$ are the wavelet coefficients and $c_{(d_0)}$ are the expansion coefficients of the coarser signal approximation $x_{(d_0)}$. The $c_{(d_0)}$ and $d_{(i)}$ can be interpreted in terms of simple filtering and down sampling operations. The 2-D DWT is computed by applying a separable filter bank to the image where $c_i(x, y)$ corresponds to low frequencies (LL), $d_i^1(x, y)$ corresponds to the vertical high frequencies (horizontal edges, LH), $d_i^2(x, y)$ horizontal high frequencies(vertical edges, HL), $d_i^3(x, y)$ the high frequencies in both directions (the corners, HH). The image $I(x, y)$ is represented at several scales by $\{$ cd_0, $d_i^n(x, y)$, n= 1,2,3, i=1,..., d_0 $\}$. The 2D-DWT transform of an $2^i \times 2^i$ image is represented as a set of shifted and dilated wavelet function resulting into subband images, each of size $2^{i-1} \times 2^{i-1}$.

The standard dyadic transform are not suitable for analysis of high frequency signals with relatively narrow band because the subband decomposition is applied on the low pass component plane of the input image. The choice of wavelet scale parameter within a DWT is 2^i, where $i \in \{0, 1, 2....\}$. Such choices are not suitable for analysis of high frequency signals with relatively narrow band.

Texture classification and segmentation require a finer frequency analysis. To obtain efficient texture properties, multiscale overcomplete packet transform is necessary. Wavelet packets comprises all possible combinations of subband decomposition applied recursively to the lowpass and high pass filter results of the previous wavelet transform step. As a result, it is possible to create arbitrary tree structures that generates various combinations of orthonormal bases [169]. Better frequency localization, while retaining the structure of a discrete decomposition can be obtained through wavelet packet decomposition.

As an example, a 2-level decomposition on the orthogonal subspaces generates bases, $P_{2,i,k}(l)$, i= 0,1,2,3, $p_{1,1,k}(l)$=1/2$\psi(l/2-k)$, $p_{2,0,k}(l)$=1/4$\phi(l/4-k)$, $p_{2,1,k}(l)$=1/4$\psi(l/4-k)$ forms orthogonal sets.

An appropriate way to perform the wavelet transform for texture feature extraction is to detect the most significant frequency channels and then decompose them further. In order to avoid the full decomposition, an adaptive decomposition algorithm using a maximum criteria of textural measures based on the statistics extracted from each of the subbands, can be used to identify the most significant subbands. This enables to select the desired frequency subspaces for further decomposition. A signal can be adaptively decomposed using a selected bases according to a selection condition, based on e.g., (energy, entropy of the subbands). The significant nodes are computed and the optimal basis is obtained based on entropy minimized tree, where the entropy function is computed between a parent and quad child. The entropy minimized tree forms a nested sequence of subspaces, $V_1 \subset V_2 \subset ...V_n$. A detailed multiresolution analysis using wavelet packets is explained in the Appendix . The implementation of wavelet packet transform has been explained in Fig. 5.2.

The detailed explanation of the feature extraction process is explained below.
The wavelet packet co-efficients are extracted from the independent color channels, from which the local spatial frequency characteristics can be computed to capture texture characteristics.

5.2.2 Integrating wavelet features for extraction of colored textures

Features extracted from different frequency bands of wavelet coefficients are shown to be effective for representing texture properties [4]. An input color image can be looked as a 3-D energy function $E(x, y, \lambda)$ where (x,y) denotes the spatial co-ordinates and λ denotes the wavelength of light energy. The local spatial frequency characteristics of $E(u, v, \lambda)$ can be used to capture texture characteristics. The properties can be combined in the wavelet domain by convolution with a wavelet filter in the spatial domain with the independent color channels [88]. The building block of feature extraction is shown in Fig. 5.1.

$$\hat{M}(u,v,\lambda) = h(u,v) * \int E(u,v,\lambda)\, d\lambda \tag{5.3}$$

The filter bank $h(u,v)$ is a set of band pass filters with frequency selective properties. The original image is available in RGB format. The RGB values encode the color information. Each of the individual components are transformed separately to obtain the wavelet packet coefficients.

5.2.3 Multiresolution feature extraction

In order to select the bases adaptively, entropy of each subbabnd decomposition of the wavelet packet tree is evaluated. The nodes marked as the part of best basis are decomposed further. As a result, specific trees can be created by only splitting each subband according to a defined splitting condition.

The best basis are selected on computation of entropy of the subbands. If E_1 is the entropy of a particular node, E_2 the sum of entropy of the children, of the node. $E_1 \leq E_2$ then the nodes are marked as the part of best basis set. The entropy here is the Shannon entropy. The analysis is performed upto second level of decomposition and this results in a set of wavelet packet bases.

Having obtained the entropy minimized tree, multiresolution analysis is performed considering the spatial similarities across subbands. In general, there are spatial similarities across subbands [169]. The pixels in each subband are linked to the pixels of the adjacent subband at the next lower level. Each pixel (x, y) from the former set of subbands acts as the root of pixels $(2x, 2y)$, $(2x + 1, 2y)$, $(2x, 2y + 1)$, $(2x + 1, 2y + 1)$. For a (n) level decomposition the same rule links the pixels of the adjacent subbands starting from LHn, HLn, HHn, respectively. The selected bases V_n are interpolated by factor (n), to ensure the identity to the sampling spaces. This generates a transform such that, every subband is having the same number of coefficients as the number of samples in the original signal. It also leads to an overcomplete representation. This approach is computationally effective and also helps in reducing blockyness arising in case of the block based methods and the

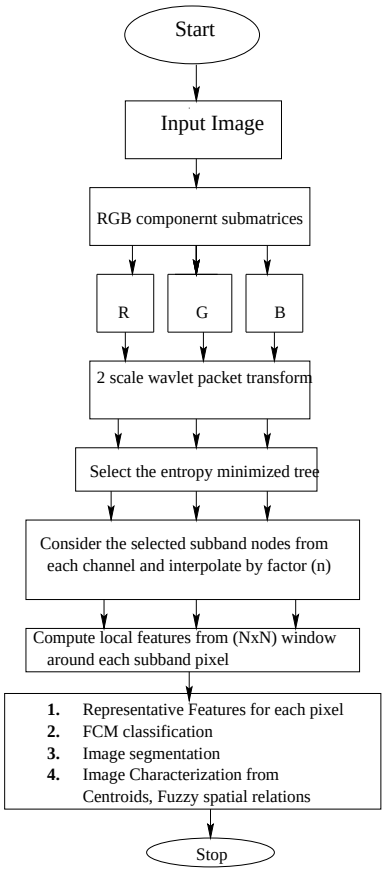


Figure 5.1: Building block of feature extraction.

cost involved in pixel wise segmentation. The local features of the filter output (wavelet domain) around each (x,y)th pixel is estimated from the selected bases.

One of the simplest approaches for describing texture is to use moments (e.g. moments of gray level histogram) of an image or region. The second moment (variance) is of particular interest in texture description [75], which can be used to establish descriptors of relative smoothness. The third moment is a measure of the skewness of the histogram, while the fourth moment is a measure of relative flatness. The moments of the wavelet coefficient of various frequency bands have also proven effective for discriminating textures [201].

In the present approach, moments have been computed on a local window, around each pixel, from each of the RGB channel, which are clustered by applying the fuzzy C-means algorithm to obtain the fuzzy partitions. To obtain the moments a Daubechies-4 wavelet transform is applied. Around the (x,y) pixels of a certain subband a $W \times W$ window is centered. The raw moments m00, m10, m01, m20, m02, m11 are computed. From which the centralized moments $\mu 20$, $\mu 02$, $\mu 11$ are computed on each window, [75]. The detailed computation of moments has already been discussed in chapter 4.

These features generated quite satisfactory segmented results, when tested upon different type of color images.

For a color input image, the moments computed along the individual channels are represented as, and $F_k\lambda$ =[$\mu 20$, $\mu 02$, $\mu 11$] The feature computed along the three channels are considered as linearly independent and represented as,

$$F_k(x, y) = \sum F_{k\lambda} \tag{5.4}$$

The segmentation obtained with the moments are shown in Figs. 5.3(b), 5.4(b), 5.5(b), 5.6(b).

Apart from $\mu 20$, $\mu 02$, $\mu 11$, the variance of the coefficients of selected the window is computed from the (R) plane. By classifying these features we approximately segment the boundaries between the homogeneous region as shown in Figs. 5.3(c), 5.4(c), 5.7(b), 5.8(c). The feature used to obtain the boundaries are represented in (5.5).

$$F_k(x,y)var = (\sum_{x=1}^{W}\sum_{y=1}^{W} w_k(x,y) - \bar{F}_k(x,y))^2 \tag{5.5}$$

where $\bar{F}_k(x,y)$ is the local mean and $w_k(x,y)$ is the coefficients at different scales k. The size of the window is an important parameter. Accurate texture measurement demands larger window size and better localization of region boundaries require smaller window size. A 5×5 window has been selected for the proposed feature extraction.

The extracted features are integrated in the feature space to produce approximate segmentation into homogeneous regions. The features need to be clustered into different categories.

We have used the standard fuzzy C- means algorithm [28] for pattern classification. Let X= $\{x_1, x_2, ..,x_N\}$ be a finite subset of n dimensional vector space R^n where x_i is a feature vector for a pixel in the image. For an integer C,$2 \leq C \leq N$. A $C \times N$ matrix U=[u_{ik}] is called the fuzzy C partition of X whenever the entries U satisfies the following conditions. $u_{ij} \in [0,1]$ for all i and j, u_{ij} is the degree of membership of x_i in the cluster j, $\sum_{i=1}^{C} u_{ij}$ =1 for all j and $0 \leq \sum_{i=1}^{N} u_{ij} \leq N$ for all i.

5.2.4 Attributes of fuzzy sets on segmented image

Index of fuzziness : From fuzzy C-means clustering of $F_k(x,y)$, or$F_k(x,y)var$ a fuzzy partition matrix U of size C by N is obtained, where C is number of clusters and N is number of data points. The feature vector usually belongs to multiple regions with different degrees of memberships as opposed to classical region representation in which the feature vector belongs to exactly one region. The measure of fuzziness between the C fuzzy partitions of X, as obtained from the fuzzy C means clustering is considered. Let the fuzzy partitions (u_1, u_2, .. u_C) on X be defined by the set of C menbership functions, $\mu u_j(x)$ for all $x \in X$ and j = 1, 2, ...C. The intersection between two fuzzy partitions u_m and u_j , ($m \neq j$) from the C collections is defined from [155] and gives the measure of Index of Fuzziness.

$$I(u_j \cap u_m) = \frac{1}{|X|} \sum_{x \in X} [min(\mu u_j(x), \mu u_m(x))] \tag{5.6}$$

where $|X|$ denotes the cardinality i.e., the number of support elements in $|X|$.
Computation of the fuzzy Spatial Relation properties are explained as follows :
Shape distance : The pixels of the fuzzy partition are assigned distinct class label for which the membership is maximum. As a result, we obtain the segmented image. Each segmented region can be looked upon a bounded fuzzy set where the membership is zero outside the region. We consider the shape properties of the fuzzy regions of the segmented image, as proposed in [48],

If $\mu(x, y)$ is the membership of the fuzzy set μ at the point (x, y) in R^2. The center of gravity of μ is the point (x_0, y_0) where $x_0 = \frac{\int_s x\mu(x,y)\,ds}{A(S)}$ and $y_0 = \frac{\int_s y\mu(x,y)\,ds}{A(S)}$ where A(S) = $\int_s u(x, y)ds$. The dissimilarity between two fuzzy sets u and v is defined as,

$$D_1(u, v) = \int_s |u - v|\,ds \tag{5.7}$$

A fuzzy set v is a rotation of a fuzzy set u, if there exits an angle $\alpha \in [0, 2\pi]$ so that any point v may be represented as,

$$v(x, y) = u(xcos\alpha - ysin\alpha, xsin\alpha + ycos\alpha) \tag{5.8}$$

The shape distance between two fuzzy sets u and v (here u and v corresponds to individual segmented regions) are defined as,

$$D_3 = min[D_1(u, v_{\alpha 1-\alpha 2}), D_1(u, v_{\alpha 1-\alpha 2+\pi})] \tag{5.9}$$

where $\alpha 1$ and $\alpha 2$ are the orientations defined as the angles that the major axis of u and v make with the x axis. $D_1(u, v_\alpha)$ is the dissimilarity with respect to rotation of v by angle α. The major and minor axes are perpendicular to each other and passes through the center of gravity. D_3 is the smaller of the D_1 values in these two orientations. The orientation θ of a fuzzy set can be obtained from the relation as follows,

$$tan2\theta = \frac{2\int_s (x-x_0)(y-y_0)\mu(x,y)\,ds}{\int_s (x-x_0)^2\mu(x,y)\,ds - \int_s (y-y_0)^2\mu(x,y)ds} \tag{5.10}$$

It can be proved that D_3 is a metric. This metric can be used in pattern recognition and matching problems in a fuzzy framework. We have not considered the relations like left to, above, etc. as the segmented regions may not represent the image at object levels.

Feature Representation : The feature representation of the segmented image can be looked as a collection of features [F_1, F_2, ...F_C] representing the variance of the centroids where C represents the number of classes. The index of fuzziness I_{mj} for $m \neq j$ between any two partitions of the C separable classes are considered. Also the shape distance S_{m1m2} between any two segmented fuzzy regions (m1, m2) are computed to generate the effective characterization of an image. In order to compute feature similarities between two images, Euclidean distance metric is used to compute the dissimilarity value between the centroids. If $X = [x_1, x_2, \cdots, x_i]$ and $Y = [y_1, y_2, \cdots, y_i]$ are two feature vectors then the Euclidean distance metric between X,Y is given by,

$$E_d = \sqrt{\sum (x_i - y_i)^2} \tag{5.11}$$

Similarity Evaluation : The intuitive similarity between two regions are computed from shape distance. The rank is computed as follows. The similarity is performed on each type of features separately. Finally the rank obtained from each individual set of features are combined to get the final result. The rank obtained from matching the centroids and index of fuzziness features are decided by measuring Euclidean distance between the features of the query and the target images. The total similarity is calculated as follows :

$$d_{qt} = \sum_{k=1}^{n} (r_{tk})/n \tag{5.12}$$

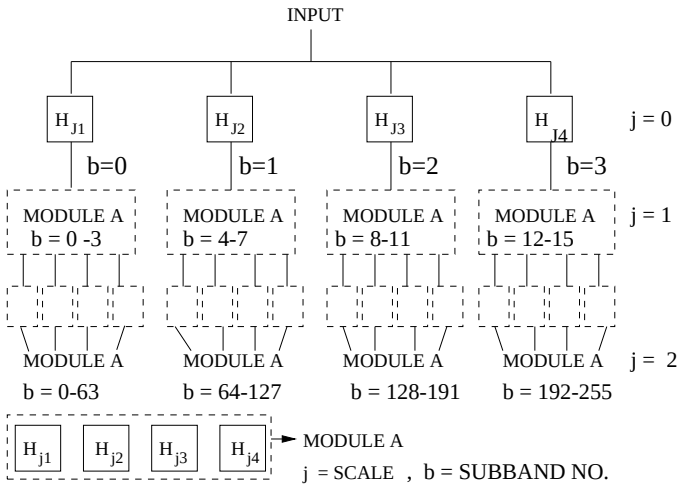


Figure 5.2: General tree structure of discrete wavelet packet frame (M) band decomposition. Down sampling by a factor of 2 is required for standard wavelet packets.

where d_{qt} is the similarity distance of the query image q to image t. n is the total number of features r_{tk} is the rank of the image t in feature k.

5.3 Experimental steps

We test our algorithm on a database consisting of 1000 images with 100 images from ten different categories down loaded http://wang.ist.psu.edu/docs/related/. The experimental results are shown from Figs. 5.3 to Fig. 5.13. The number of distinct regions in natural images vary and also depend upon the individual perception. The segmentation results greatly depend upon the number of specified classes. We grossly partition the image into three meaningful classes to analyze the region attributes and spatial relations between these regions. Fixing the number of classes may not effectively partition all database images at individual object level. It is expected that similar partitions will be generated for images with similar semantics for use in computing overall image to image similarity. If the segmentation becomes finer, the uncertainty in characterizing the perceptual contents increases due to detailed classifications within regions. This may degrade the overall

impression of the object nature.

The segmentation results are shown from Figs. 5.3 to Fig. 5.8. The image shown as Fig. 5.3(b) is the segmented output of Fig.5.3(a) into three regions. Individual regions are characterized with different gray values. Fig.5.3(c) is the segmented output which broadly classifies the boundary between the regions using the features obtained from (5.5). The regions boundaries are assigned a common class label (plotted as dark boundary pixels). The segmented output of Fig. 5.4(a) is shown in Fig. 5.4(b) using the centralized moments $\mu 20$, $\mu 02$, $\mu 11$, as explained in Section 5.2.3. The classified regions boundaries plotted as dark pixels are shown in Fig. 5.4(c). Similarly the segmented output for Fig. 5.5(a) is shown in Fig. 5.5(b) with number of classes equal to three. The segmented output of Fig. 5.6(a) with the number of classes equal to three is shown in Fig. 5.6(b). In this case with the assigned number classes equal to three, the regions are not able to represent the individual objects. However the interrelation computed between the regions, may capture discriminating information for images of similar type. The region boundaries as obtained from classification of the image of Fig. 5.7(a) with three number of classes, are shown in Fig. 5.7(b). The segmentation results as shown in Fig. 5.8(b), (c) for Fig. 5.8(a) are quite satisfactory in classifying the images into the assigned regions. By intelligent selection of the basis, the number of features have been reduced for the desired segmentation. In our problem, the description of the entire image is used in querying rather than taking the attribute of individual regions. We have taken query examples from all categories of images and tested the performance with the proposed features. We try two explain the results obtained on images with some distinctive properties. The query image in the case of Fig. 5.9 is of a horse in green background. Fig. 5.9(a) represent the results obtained by matching the centroids of the segmented regions (three regions). Fig. 5.9(b) shows the results after taking into the account, the spatial relation features. As observed from Fig. 5.9(b), the results did not improve significantly. This may be explained as follows. In this case, the object of interest is horse which may or may not be in similar background. As the color and texture properties of significant regions are quite similar, the properties

(color and texture) extracted from the significant regions serves quite satisfactorily to retrieve images from the same category. Here, the intersection between the fuzzy partitions could not reflect discriminating spatial properties because the regions are not always representing the individual objects.

The query as shown in Fig. 5.10 is of flower. The retrieval results as obtained considering the centroids are shown in Fig. 5.10(a) and the results obtained after considering the spatial properties are shown in Fig. 5.10(b). An improvement of result is obtained, as seen from Fig. 5.10(b). In such cases, shape plays an important role along with color and texture. As a result, the spatial relationship between the regions are also expected to play a distinctive feature in this case, as the segmented regions fairly describes different objects.

Images in Fig. 5.11(a) are constituting of a significant object with some background object. The main object almost occupies the major portion of the scene. We have tested the results considering the features (centroids) and other two spatial features separately. The combined result is shown in Fig. 5.11(a) with the top left image as as the query. The results obtained from the combined features outer performed the the individual feature retrieval. Intuitively the shape of the object itself is more meaningful than the color or texture information in the example. For the query example in Fig. 5.12(a), both the background and object are important. The object region has quite a common color and texture. The centroid features which carry the region attributes in terms of color and texture provide a better index than the shape distance. In this case, the color and texture contain most of the useful information. Combining shape distance did not make significant difference to the query results. The retrieval results as shown in Fig. 5.12(b) is obtained by a querying with a dinosaur. As the images are much simpler with almost regular shape and background, the combined features (color, texture and spatial relations) performs highly satisfactorily in such cases.

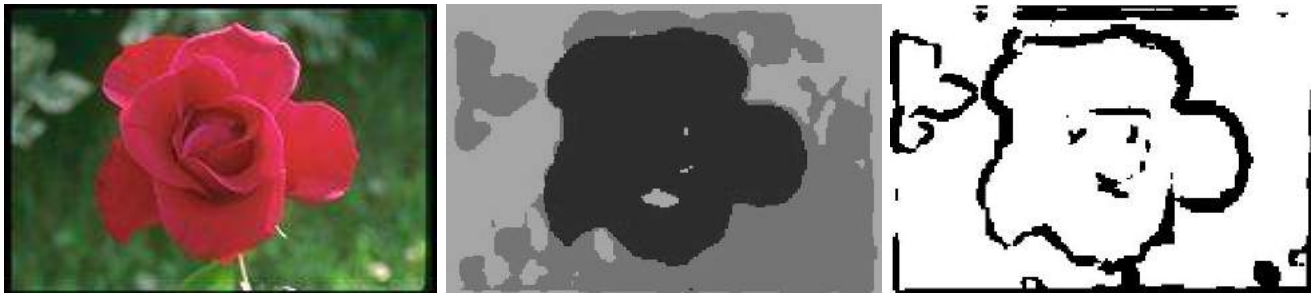
In order to allow comparison between different retrieval features, we present a recall precision graph, computed from different numbers of return images. The system retrieves

r images that belongs to the same class $C1$ as the query ($r \leq n$). There are N_{c1} images in the class $C1$ of the query. Then P=r/n is the precision and R = r/N_{c1} the recall for this query. With each test set image as the query, the average precision and recall for individual feature retrieval is shown in the Fig. 5.13(a). We found that the similarity as obtained from index of fuzziness alone is not satisfactory. Intuitively this feature measures the fuzziness in clustering and indirectly captures dissimilarity between the regions. However the combined results improves the precision.

The properties extracted segmented regions play an important role in characterizing the color and texture contents of different regions from an image. As the robustness of the proposed algorithm is tested in the case of computing overall closeness between images, we compare our results which also evaluate image to image similarities mostly based on the features like color and texture. We benchmark our results with retrieval algorithms using global color texture histograms as used in [151]. Color is represented using a 2D histogram over the HS coordinates of the HSV space. Texture is represented by two histograms (coarseness and directionality), of the image. The similarity distance between two color histograms is computed by histogram intersection. The similarity between two textures of the whole image is determined by a weighted sum of the Euclidean distance. We also provide a comparison in retrieval results with Gabor features computed around randomly selected points. We tested the results in almost all categories of images as shown in Fig. 5.13(b). Histogram gave satisfactory results for images objects having significant objects with quite a common color and texture. In case of general scenes our algorithm generated better results. This proves that the proposed features captures additional spatial information of the segmented regions.

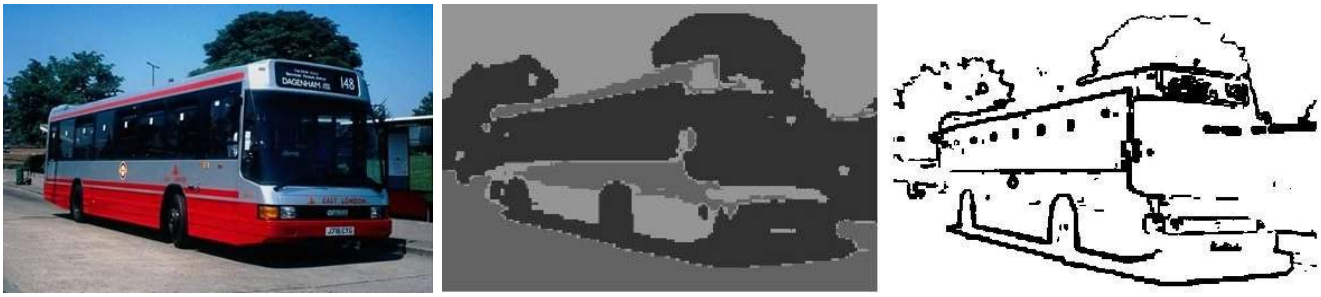
5.4 Conclusion

In this work, a segmentation algorithm has been developed using wavelet packet based features and the fuzzy C- means statistical clustering algorithm. The color and texture



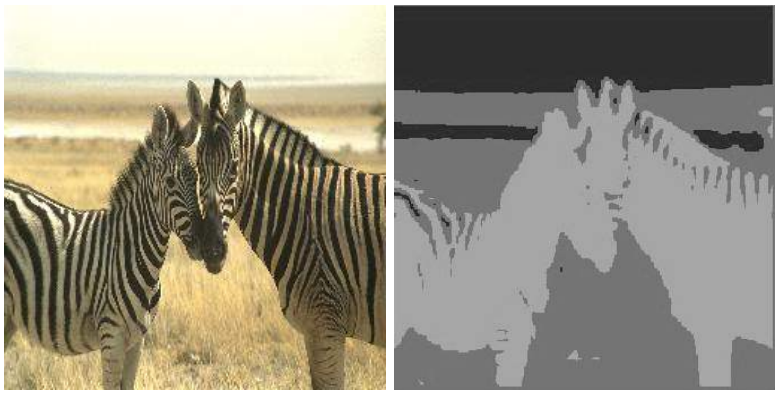
(a) (b) (c)

Figure 5.3: (a) Original image (b) segmented output (3 regions) (c) segmented output (with 3 region boundaries assigned a common black label)



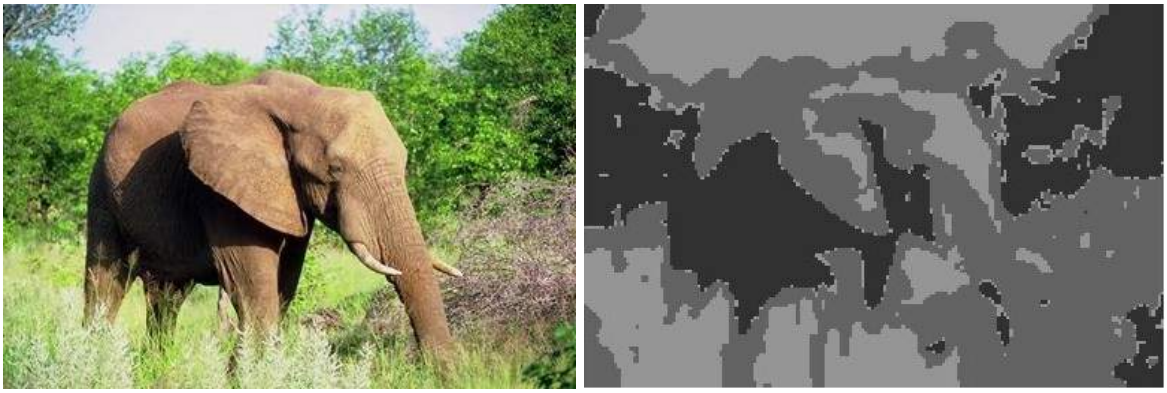
(a) (b) (c)

Figure 5.4: (a) Original image (b) segmented output (3 regions) (c) segmented output (with 3 region boundaries assigned a common black label)



(a) (b)

Figure 5.5: (a) Original image (b) segmented output (3 regions)

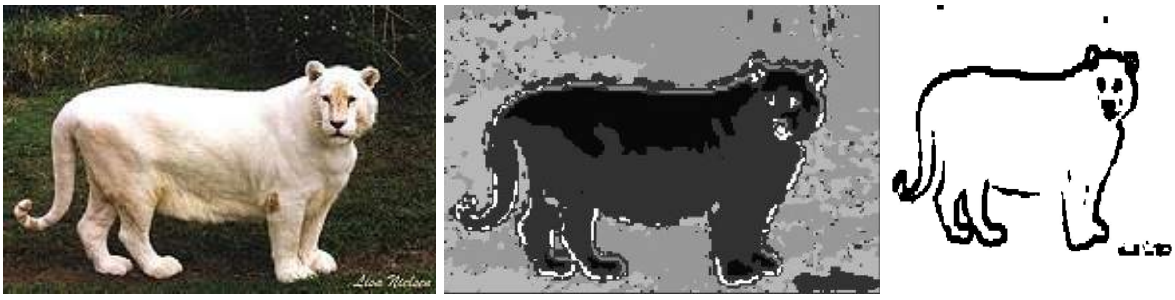


(a) (b)

Figure 5.6: (a) Original image (b) segmented output (3 regions)

(a) (b)

Figure 5.7: (a) Original image (b) segmented output (with 3 region boundaries assigned a common black label)



(a) (b) (c)

Figure 5.8: (a) Original image (b) segmented output (5 regions) (c) segmented output (with 5 region boundaries assigned a common black label)

(a) (b)

Figure 5.9: (a) Retrieval results. (a) Color and textural properties only (b) after combining Fuzzy Spatial Relations. With top left image as the query image

(a) (b)

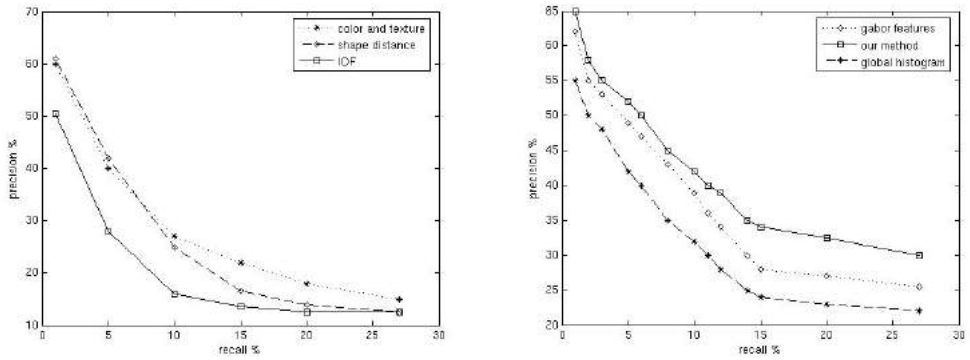
Figure 5.10: (a) Retrieval results. (a) Color and textural properties only (b) after combining Fuzzy Spatial Relations. With top left image as the query image

(a) (b)

Figure 5.11: (a) Retrieval results. (a) Color and textural properties only (b) after combining Fuzzy Spatial Relations. With top left image as the query image

(a) (b)

Figure 5.12: Retrieval results, combining color, texture and Spatial Relation properties.



(a) (b)

Figure 5.13: (a) Recall, Precision curve from results with individual features (b) comparative results

properties of the fuzzy regions are unified with the topological features, which is effective in capturing spatial relation between the regions. In the proposed approach, querying is based on using the description of the entire image rather than on the attributes of one or two regions of interest. The spatial relation features may be more effective if the segmented regions grossly partitions an image into its individual objects. However, it is difficult to decide the number of classes for all images within the database. With the increase in number of classes, crucial details of different region may be obtained, which will reduce the importance of the spatial relation features. Such a choice may not be much helpful in case of evaluating overall similarity between images.

In the present case, the number of classes used to segment an image is set equal to three. This choice has been made as a compromise between computational time required and accuracy. We computed the retrieval time for every image in the database as the query. The proposed approach was found to require an average CPU time of 0.3 secs for retrieval. The average time required for feature extraction is 30secs. The retrieval time primarily depends upon the size of the image, the software and the hardware versions the system runs on, size of the feature vector and the similarity measure used. The time complexity involved for matching and sorting $T(n) \simeq O(NlogN)$ in the present case. The imple-

mentation software includes Matlab codes in a SUN Blade system. The aspect of reducing computational time has to be taken care of in future, to make large-scale retrieval feasible in high dimensional feature space.

The features, proposed to describe the region boundaries have not yet been incorporated for image retrieval applications. Better experiments may be performed using these boundary based features in association with other spatial relationship features, like relative positions between regions. Other interesting extension of the work could be to incorporate suitable weighting factors in order evaluate importance of individual regions, for enhancing the retrieval accuracy. These extensions may provide suitable future directions for research.

Chapter 6

Conclusions and Scope for Further Research

6.1 Discussions and conclusions :

The present thesis has demonstrated the effectiveness of some newly proposed feature extraction methods and a feature evaluation mechanism for image retrieval applications. We have tried to frame the problems in soft decision making paradigm, because it plays an important role in handling various forms of uncertainties like, in data characteristics, similarity evaluations, relevance feedback and other stages of a CBIR system. In the process of designing CBIR systems, four different problems have been considered. These are (i) Fuzzy edge based feature extraction for image retrieval. (ii) Extraction of visually significant points including the corner points by using fuzzy set theoretic approach and a classification based approach, using Support Vector Machines. (iii) Retrieval of color images using features around the extracted visually significant points, along with fuzzy feature evaluation mechanism from relevance feedback, to enhance the retrieval accuracy. (iv) An image retrieval scheme using wavelet based color and textured properties of regions combined with fuzzy spatial relations, to facilitate a better image retrieval scheme has been proposed.

To increase the robustness of an image retrieval system, invariant features are needed. It is desirable that the features used for characterizing an image should be invariant to several transformations like rotation, translation, scale, partial occlusions and also tolerant to noise corruption. These problems have been extensively studied by different researchers in the field of image processing and computer vision applications. In the current thesis, we have taken care of some of these aspects in designing the CBIR systems. We have addressed the problem of feature extraction process when the input images have undergone imaging defects like, overexposed or underexposed lighting conditions, blurring and noise corruption.

In the first part of our investigation, we have tested our algorithm on gray level images. The study in chapter 2, involves feature extraction based on fuzzy edge detection. Topological feature namely, fuzzy compactness computed from the edge map, is invariant to rotation translation and scaling by definition. The invariance properties have been illustrated with examples. It is found that, the fuzzy edge map generated at different thresholded planes, are quite rich in capturing significant information of different objects within an image. No prior region segmentation is necessary for evaluating overall simialrty between images in the database.

The performance of an image retrieval process depends on the kind of features and their representation scheme. When there are large number of images in a database, parameter and threshold selection become an important task. This particular problem is handled in the following way. The edge points are assigned membership values based on the local contrast, computed from difference in gray values over entire range (0 to 255) for an image. The selected edge points are categorized as strong, medium and weak, based on their respective membership values. A judicious choice of threshold can be made following the criteria. Using a threshold at higher $\alpha - cut$ planes say (0.9), the pixels with higher contrast membership values are extracted which are termed as strong boundary pixels. Similarly, selecting threshold value of (0.7), the edge map will extract medium and strong type pixels. Setting a threshold value less than (0.5) will select spurious noise edges along

with significant edge points, which may reduce the accuracy of retrieval.

In the proposed method, three fixed threshold values have been used for detecting edges for all images in the database. A key aspect of the proposed technique is that local gray level contrast information is also embedded with shape information. It is evident from the results that the algorithm can produce fairly satisfactory results for all types of images using the properly selected values. In terms of computational complexity the proposed method is lower than standard edge detection methods like Canny's [40] edge detector. In the proposed algorithm, the parameter (k_d), which determines the degree of connectedness between the adjacent edge points, depends upon the choice of window size. A suitable choice of local window size, from which the value of k_d is determined is important. A trade off between the computation time and accuracy is desirable. Although estimation of $\mu_d(P)$ involves more computational time it helps in generating a refined edge map which is more noise tolerant. The comparison of retrieval performances between the systems that uses Fourier descriptor used with Canny's edge detector as feature and the algorithm that uses the above proposed features, establishes that our features offer better results in most of the situations. In our method better characterization of images may be obtained by virtue of matching edge maps at different thresholding levels which carry contrast information also. The proposed algorithm may also be used for any image processing task where extraction of edge maps is required for a large set of images.

There are several other aspects which have not been explored in the present study like, retrieval of similar images under different illuminations. Also if an image contains multiple objects, the feature descriptor although having the invariance properties, (rotation, translation and scale) can give a general description of the entire image. Such features can be used to get an acceptable result for evaluating overall similarity between images having multiple objects and also single object image. To utilize shape features effectively, information regarding deformation and occlusion of object is required. The performance of the fuzzy edge extraction scheme may be enhanced to get better representation, if it is combined with other properties, which may include projective invariance along with

rotational, translational and scale invariance for challenging applications of object recognition.

In chapter 3, the problem of extracting visually significant points, such as the corner points and points where considerable change in curvature occurs, has been considered with the fuzzy edge map as input. Local information is computed from the extracted visually significant points. The information obtained from the visually significant points are used to generate subtle discrimination between images. We are interested in such a procedure because the corner points and the points near about to it, holds a great potential in generating perceptually significant information. These points have a wide range of application in the area of image retrieval, computer vision with an additional advantage of reducing the search space.

A fuzzy rule based approach is used to extract cluster of point set, where good change of curvature occurs along with the true corner points (very high curvature). The algorithm is tested on a large number of images. To emphasize its real life applications, the problems of extracting features under different illuminations and blurring conditions are also studied. Although the problem of handling uncertainties have been addressed in [61], [112], but the problems of handling images under different imaging conditions and also the problem of selecting different tuning parameters have not been addressed. Corner detection is considered as an important problem in Computer Vision over more than decades. Also there are quite a of number of corner detection methods reported in the literature as already mentioned in chapter3. One main point to be noted is that, the performance of the detectors are dependent on several tuning parameters. Selection of right parameters, suitable for particular image is extremely important. Automatic selection of threshold parameters are therefore necessary to enhance the accuracy of feature extraction. The proposed algorithm takes care of this problem. Selection of threshold, above which the candidate edge points are selected for testing corner like properties is made from pixel contrast histogram. It is shown that algorithm performs fairly well under varying imaging conditions, resulting into blurred, overexposed and underexposed images. Results

are also tested on binary images, which has been changed to gray scale with the effect of smoothing, required in the preprocessing step of generating the fuzzy edge map. One advantage of the methodology is that, the minimum and the maximum pixel contrast ratio as obtained from each image determines the dynamic range between the feet and shoulder point, which maps the membership values between 0 to 1. This increases the flexibility of the algorithm for handling wide variety of images.

We have also compared the results of the proposed algorithm with that of standard benchmark corner detector of Harris [84] and SUSAN [190] corner detector. The complexity involved in the proposed method is also analyzed, which shows that although the proposed algorithm is not computationally superior, but the results are quite satisfactory in terms of accuracy in location of the corner points and handling impreciseness in image data.

A more detailed study is necessary to exploit an automatic threshold selection scheme, which we believe will result into a more useful and flexible corner detection algorithm. In the present work, the problem of feature extraction from noisy images have not been addressed in details. Although, through selection of an exponential function and by estimating degree of connectedness it is possible to make the edge detection algorithm more noise tolerant but such processing involved more computational time. The problem of handling noisy images, in the context of proper choice of membership function can be explored further in future works.

As a further extension of this work we have made a preliminary investigation, using SVM classifier to speed up the process and to tackle the threshold selection problem. This was done from investigation of learning tools to classify the edge points and extract distinctive information. Some interesting observations of this study was noted by considering the edge map as input, that the support points correspond to representative corners. This problem is expected to have interesting extension as future works.

In chapter 4, retrieval of color images using the visually significant points is demonstrated. Several methodologies are proposed in the literature, for effective image retrieval.

However, most of the techniques involve image segmentation into homogeneous regions, where the individual regions grossly represent an image at different object levels. The advantages of such methods are that, spatial relations can be computed if the the images having multiple objects. With such characterization it is also possible for a system to handle sub queries. Such representation although robust but needs to handle certain problems, because of difficulties to estimate the number of classes required for segmentation of each image in an unknown database. To tackle such problems an alternative approach is popular, in which features from visually significant portions are used for effective image characterization [121], [141]. In the proposed image representation scheme, the extracted cluster of points, which represent the fuzzy corner points are used as candidates for computation of features. Along with these features some global measurements have also been computed as component features. Selection of optimal set of features is done from the view point of feature ranking based on non-redundant mutual information criteria. It has also been shown that the global information when combined with information derived form visually significant points (local), can capture better information for characterizing an image. This set of features are used to retrieve images.

We have also addressed the issue of determining suitable weighting factors to be used for various features for obtaining better retrieval accuracy. To tackle such problem a fuzzy relevance feedback method is proposed. The importance of each selected component feature is evaluated from fuzzy feature evaluation index, computed from both the relevant and irrelevant set of images. The invariance properties under rotation, translation, varying illumination, etc. are shown with various examples. It has been shown that, it is not possible to achieve satisfactory results for all type of queries using a fixed set of features. The results are compared with some popular methods like FIRM [51], SIMPLIcity [206] which also compute overall similarity between images using the same database. Considerable success can be achieved using our method, which is shown from the comparative results. To test the robustness of the proposed methodology we, have also compared our retrieval results, with that of the visual features recommended for MPEG-7 standards.

Comparative results so obtained for some of the common categories of images are quite encouraging. To obtain further improved results, a detailed investigation of the proposed feature weighing scheme is necessary. This may provide better compatibility with MPEG-7 feature standards so that the algorithm may be extended for video domain to achieve a better comparable quality.

One of the important problems in image retrieval is that the searching process for each query becomes computationally expensive when there are large number of images in the data base. A proper solution to overcome such a problem is a challenging issue. A suitable indexing method may be helpful, where by searching over a range, may reduce the search space. Several methodologies have been proposed using different machine learning tools like, Neural networks, Support Vector machines (SVM) to classify images at different levels of semantic categories. The algorithms proposed in FIRM and SIMPLIcity have also taken care of this points. Our proposed algorithm may be extended to obtain a better representation scheme, by considering a similar strategy.

In chapter 5, wavelet based region based properties are used in conjunction with fuzzy spatial relations for better characterization of an image. An algorithm for segmenting an image into uniform colored textured regions, using co-efficients of wavelet packet transform has been proposed. Fuzzy spatial interrelations between the segmented regions have also been computed as a measure to incorporate spatial relations, between the segmented regions. The features like color, texture, shape as well as the spatial relations between the individual regions play a significant role in representing the overall meaning associated with an image. Methodologies to compute spatial relationships [66], [152], [102] mostly deal with the geometric attributes of a region like (area, shape, adjacency, surroundedness, etc.). In our proposed approach, a new idea of extracting topological interrelation between fuzzily partitioned regions, of an image has been introduced. The performance is tested on various types of images. The effectiveness of the proposed features after combining spatial relation features with region properties is shown from the improvement in results for different cases. Comparisons with the other methods, which use visual

descriptors like colored and texture properties of homogeneous regions, for evaluating similarities between images are also done.

6.2 Future scope of work :

In the present thesis we have put our attention in developing suitable algorithms which mainly deal with soft decision making in feature extraction, capturing spatial relations between objects, relevance feedback, etc. We have shown various applications on gray level and color images for retrieving relevant images, from image databases.
The work reported in the present thesis may be extended for better results and interesting applications.

1. In the present thesis, we have considered 2D scene images. The applications could be extended to 3D domain, for evaluating shape similarity for 3D objects if we consider additional features which are having the property of geometrical invariance, beside the normal 2D invariance properties like, rotation, translation and scaling transformations which are used in the proposed methods. 3D Zernike invariants may be used as descriptors for content based 3D shape retrieval.

2. Extraction of high curvature points (containing significant shape information) is done using the fuzzy edge map as input containing gradient magnitude and direction information. However, this approach requires a heuristic selection of thresholds for discriminating true corner points from false protrusions. This tuning requires external intervention for different class of images. To avoid this problem, we have used Support Vector Machines which take on the fuzzy edge map as input with gradient and magnitude information and extracts the true corner points as its support points. This kind of approach may be adopted for suitable modification for color images.

3. In the present thesis, we have tested the proposed methodologies for image retrieval, i.e., a single frame. Evaluating similarities between the video clips comprising of

several frames becomes an important research problem. It is very natural that algorithm developed for still images could be extended to video domain by incorporating suitable properties e.g., motion vectors. To test the robustness on the proposed features for possible application in video, we have at first compared the results on 2D images with the contest descriptors recommended as MPEG-7 visual feature standards. On extensive experiments it is found that our experiments are quite compatible with the results as obtained from MPEG-7 visual descriptors, with an additional benefit of handling uncertainties due to blurring, noise corruption and intensity variations. This also justifies our idea for extending the present algorithm for video domain. The results could be further improved with relevance feedback which has also been tested on single frames.

4. The algorithm developed for color texture segmentation using wavelet packets and combining fuzzy spatial relationship may be further modified for better characterization of semantic concepts. In case of natural images some kind of subjective evaluation is possible if importance of individual regions can be calculated. The number of occurrences of individual regions may be a measure, to compute a region weighting factor. Moreover a knowledge based scheme could also be exploited to select the classes automatically. Solutions which take care of these two aspects may result into much better scheme. The system may be further modified which is capable to handle querying on individual regions or blobs.

5. In real life situations the databases consist of very large number with wide variety of images. With a given query, similarity evaluation is basically a search with the nearest neighbors around the query points or a range search around the query points images. Most of the nearest neighbors search techniques employ multidimensional indexing techniques to reduce the search time. Suitable data structures [29] for space partitioning (e.g. Kd-tree) may be used. The proposed features may be coupled with suitable indexing techniques, for handling very large databases while invoking less computational time.

6. In the present thesis, we have mainly concentrated on suitable feature extraction for image retrieval applications. Euclidean distance has been chosen as a similarity measure

for all the methods. The experiments using some other similarity measures like, Mahalanobis or Bhattacharyya distance may be further tested with the proposed features as a future scope.

Chapter 7

Appendix

7.0.1 Multiresolution Analysis using wavelets

The full discrete wavelet expansion of a signal that forms an orthonormal basis for $L^2(R^2)$ is given as,

$$\psi_{j,k} = s^{-j/2}\psi(s^{j/2}t - k\tau) \tag{7.1}$$

The effect of discretizing the wavelet is that the time scale space is sampled at discrete intervals. s is usually chosen as 2 and $\tau = 1$ so that the sampling of the frequency axis and the time axis correspond to dyadic sampling.

An arbitrary function can be represented with linear combination of a set of wavelets or basis functions. These basis functions are obtained from a single prototype wavelet called mother wavelet by dilations(scale) and translations(shift). Wavelet transformation can be used to analyze the space and frequency content of an image. Mallat has shown that a signal at any resolution, can be decomposed into a signal at a coarser resolution, plus the associated wavelet components. The multiresolution characteristics of wavlet transform is very attractive to different image processing and vision applications [130]. The wavelets can be constructed from a scaling function ϕ. The discrete normalized scaling and wavelet basis functions for 2D-DWT are defined as [130],

$$\begin{aligned}\phi_{i,k}(l) &= 2^{i/2}h_i(2^i l - k)\\ \psi_{i,k}(l) &= 2^{i/2}g_i(2^i l - k)\end{aligned} \tag{7.2}$$

where i and k are the dilation and translation parameters and h_i and g_i are respectively the sequence of lowpass and bandpass filters of increasing width indexed by i. (i, $k \in Z$) a positive set of integers. The full discrete wavelet expansion of a signal that form an orthonormal basis for $L^2(R^2)$ is given as,

$$x(l) = \sum_{k\in z} c_{(d_0)}(k)\phi_{d_0,k} + \sum_{i=1}^{d_0}\sum_{k\in z} d_{(i)}(k)\psi_{i,k} \tag{7.3}$$

$d_{(i)}$ are the wavelet coefficients and $c_{(d_0)}$ are the expansion coefficients of the coarser signal approximation $x_{(d_0)}$. The $c_{(d_0)}$ and $d_{(i)}$ can be interpreted in terms of simple filtering and down sampling operations. The 2-D DWT is computed by applying a separable filter bank to the image where $c_i(x, y)$ corresponds to low frequencies (LL) $d_i^1(x, y)$ corresponds to the vertical high frequencies (horizontal edges, LH), $d_i^2(x, y)$ horizontal high frequencies(vertical edges, HL), $d_i^3(x, y)$ the high frequencies in both directions (the corners, HH).

The image $I(x, y)$ is represented at several scales by $\{\ cd_0$, $d_i^n(x, y)$, n= 1,2,3, i=1,..., $d_0\ \}$. The 2D-DWT transform of an $2^i \times 2^i$ image is represented as a set of shifted and dilated wavelet function resulting into subband images, each of size $2^{i-1} \times 2^{i-1}$.

Wavelets are generally expressed by means of filter co-efficients h_n Wavelet decomposition scheme (one level, and three level) are shown in Figs. 7.1 and 7.2.

The multiresolution analysis using wavelet packets can be explained as follows. Multiresolution analysis (MRA) analyzes the signal at different frequency with different resolution. Every spectral component is not resolved equally. This approach makes sense because high frequency component has short duration and low frequency component has long duration. Width of the window is changed as the transform is computed for every single spectral component.

Let us consider a signal $f \in L^2(R)$ and introduce a projection operator P_i such that

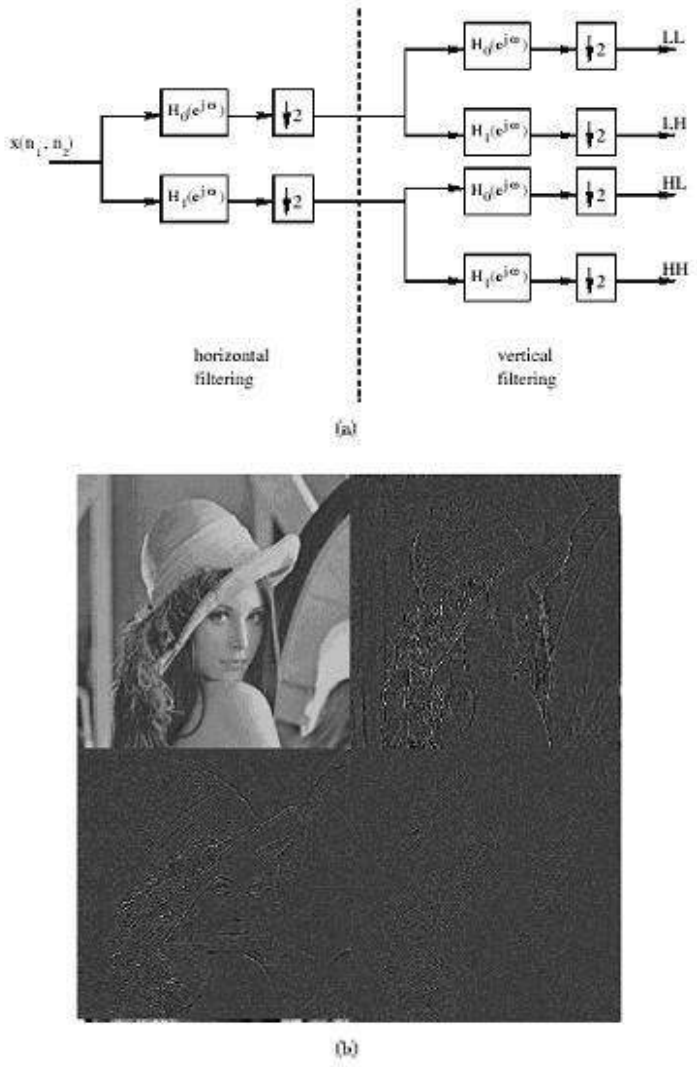


Figure 7.1: Wavelet decomposition scheme

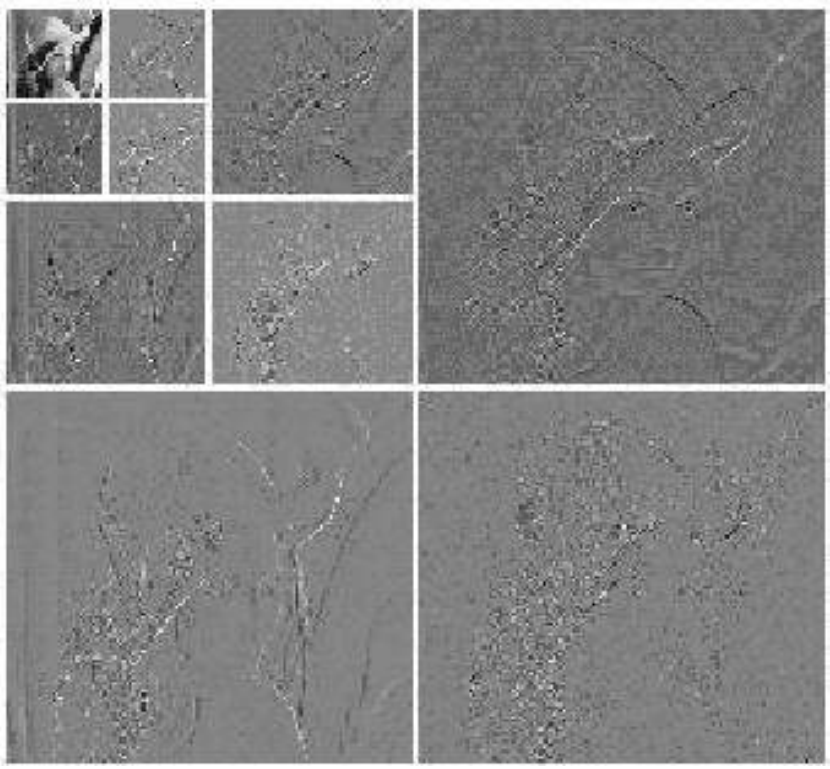


Figure 7.2: 3- level Wavelet decomposition scheme

$P_i f$ is the approximation of f at the resolution i. If we associate this i with a such that, a= 2^i then higher i means "lower resolution".

$$V_j = \{P_j f | f \in L^2(R)\} \tag{7.4}$$

A multiresolution analysis of f is then performed by projecting f on successive lower resolution approximation. Formally, a multiresolution analysis of $f \in L^2(R)$ consists of series of nested subspaces $V_i \in L^2(R)$ for which $V_1 \subset V_2 \subset ...V_n$.

$$\bar{\cup} V_i = L^2(R) \quad and \quad \cap V_i = 0 \tag{7.5}$$

In order to make this ladder of subspaces a multirtesolution representation, it is required that,

$$f(l) \in V_i \quad \Leftrightarrow f(l/2) \in V_{i+1} \ \forall_i \in Z \tag{7.6}$$

The required invariance of V_o under translation is such that,

$$f(l) \in V_0 \quad \Leftrightarrow f(l-k) \in V_0 \ \forall_K \in Z \tag{7.7}$$

This means that the space containing high resolution signals will consist of those of low resolution signals. This ensures elements in a space are simply the scaled versions of the elements in the next space. Which may be stated that, if a function $\phi(l) \in V_0$, then it is also in V_1, the space spanned by $\phi(l/2)$. The full wavelet packet decomposition iteratively decomposes both the approximation and detail subbands at each level. This results in a larger binary tree decomposition that generates two branches from each node. As a result, a very large number of packet bases can be generated for the analysis of many frequency ranges.

Wavelet packets may be described by the collection of $\{WP_i(t)|i \in Z^+\}$ such that, Z^+ is the set of positive integers.

An approximation of the function $f(l)$ using wavelet packet WP_n at scale i with a down sampling factor m is,

$$f(t) = \sum_k WP_{n,i,k} 2^{-p/2} WP_n(2^{-i}l - k) \tag{7.8}$$

Here $WP_{n,i,k}$ are the wavelet packet coefficients.

Wavelet packet coefficients may be efficiently in terms of coefficients at coarser scale as,

$$WP_{2ni,k} = \sum_m h(m-2k)WPn, i-1, m \tag{7.9}$$

$$WP_{2n+1,i,k} = \sum_m g(m-2k)WPn, i-1, m \tag{7.10}$$

As an example, a 2-level decomposition on the orthogonal subspaces generates bases, $P_{2,i,k}(l)$, i= 0,1,2,3 , $p_{1,1,k}$(l)=1/2$\psi(l/2-k)$, $p_{2,0,k}$(l)=1/4$\phi(l/4-k)$, $p_{2,1,k}$(l)=1/4$\psi(l/4-k)$ forms orthogonal sets.

For standard wavelet decomposition, n is restricted to 0. Only WP_{0i} (are decomposed.

The choice of wavelet scale parameter within a DWT is 2^i, where $i \in \{0, 1, 2....\}$. Such choices are not suitable for analysis of high frequency signals with relatively narrow band because the subband decomposition is applied on the low pass component plane of the input image. However, many analysis applications(such as texture classification and segmentation require a finer frequency analysis.)

Since Wavelet packets comprise all possible combinations of subband decomposition applied recursively to the lowpass and high pass filter results of the previous wavelet transform step, it is possible to create arbitrary tree structures that generates various combinations of orthonormal bases [169]. As a result, better frequency localization, while retaining the structure of a discrete decomposition can be obtained trough wavelet packet decomposition.

An appropriate way to perform the wavelet transform for texture feature extraction is to

detect the most significant frequency channels and then decompose them further. In order to avoid the full decomposition, an adaptive decomposition algorithm using a maximum criteria of textural measures based on the statistics extracted from each of the subbands, and identify the most significant subbands can be done. This enables to select the desired frequency subspaces for further decomposition. A signal can be adaptively decomposed using a selected bases according to a selection condition, based on e.g. (energy, entropy of the subbands).

Bibliography

[1] C. Achard, E. Bigorgne, and J. Devars, "A sub-pixel and multispectral corner detector," *ICPR 2000*, vol. 3, pp. 971–974, 2000.

[2] C. Achard, J. Devars, and L. lacassagne, "Object image retrieval with image compactness vectors." *15th International Conference on Pattern Recognition, Barcelona, Spain*, pp. 271–274, 2000.

[3] M. Acharyya and M. K. Kundu, "Robust texture classification using wavelet frames," *Image Processing and communications*, vol. 5(2), pp. 19–37, 1999.

[4] M. Acharyya and M. K. Kundu, "An adaptive approach to unsupervised texture segmentation using m-band wavelet transform," *Signal Processing*, vol. 81, pp. 1337–1356, 2001.

[5] M. Acharyya, M. K. Kundu, and R. K. De, "Document segmentation using wavelet scale-space features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12(12), pp. 1117–1127, 2002.

[6] M. Acharyya, M. K. Kundu, and R. K. De, "Extraction of features using m- band wavelet packet frames and their neuro-fuzzy evaluation for multi-texture segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25(12), pp. 1639–1644, 2003.

[7] M. Acharyya, M. K. Kundu, and R. K. De, "Segmentation of remotely sensed images using wavelet features and their evaluation in soft computing framework,"

IEEE Transactions on Geoscience and Remote Sensing, vol. 41(12), pp. 2900–2905, 2003.

[8] G. Aggarwal, T. V. Ashwin, and S. Ghosal, "An image retrieval system with automatic query modification," *IEEE Transactions on Multimedia*, vol. 4(2), pp. 201–214, 2002.

[9] K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of affine-invariant fourier descriptors to recognition of 3d objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 640–647, 1990.

[10] S. Arivazhagana, L. Ganesanb, and S. P. Priyala, "Texture classification using gabor wavelets based rotation invariant features," *Pattern Recognition Letters*, vol. 27(16), pp. 1976–1982, 2006.

[11] E. M. Arkin, L. Chew, D. Huttenlocher, K. Kedem, and J. S. B. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13(3), pp. 209–226, 1991.

[12] J. R. Bach, C. Fuller, A. Gupta, A. Hamapur, B. Horowitz, R. Humph, R. Jain, and C. F. Shu, "The virage image search engine: an open frame work for image management," *Proceedings SPIE*, vol. 2670, pp. 76–87, 1996.

[13] S. C. Bae, I. S. Kweon, and C. D. Yoo, "Cop : a new corner detector," *Pattern Recognition Letters*, vol. 20, pp. 1349–1360, 2002.

[14] M. Banerjee and M. K. Kundu, "Content based image retrieval with multiresolution salient points," *Fourth Indian Conference Computer Vision, Graphics and Image Processing, ICVGIP 2004, India*, pp. 399–404, 2004.

[15] M. Banerjee and M. K. Kundu, "Handling of impreciseness in gray level corner detection using fuzzy set theoretic approach," *Applied Soft Computing (Article in press)*, 2007.

[16] M. Banerjee, M. K. Kundu, and P. K. Das, "Image retrieval with visually prominent features using fuzzy set theoretic evaluation," *IEE International Conference on Visual Information Engineering VIE 2006, India*, vol. 298-303, 2006.

[17] M. Banerjee, M. K. Kundu, and P. Mitra, "Corner dectection using support vector machine," *17th International Conference on Pattern Recognition ICPR(2004), U.K*, vol. 2, pp. 819–822, 2004.

[18] M. Banerjee and M. K. Kundu, "Content based image retrieval with fuzzy geometrical features," *IEEE International conference in fuzzy systems*, pp. 932–937, 2003.

[19] M. Banerjee and M. K. Kundu, "Edge based features for content based image retrieval," *Pattern Recognition*, vol. 36(11), pp. 2649–2661, 2003.

[20] M. Banerjee and M. K. Kundu, "Content- based image retrieval using wavelet packet and fuzzy spatial relations," *Proceedings of 4th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2006), Lecture Notes in Computer Science series (LNCS-4338)*, pp. 861–871, 2006.

[21] M. Banerjee and M. K. Kundu, "Image retrieval using fuzzy relevance feedback and validation with MPEG-7 content descriptors," *Proceedings of 2nd International Conference on Pattern recognition and Machine Intelligence(PREMI2007), Lecture Notes in Computer Science series (LNCS 4815)*, pp. 144–152, 2007.

[22] M. Banerjee and M. K. Kundu, "Content-based image retrieval using significant features and their evaluation with fuzzy model," *Pattern Recognition (under revision)*, 2008.

[23] M. Banerjee, M. K. Kundu, and P. Maji, "Content-based image retrieval using visually significant point features," *Fuzzy Sets and Systems (revised and submitted)*, 2008.

[24] I. Bartolini, P. Ciaccia, and M. Patella, "Accurate retrieval of shapes using phase of fourier descriptors and time warping distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27(1), pp. 142–147, 2005.

[25] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Transactions on Neural Network*, vol. 5(4), pp. 537–550, 1994.

[26] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(4), pp. 509–522, 2002.

[27] S. Berretti, A. D. Bimbo, and P. Pala, "Retrieval by shape similarity with perceptual distance and effective indexing," *IEEE Transactions on Multimedia*, vol. 2(4), pp. 225–240, 2000.

[28] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, Plenium Press, 1981.

[29] P. K. Bhunre, C. A. Murthy, A. Bishnu, and B. B. Bhattacharya, "A hybrid data and space partitioning technique for similarity queries on bounded cluster," *Proc.1st International Conference on Pattern recognition and Machine Intelligence(PREMI2005), Lecture Notes in Computer Science series (LNCS 3776)*, pp. 544–550, 2005.

[30] A. D. Bimbo, *Visual Information Retrieval*. San francisco, USA: Morgan Kaufmann Publishers, Inc., 2001.

[31] A. D. Bimbo and P. Pala, "Visual image retrieval by elastic matching of user sketches," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19(2), pp. 121–132, 1997.

[32] A. D. Bimbo and P. Pala, "Shape indexing by multi-scale representation," *Image and Vision Computing*, vol. 17, pp. 245–261, 1999.

[33] E. Binaghi, I. Gagliardi, and R. Schettini, "Image retrieval using fuzzy evaluation of color similarity," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 8, pp. 945–968, 1994.

[34] E. Binaghi, I. Gagliardi, and R. Schettini, "Image retrieval using fuzzy representation of colors," *In Proceedings of Ingenta, SoftComputing, Springer*, vol. 11(3), pp. 287–298, 2007.

[35] I. Bloch, "Fuzzy spatial relationships for image processing and interpretation: a review," *Image and Vision Computing*, vol. 23(2), pp. 89–110, 2005.

[36] J. Bracamonte, M. Ansorge, F. Pellandini, and P. A. Farine, "Efficient compressed domain target image search and retrieval," *4th International Conference on Image and Video Retrieval, Singapore*, pp. 154–163, 2005.

[37] C. Brambilla, D. A. Ventura, I. Gagliardi, and R. Schettini, "Multiresolution wavelet transform and supervised learning for content-based image retrieval," *IEEE International Conference on Multimedia Computing and Systems*, vol. 1, pp. 183–188, 1999.

[38] O. M. Bruno, R. M. Cesar, and L. A. Consularo, "Automatic feature selection for biological shape classification in synergos," *SIBGRAPI-98 IEEE.Comput. Society Press*, pp. 363–370, 1998.

[39] C. J. C. Burges, "A tutorial on support vector machines," *Knowledge Discovery and Data Mining 1998*, vol. 2(2), 1998.

[40] J. A. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8(6), pp. 679–698, 1986.

[41] G. Carneiro, A. B. Chan, and P. J. Moreno, "Supervised learning of semantic classes for image annotation and retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29(3), pp. 394–410, 2007.

[42] C. Carson, S. Belongie, H. Greenspan, and J.Malik, "Region - based image querying," *Proceedings of CVPR'97 Workshop on Content- Based access of Image and Video Libraries*, 1997.

[43] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, "Blobworld a system for region - based image indexing and retrieval," *Proceedings of Visual Information Systems*, pp. 509–516, 1999.

[44] Y. K. Chan and C. C. Chang, "A color image retrieval method based on color moment and color variance of adjacent pixels," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 16(1), pp. 113–125, 2002.

[45] F. C. Chang and H. M. Hang, "A relevance feedback image retrieval scheme using multi-instance and pseudo image concepts," *IEICE Transactions on Information and Systems*, vol. D(5), pp. 1720–1731, 2006.

[46] H. S. Chang and K. Kang, "A compressed domain scheme for classifying block edge patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14(2), pp. 145–151, 2005.

[47] S. K. Chang, Q. Y. Shi, and C. W. Yan, "Iconic indexing by 2-d strings," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9(3), pp. 413–428, 1987.

[48] B. B. Chaudhuri, "Fuzzy geometry and shape relations in image spaces," *IETE Journal of Research*, vol. 44(4-5), pp. 161–175, 1998.

[49] B. B. Chaudhury and B. U. Shankar, "An efficient algorithm for extrema detection in digital images." *Pattern Recognition Letters*, vol. 10, pp. 81–85, 1989.

[50] Y. Chen, J. Z. Wang, and R. Krovetz, "CLUE: Cluster-based retrieval of images by unsupervised learning," *IEEE Transactions on Image Processing*, vol. 14(8), pp. 1187–1201, 2005.

[51] Y. Chen and J. Z.Wang, "A region-based fuzzy feature approach to content-based image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(9), pp. 1–16, 2002.

[52] K. O. Cheng, N. F. law, and W. C. Siu, "Multiscale directional filter bank with applications to structured and random texture retrieval," *Pattern Recognition*, vol. 40(4), pp. 610–621, 2007.

[53] S. C. Cheng, "Content-based image retrieval using moment-preserving edge detection," *Image and Vision Computing*, vol. 21(9), pp. 809–826, 2003.

[54] S. C. Cheng, C. T. Kuo, and H. J. Chen, "Visual object retrieval via block-based visual-pattern matching," *Pattern Recognition*, vol. 40(4), pp. 1695–1710, 2007.

[55] Y. Chi and K. H. Maylor, "A local structure matching approach for large image database," *International Conference on Image Analysis and Recognition 2004, In Lecture notes in computer science ISSN 0302-9743*, vol. 3211, pp. 761–768, 2004.

[56] J. M. Corridoni, A. D. Bimbo, and P. Pala, "Image retrieval by color semantics," *ACM Multimedia Systems*, vol. 7(3), pp. 175–183, 1999.

[57] G. C. Cross and A. K. Jain, "Markov random field texture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, pp. 25–39, 1983.

[58] L. da Fontoura costa and R. Marcondes, *Shape analysis and classification*. USA: CRC Press, 2001.

[59] M. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance," *IEEE Transactions on Image Processing*, vol. 11(2), pp. 146–158, 2002.

[60] A. Ferro, G. Gallo, R. Giugno, and A. Pulvirenti, "Best match retrieval for structured images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(7), pp. 707 –718, 2001.

[61] M. A. Fischler and H. C. Wolf, "Locating perceptually salient points on planar curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16(2), pp. 113–129, 1994.

[62] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. S, and P. Yanker, "Query by image and video content: the QBIC system," *IEEE Computer*, vol. 28(9), pp. 23–32, 1995.

[63] J. F.Omhover and M. Detyniecki, "Fast gradual matching measure for image retrieval based on visual similarity and spatial relations," *International Journal on Intelligent Systems*, vol. 21(7), pp. 711–723, 2006.

[64] J. M. Francos, A. Narasimhan, and J. W. Woods, "Maximum likelihood parameter estimation of textures using a wold-decomposition based model," *IEEE Transactions on Image Processing*, vol. 4, pp. 1655–1666, 1995.

[65] H. Freeman and L. S. Davis, "A corner-finding algorithm for chain-coded curves," *IEEE Transaction on Computers*, vol. C-26, pp. 297–303, 1977.

[66] J. Freeman, "The modeling of spatial relations," *Computer Graphics and Image Processing*, vol. 4, pp. 156–171, 1975.

[67] B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 17(5), pp. 522–529, 1995.

[68] I. Gagliardi and R. Schettini, "A method for the automatic indexing of color images for effective image retrieval," *The New Review of Hypermedia and Multimedia*, vol. 3, pp. 201–224, 1997.

[69] T. Gevers and A. W. M. Smeulders, "Color - based object recognition," *Pattern Recognition*, vol. 32, pp. 453–464, 1999.

[70] T. Gevers and A. W. M. Smeulders, "Combining color and shape invariant features for image retrieval." *Image and Vision computing.*, vol. 17(7), pp. 475–488, 1999.

[71] T. Gevers and A. W. M. Smeulders, "Content based image retrieval by viewpoint invariant image indexing," *Image and Vision Computing*, vol. 17(7), pp. 475–488, 1999.

[72] T. Gevers and A. W. M. Smeulders, "PicToSeek: combining color and shape invariant features for image retrieval," *IEEE Transactions on Image Processing*, vol. 9(1), pp. 102–119, 2000.

[73] G. Giacinto and F. Roli, "Bayesian relevance feedback for content-based image retrieval," *Pattern Recognition*, vol. 37(7), pp. 1499–1508, 2004.

[74] Y. Gong, C. H. Chuan, and G. Xiaoyi, "Image indexing and retrieval based on color histograms," *Multimedia Tools and Applications*, vol. 2(2), pp. 133–156, 1996.

[75] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. New York: Wiley, 1985.

[76] V. Gouet and N. Boujemma, "About optimal use of color points of interest for content-based image retrieval, Tech report,RR-4439, INRIA, France May 2006."

[77] A. Gross and L. Latecki, "Digital geometric invariance and shape representation," *International Symposium on Computer Vision, Florida , USA*, pp. 121–126, 1995.

[78] V. N. Gudivada and V. V. Raghavan, "Content based image retrieval systems," *IEEE Computer Society*, vol. 28(9), pp. 18–22, 1995.

[79] J. Han and K. K. Ma, "Fuzzy color histogram and its use in color image retrieval," *IEEE Transactions on Image Processing*, vol. 11(10), pp. 944– 952, 2002.

[80] J. Han, K. N. Ngan, M. Li, and H. J. Zhang, "A memory learning framework for effective image retrieval," *IEEE Transactions on Image Processing*, vol. 14(4), pp. 521–524, 2005.

[81] J. W. Han and L. Guo, "A shape-based image retrieval method using salient edges," *Signal processing*, vol. 18, pp. 141–156, 2003.

[82] R. M. Haralick, "Ridges and valleys on digital images," *Computer Vision, Graphics and Image Processing*, vol. 22, pp. 28–38, 1983.

[83] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3(6), pp. 610–621, 1973.

[84] C. Harris and M. Stephens, "A combined corner and edge detector," *4 th Alvey Vision Conference*, pp. 147–151, 1988.

[85] X. He, O. King, W. Ma, M. Li, and H. J.Zhang, "Learning a semantic space from user's relevance feedback for image retrieval," *IEEE transactions on Circuits and Systems for Video technology 2003*, vol. 13(1), 2003.

[86] M. Heath, S. Sarkar, T. Sanocki, and K. Bowyer, "Assessing the relative performance of edge-detection algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19(12), pp. 1338–1359, 1997.

[87] G. Heidemann, "Combining spatial and colour information for content based image retrieval," *Computer Vision and Image Understanding*, vol. 94(3), pp. 234–270, 2004.

[88] M. A. Hoang, J. M. Geusebroek, and A. W. M. Smeulders, "Color texture measurement and segmentation," *Signal Processing*, vol. 85, pp. 265–275, 2005.

[89] M. K. Hu, "Visual pattern recognition by moment invariants," *IEEE computer Society, Los Angeles, CA, 1977*, 1977.

[90] J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih, "Image indexing using color correlograms," *IEEE International Conference on Computer Vision and Pattern Recognition,*, pp. 762–768, 1997.

[91] A. K. Jain and Vailaya, "Image retrieval using color and shape," *Pattern Recognition*, vol. 29, pp. 1233–1244, 1996.

[92] S. Jeong, C. S. Won, and R. M. Gray, "Image retrieval using color histograms generated by gauss mixture vector quantization," *Computer Vision and Image Understanding*, vol. 94, pp. 44–66, 2004.

[93] N. Jhanwar, S. Chaudhury, G. Seetharaman, and B.Zavidovique, "Content based image retrieval using motif cooccurrence matrix," *Image and Vision Computing*, vol. 22(14), pp. 1211–1220, 2004.

[94] Z. Jin, I. King, and X. Q. Li, "Content-based retrieval by relevance feedback," *Lecture Notes in Computer Science*, vol. 1929, pp. 521–529, 2000.

[95] T. Josta, N. Ouerhania, R. Wartburgb, R. Murib, and H. Huglia, "Assessing the contribution of color in visual attention," *Computer Vision and Image Understanding*, vol. 100(1), pp. 107–123, 2005.

[96] H. Kauppinen, T. Seppnaen, and M. Pietikainen, "An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17(2), pp. 201–207, 1995.

[97] A. Khotanzad and Y. Hong, "Invariant image recognition by zernike moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 489–497, 1990.

[98] L. Kitchen and A. Rosenfeld, "Gray-level corner detection," *Pattern Recognition Letters*, vol. 1, pp. 95–102, 1982.

[99] B. C. Ko and H. Byun, "Multiple regions and their spatial relationship-based image retrieval," *International Conference on Image and Video Retrieval, Lecture notes in computer science*, vol. 2383, pp. 81–90, 2002.

[100] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture image retrieval using rotated wavelet filters," *Pattern Recognition Letters*, vol. 28(10), pp. 1240–1249, 2007.

[101] M. Kokare, B. N. Chatterji, and P. Biswas, "Cosine-modulated wavelet based texture features for content-based image retrieval," *Pattern Recognition Letters*, vol. 25(4), pp. 391–398, 2004.

[102] R. Krishnapuram, J. M. Keller, and Y. Ma, "Quantitative analysis of properties and spatial relations of fuzzy image regions," *IEEE Transactions on Fuzzy Systems*, vol. 1(3), pp. 222–233, 1993.

[103] R. Krishnapuram, S. Medasani, S. H. Jung, Y. S. Choi, and R. Balasubramaniam, "Content-based image retrieval based on a fuzzy approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16(10), pp. 1185–1199, 2004.

[104] S. Kulkarni, "Image retrieval based on fuzzy mapping of image database and fuzzy similarity distance," *International Conference on Computer and Information Science, ICIS 2007*, pp. 812–817, 2007.

[105] A. Kumar, "G-tree: A new data for organizing multidimensional data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 6(2), pp. 341–347, 1994.

[106] M. K. Kundu and M. Acharyya, "Application to texture segmentation for real life image analysis," *International Journal of wavelets, Multiresolution and Information Processing (IJWMIP)*, vol. 1(1), pp. 115–149, 2003.

[107] I. Kunttu, L. Lepisto, J. Rauhamaa, and A. Visa, "Multiscale fourier descriptors for defect image retrieval," *Pattern Recognition Letters*, vol. 27(2), pp. 123–132, 2006.

[108] J. Laaksonen, M. Koskela, S. Laakso, and E. Oja, "PicSOM -content based retrieval with self-organizing maps," *Pattern Recognition Letters*, vol. 21, pp. 1199–1207, 2000.

[109] L. J. Latecki, V. Megalooikonomoua, Q. Wanga, and D.Yu, "An elastic partial shape matching technique," *Pattern Recognition*, vol. 40(11), pp. 3069–3080, 2007.

[110] C. Laurent, N.Laurent, and M. Maurizot, "In depth analysis and evaluation of saliency-based color image indexing methods using wavelet salient features," *Multimedia Tools and Applications*, vol. 31(1), pp. 73–94, 2006.

[111] T. Law, H. Itoh, and H. Seki, "Image filtering, edge detection, and edge tracing using fuzzy reasoning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18(5), pp. 481–491, 1996.

[112] K. J. Lee and Z. Bien, "A gray -level corner detector using fuzzy logic," *Pattern Recognition Letters*, vol. 17, pp. 939–950, 1996.

[113] W. K. Leow and R. Li, "The analysis and applications of adaptive-binning color histograms," *Computer Vision and Image Understanding*, vol. 94, pp. 67–91, 2004.

[114] J. Li and J. Z. Wang, "Real-time computerized annotation of pictures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30(6), 2008.

[115] L. Li and W. Chen, "Corner detection and interpretation on planar curves using fuzzy reasoning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14(4), pp. 1204–1209, 1999.

[116] X. Li and X. Qu, "Matching spatial relations using db-tree for image retrieval," *International Conference on Pattern Recognition*, vol. 2, pp. 1230–1234, 1998.

[117] J. H. Lim and J. S. Jin, "Combining intra-image and inter-class semantics for consumer image retrieval," *Pattern Recognition*, vol. 38(6), pp. 847–864, 2005.

[118] F. Liu and R. Picard, "Periodicity, directionality, and randomness: Wold features for image modeling and retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18(3), pp. 722–733, 1996.

[119] Y. Liu, D. Zhang, G. Lu, and W. Y. Ma, "A survey of content-based image retrieval with high-level semantics." *Pattern Recognition*, vol. 40(1), pp. 262–282, 2007.

[120] S. Loncaric, "A survey of shape analysis techniques," *Pattern Recognition Letters*, vol. 31(8), pp. 983–1001, 1998.

[121] E. Loupias and N. Sebe, "Wavelet-based salient points: Applications to image retrieval using color and texture features," *In Advances in visual Information Systems, Proceedings of the 4th International Conference, VISUAL 2000*, pp. 223–232, 2000.

[122] D. Lowe, "Distinctive image features from scale invariant keypoints," *International Journal of Computer vision*, vol. 2(60), pp. 91–110, 2004.

[123] D. G. Lowe, *Perceptual Organization and Visual Recognition*. USA: Kluwer Academic Publishers., 1985.

[124] G. Lu, "Techniques and data structures for efficient multimedia retrieval based on similarity," *IEEE Transactions on Multimedia*, vol. 4(3), pp. 372–384, 2002.

[125] S. Y. Lu and K. S. Fu, "A syntactic approach to texture analysis," *Computer Graphics and Image Processing*, vol. 7, pp. 303–330, 1978.

[126] J. Luo, M. Boutell, R. T. Gray, and C. Brown, "Image transform bootstrapping and its applications to semantic scene classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 35(3), pp. 563– 570, 2005.

[127] W. Y. Ma and M. Manjunath, "NeTra :a toolbox for navigating large image databases," *Proc. IEEE International Conference on Image Processing*, pp. 568–571, 1997.

[128] F. Mahmoudi, J. Shanbehzadeh, A. Moghadam, and H. Zadeh, "Image retrieval based on shape similarity by edge orientation autocorrelogram," *Pattern Recognition*, vol. 36, pp. 1725–1736, 2003.

[129] S. P. Maity, M. K. Kundu, and T. S. Das, "Robust SS watermarking with improved capacity," *Pattern Recognition Letters*, vol. 28(4), pp. 350–356, 2007.

[130] S. Mallat, "A theory for multiresolution signal decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11(7), pp. 674–693, 1989.

[131] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Fast wavelet histogram techniques for image indexing," *Journal of Computer Vision and Image Understanding*, vol. 75(1), pp. 99–110, 1999.

[132] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18(8), pp. 837–841, 1997.

[133] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11(6), pp. 703–715, 2001.

[134] B. S. Manjunath, P. Salembier, and T. S. (editors), *Introduction to MPEG-7: Multimedia Content description Interface*. John Wiley and Sons, Inc., USA, 2002.

[135] J. Mao and A. K. Jain, "Texture classification and segmentation using multiresolution simultaneous autoregressive models," *Pattern Recognition*, vol. 25(2), pp. 173–188, 1992.

[136] J. M. Martnez, *MPEG-7 Multimedia Description Schemes (XM)*. ISO/IEC JTC1/SC29/WG11 N3914, 2001.

[137] J. M. Martnez, *MPEG-7 Overview (version 10)*. ISO/IEC JTC1/SC29/WG11N6828, 2004.

[138] J. C. Martnez, J. M. Medina, C. D. Barranco, G. Perales, and J. M. S. Hidalgo, "Retrieving images in fuzzy object-relational databases using dominant color descriptors," *Fuzzy Sets and Systems*, vol. 158(3), pp. 312–324, 2007.

[139] J. Matas, R. Marik, and J. Kittler, "On representation and matching of multi-colored objects," *5th International Conference on Computer Vision*, vol. 31, pp. 1369–1390, 1998.

[140] B. M. Mehtre, M. S. Kankanhalli, and W. F. Lee, "Shape measures for content based image retrieval: a comparison," *Information Processing and Management*, vol. 33(3), pp. 319–337, 1997.

[141] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer vision*, vol. 1(60), pp. 63–86, 2004.

[142] P. Mitra, C. Murthy, and S. Pal, "A probabilistic algorithm for active support vector learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26(3), pp. 413–418, 2004.

[143] P. Mitra, B. U. Shankar, and S. K. Pal, "Segmentation of multispectral remote sensing images using active support vector machines," *Pattern Recognition Letters*, vol. 25(9), pp. 1067–1074, 2004.

[144] H. A. Moghaddam, T. K. Taghizadeh, A. H. Rouhi, and M. T. Saadatmand, "Wavelet correlogram :a new approach for image indexing and retrieval," *Pattern Recognition*, vol. 38(12), pp. 2506–2518, 2005.

[145] H. A. Moghaddam and M. S. Tarzjan, "Gabor wavelet correlogram algorithm for image indexing and retrieval," *IEEE International Conference on Pattern Recognition (ICPR'06)*, pp. 925–928, 2006.

[146] H. Moravec, "Towards automatic visual obstacle avoidance," *Proc. 5th International Joint Conference on Artificial Intelligence*, p. 584, 1997.

[147] H. Muller, W. Muller, D. Squire, S. M. Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern Recognition Letters*, vol. 22, pp. 593–601, 2001.

[148] M. S. Nixon and A. S. Aguado, *Feature extraction and Image Processing*. Oxford, England: Reed Educational and Professional Publishing Ltd., 2002.

[149] M. Novotni and R. Klein, "3d zernike descriptors for content based shape retrieval," *In Proceedings of ACM, Seatle, Washington, USA 2003*, 2003.

[150] P. J. Oonincx and P. de Zeeuw, "Adaptive lifting for shape-based image retrieval," *Pattern Recognition*, vol. 36(11), pp. 2663–2672, 2003.

[151] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T. S. Huang, "Supporting similarity queries in MARS," *In Proceedings of the 5th ACM International Multimedia Conference, Seattle, Washington*, vol. 403-413, 1997.

[152] S. K. Pal, "Fuzziness, image information and scene analysis," in *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, R. R. Yager and L. A. zadeh, Eds. Norwell: Kluwer Academic, 1992, pp. 147–184.

[153] S. K. Pal and B. Chakraborty, "Intraclass and interclass ambiguities (fuzziness) in feature evaluation," *Pattern Recognition Letters*, vol. 2, pp. 275–279, 1984.

[154] S. K. Pal and B. Chakraborty, "Fuzzy set theoretic measures for automatic feature evaluation," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 16(5), pp. 754–760, 1986.

[155] S. K. Pal and D. D. Majumder, *Fuzzy mathematical Approach to Pattern Recognition*. New York: Willey Eastern Limited, 1985.

[156] S. K. Pal and A. Rosenfeld, "Image enhancement and thresholding by optimization of fuzzy compactness," in *Fuzzy Models For pattern Recognition*, J. C. Bezdek and S. K. Pal, Eds. IEEE PRESS, 1991, pp. 369–379.

[157] S. K. Pal, A. Ghosh, and M. K. Kundu, *Soft Computing for Image Processing*. Physica-Verlag, 2000, ch. 1, pp. 44–78.

[158] S. K. Pal and P. Mitra, *Pattern Recognition Algorithms for Data Mining*. CRC Press, Boca Raton, Florida, 2004.

[159] M. Partio, B. Cramariuc, and M. Gabbouj1, "An ordinal co-occurrence matrix framework for texture retrieval. ID 17358," *EURASIP Journal on Image and Video Processing*, 2007.

[160] G. Pass and R. Zabith, "Comparing images using joint histograms," *Multimedia Systems*, vol. 7, pp. 234–240, 1999.

[161] A. Pentland, R. Picard, and S. Sclaroff, "Content-based manipulation of image databases," *International Journal of Computer Vision*, vol. 18(3), pp. 233–254, 1996.

[162] E. G. M. Petrakis and C. Faloutsos, "Similarity searching in large image databases," College Park, MD, USA, Tech. Rep. CS-TR-3388, 1994. [Online]. Available: citeseer.ist.psu.edu/petrakis95similarity.html

[163] I. Popovici and W. D. Withers, "Custom-built moments for edge location," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28(4), pp. 637–642, 2006.

[164] H. N. Pour and E. Kabir, "Image retrieval using histograms of uni-color and bi-color blocks and directional changes in intensity gradient," *Pattern Recognition Letters*, vol. 25(14), pp. 1547–1557, 2004.

[165] F. Qian, B. Zhang, and F. Lin, "Constructive learning algorithm-based rbf network for relevance feedback in image retrieval," *In Lecture Notes in Computer Science*, vol. 2728, pp. 352–361, 2003.

[166] J. Quinlan, *Programs for Machine Learning*. Morgan Kaufmann, CA, 1993.

[167] M. K. M. Rahman, W. P. Yang, T. W. S. Chow, and S. Wu, "A flexible multi-layer self-organizing map for generic processing of tree-structured data," *Pattern Recognition*, vol. 40(5), pp. 1406–1424, 2007.

[168] Rajashekhar, S. Chaudhuri, and V. P. Namboodiri, "Retrieval of images of man-made structures based on projective invariance," *Pattern Recognition*, vol. 40(1), pp. 296–308, 2007.

[169] R. M. Rao and A. S. Bopardikar, *Wavelet Transforms Introduction to Theory and Applications*. Pearson Education, Asia (Singapore), 2002.

[170] A. Rattarangsi and R. T. Chin, "Scale-based detection of corners of planar curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14(4), pp. 430–449, 1992.

[171] S. Ravela, R. Manmatha, and W. Croft, "Retrieval of trademark and gray-scale images using global similarity," 1998. [Online]. Available: citeseer.nj.nec.com/ravela98retrieval.html

[172] S. Ravela, R. Manmatha, and E. M. Riseman, "Image retrieval using scale-space matching," *Proceedings of the 4th European Conference on Computer Vision, In Lecture notes in computer science*, vol. 1064, pp. 273–282, 1996.

[173] A. Rosenfeld, "Fuzzy digital topology," in *Fuzzy Models For pattern Recognition*, J. C. Bezdek and S. K. Pal, Eds. IEEE press, 1991, pp. 331–339.

[174] A. Rosenfeld and E. Johnston, "Angle detection on digital curves," *IEEE Transactions on Computers*, vol. C-22, pp. 875–858, 1973.

[175] I. E. Rube, M. Ahmed, and M. Kamel, "Wavelet approximation-based affine invariant shape representation functions," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 28(2), pp. 323–327, 2006.

[176] Y. Rubner, L. J. Guibas, and C. Tomasi, "The earth mover's distance, multi- dimensional scaling, and color- based image retrieval," *Proceedings of DARPA Image understanding Workshop*, pp. 661–668, 1997.

[177] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," *In Proceedings of the IEEE International Conference on Image Processing*, pp. 815–818, 1997.

[178] Y. Rui, T. S. Huang, and S. Mehrotra, "Relevance feedback : a power tool for interactive content-based image retrieval," *IEEE transactions on Circuits and Systems for Video technology*, vol. 8(5), pp. 644–655, 1998.

[179] S. Santini and R. Jain, "Similarity measures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21(9), pp. 871–883, 1999.

[180] S. Santini, *Exploratory Image Databases : Content Based retrieval*. San Diego, USA: Academic Press, 2001.

[181] S. Sclaroff and A. Pentland, "Modal matching for correspondence and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17(6), pp. 545–561, 1995.

[182] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26(5), pp. 550–571, 2004.

[183] S. Shim and T. Choi, "Edge color histogram for image retrieval," *In Proc. of IEEE International Conference on Image Processing*, vol. 3, pp. 957–960, 2002.

[184] K. Siddiqi, A. Shokoufandeh, S. Dickinson, and S. Zucker, "Shock graphs and shape matching," *International Journal of Computer Vision*, vol. 35(1), pp. 13–32, 1999.

[185] D. G. Sim, H. K. Kim, and R. H. Park, "Invariant texture retrieval using modified zernike moments," *Image and Vision Computing*, vol. 22(4), pp. 331–342, 2004.

[186] N. Sladoze, I. Nystrom, and P. K. Saha, "Measurements of digitized objects with fuzzy borders in 2d and 3d," *Image and Vision Computing*, vol. 23(2), pp. 123–132, 2005.

[187] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(12), pp. 1349–1380, 2000.

[188] J. R. Smith and S. F. Chang, "Single color extraction and image query," *International Conference on Image Processing*, vol. 3, pp. 3528–3531, 1995.

[189] J. R. Smith and S. F. Chang, "VisualSEEK a fully automated content-based image query system," *In Proc. of the fourth ACM International Conference on Multimedia*, pp. 87–98, 1997.

[190] S. Smith and M. Brady, "A new approach to low level image processing," *International Journal of Computer Vision*, vol. 23(1), pp. 45–78, 1997.

[191] J. J. Song and F. Golshani, "Shape-based 3d model retrieval," *IEEE International Conference on Tools with Artificial Intelligence*, 2003.

[192] D. M. Squire, W. Muller, H. Muller, and T.Pun, "Content based query of image databases: inspirations from text retrieval," *Pattern Recognition Letters*, vol. 21, pp. 1993–1998, 2000.

[193] M. Stricker and M. Orengo, "Similarity of color images," *SPIE Storage and Retrieval for Image and Video Databases III*, vol. 2185, pp. 381–392, 1995.

[194] M. A. Stricker and A. Dimai, "Color indexing with weak spatial constraints," *In Proceedings of Storage and Retrieval for Image and Video Databases (SPIE)*, pp. 29–40, 1996.

[195] K. T. Su, K. S. Jin, and K. I. Lee, "Content-based image retrieval using color and pattern histogram adaptive to block classification characteristics," *International conference on Computer analysis of images and patterns, In Lecture notes in computer science ISSN 0302-9743*, vol. 3691, pp. 120–127, 2005.

[196] M. Swain and D. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7(1), pp. 11–32, 1991.

[197] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transactions on Systems Man and Cybernetics*, vol. 3(6), pp. 610–621, 1978.

[198] J. Tarel and S. Boughorbel, "On the choice of similarity measures for image retrieval by example," *In ACM International Conference on Multimedia*, pp. 107–118, 2002.

[199] C. H. Teh and R. T. Chin, "On the detection of dominant points on digital curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11(8), pp. 859–872, 1989.

[200] J. K. Udupa and G. J. Grevera, "Go digital, go fuzzy," *Pattern Recognition Letters*, vol. 23, pp. 743–754, 2002.

[201] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on Image Processing*, vol. 4(11), pp. 1549–1560, 1995.

[202] V. Vapnik, *The Nature of Statistical Learning Theory*. NewYork: Springer Verlag, 1995.

[203] E. D. Ves, J. Domingo, G. Ayala, and P. Zuccarello, "A novel bayesian framework for relevance feedback in image content-based retrieval systems," *Pattern Recognition*, vol. 39(9), pp. 1622–1632, 2006.

[204] J. Wang, H. Zha, and R. Cipolla, "Combining interest points and edges for content-based image retrieval," *In Proc. of IEEE International Conference on Image Processing*, vol. 3, pp. 1256–1259, 2005.

[205] J. Z. Wang, R. M. Gray, and G. Wiederhold, "System for screening objectionable images." *Computer Communications Journal*, vol. 21(15), pp. 1355–1360, 1998.

[206] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive integrated matching for picture libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(9), pp. 947–963, 2001.

[207] L. Wang and T. Pavlidis, "Direct gray-scale extraction of features for character recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15(10), pp. 1053–1066, 1993.

[208] Y. Wang and F. Makedon, "R-histogram: Quantitative representation of spatial relations for similarity-based image retrieval," *The 11th Annual ACM International Conference on Multimedia, Berkeley, California*, pp. 323–326, 2003.

[209] J. Weijer, T. Gevers, and A. Bagdanov, "Boosting color saliency in image feature detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28(1), pp. 150–156, 2006.

[210] J. Weijer, T. Gevers, and J. Geusebroek, "Edge and corner detection by photometric quasi-invariants," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27(4), pp. 625–629, 2005.

[211] S. Wu, M. K. M. Bahman, and T. W. S. Chow, "Content-based image retrieval using growing hierarchical self-organizing quadtree map," *Pattern Recognition*, vol. 38(5), pp. 707–722, 2005.

[212] L. Yanmei, Z. Hongchen, and M. Guoguang, "Fuzzy color image retrieval based on correlation," *In Proc. of SPIE*, vol. 4922, pp. 122–126, 2002.

[213] C. H. Yao and S. Y. Chen, "Retrieval of translated rotated and scaled color textures," *Pattern Recognition*, vol. 36(4), pp. 913–929, 2003.

[214] P. Y. Yin, B. Bhanu, K. C. Chang, and A. Dong, "Integrating relevance feedback techniques for image retrieval using reinforcement learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27(10), pp. 1536–1551, 2005.

[215] Y. H. Yua and C. C. Chang, "A new edge detection approach based on image context analysis," *Image and Vision Computing*, vol. 24(10), pp. 1090–1102, 2006.

[216] L. A. Zadeh, "Fuzzy sets." *Information and Control*, vol. 8, pp. 338–353, 1965.

[217] G. Zaji, N. Koji, V. Radosavljevi, I. Reljin, I. Reljin, and B. Reljin, "Accelerating of image retrieval in CBIR system with relevance feedback. article ID 62678," *EURASIP Journal on Advances in Signal Processing*, 2007.

[218] D. Zhang, M. Chen, and Y. Lim, "An efficient and robust technique for region based shape representation and retrieval," *International Conference on Computer and Information Science, ICIS 2007*, pp. 801–806, 2007.

[219] D. Zhang and G. Lu, "Study and evaluation of different fourier methods for image retrieval," *Image and Vision Computing*, vol. 23(1), pp. 33–49, 2005.

[220] H. Zhang, R. Rahamani, S. R. Cholleti, and S. A. Goldman, "Local image representations using pruned salient points with applications to CBIR," *Proceedings of the 14th annual ACM international conference on Multimedia, USA*, pp. 287–296, 2006.

[221] J. Zhang and T. N. Tan, "Brief review of invariant texture analysis methods," *Pattern Recognition*, vol. 3, pp. 735–747, 2002.

[222] Z. Zheng, H. Wang, and E. Teoh, "Analysis of gray level corner detection," *Pattern Recognition Letters*, vol. 20(2), pp. 149–162, 1999.

[223] X. S. Zhou and T. S. Huang, "Edge-based structural features for content-based image retrieval," *Pattern Recognition Letters*, vol. 22, pp. 457–468, 2001.

[224] X. S. Zhou and T. S. Huang, "Relevance feedback in content based image retrieval: some recent advances," *Information Sciences*, vol. 148, pp. 129–137, 2002.

[225] J. Zunic and P. L. Rosin, "A new convexity measure for polygons," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26(7), pp. 923–934, 2004.