Content-Based Image Retrieval

Bhavana Kakde

Department of Electronics and Communication Engineering
National Institute of Technology Rourkela

Content-Based Image Retrieval

Thesis submitted in partial fulfillment

of the requirements for the degree of

Master of Technology

in

Electronics and Communication Engineering
(*Specialization*: *Signal and Image Processing*)

by

Bhavana Kakde

(Roll Number: 216EC6250)

based on research carried out

under the supervision of

Prof. Manish Okade

May, 2018

Department of Electronics and Communication Engineering
National Institute of Technology Rourkela

Department of Electronics and Communication Engineering
National Institute of Technology Rourkela

Prof. Manish Okade
Professor

May 29, 2018

Supervisor's Certificate

This is to certify that the work presented in the thesis entitled *Content-Based Image Retrieval* submitted by *Bhavana Kakde*, Roll Number 216EC6250, is a record of original research carried out by her under my supervision and guidance in partial fulfillment of the requirements of the degree of *Master of Technology* in *Electronics and Communication Engineering*. Neither this thesis nor any part of it has been submitted earlier for any degree or diploma to any institute or university in India or abroad.

Manish Okade

Dedication

I dedicate my entire thesis to my parents, my supervisor, brother and friends who always had my back. Signature.

Signature

Declaration of Originality

I, *Bhavana Kakde*, Roll Number *216EC6250* hereby declare that this thesis entitled *Content-Based Image Retrieval* presents my original work carried out as a postgraduate student of NIT Rourkela and, to the best of my knowledge, contains no material previously published or written by another person, nor any material presented by me for the award of any degree or diploma of NIT Rourkela or any other institution. Any contribution made to this research by others, with whom I have worked at NIT Rourkela or elsewhere, is explicitly acknowledged in the dissertation. Works of other authors cited in this dissertation have been duly acknowledged under the sections "Reference" or "Bibliography". I have also submitted my original research records to the scrutiny committee for evaluation of my dissertation.

I am fully aware that in case of any non-compliance detected in future, the Senate of NIT Rourkela may withdraw the degree awarded to me on the basis of the present dissertation.

May 29, 2018
NIT Rourkela

Bhavana Kakde

Acknowledgment

I express my deep sense of gratitude to my esteem and motivational guide Mr. Manish Okade for their precious help and guidance, I am thankful to him for giving encouragement he has given to me in finishing my thesis work.

His trust and bolster helps me in the making right decision during the difficult phase of my thesis work. My special thanks goes to my guide and Prof. T K Dan sir head of the Electronics and Communication Engineering department for providing best opportunities and facilities for presenting my work in the IEEE conference.

I am also grateful to all my teachers Prof. S. K. Patra, Prof. K. K. Mahapatra, Prof. A. K. Sahoo, Prof. Manish Okade, Prof. S. K. Das, Prof. L. P. Roy, other faculties and staff member of our department for their kind support , motivation. They also help in making my fundamentals of theoretical knowledge strong and research thereafter. Thankful affirmation is made to my PhD researcher who gave me significant help by methods for proposal, remarks and feedback. Also, I profoundly welcome the commitment made in different routes by my companions and colleagues.

May 29, 2018
NIT Rourkela

Bhavana Kakde
Roll Number: 216EC6250

Abstract

Content-Based Image Retrieval (CBIR) has a useful role in image retrieval framework. It is a commonly received solution for an efficient and effective method that can look up for the required image from the large database without human interaction. There occurs a need of CBIR because the development digital images, due to widespread capturing of images using web cameras, mobile phones enable with the camera, and digital cameras is rendering the management of image database tedious. It can also be used in other application such as web engines and social media which stores a large number of images and requires fast retrieval of the image selected by the user.

The extraction of a feature in CBIR is a noticeable step whose viability is dependent upon the strategy used for feature extraction from given images. These features can be arranged in classes like as histogram, spatial layout, shape, texture, color, etc. The CBIR uses these features to retrieve and index the image database. The objective is to find a unique representation for all different variation because the user captures images in various conditions such as occlusion and varying illumination etc. Feature extraction method used for CBIR are Local tetra patterns (LTrP) and Dual-Cross Patterns (DCP). Local tetra patterns (LTrP) is a method which acquires more detailed information by using four possible directions of every center pixel in an image, and is calculated from first order derivatives in horizontal and vertical directions. DCP encodes second order information in the vertical, horizontal and diagonal direction, by performing the encoding of sample points in the local surrounding region of every center pixel in an image.

Simulation is performed in Matlab 8.6 to evaluate the performance of retrieval framework using LTrP and DCP, and Corel 900 database(D) is used for this purpose. The simulation performance of presented technique is evaluated in terms of average recall and average precision. The average recall for LTrP is 42% and the average precision for DCP is 63% The average recall for DCP is 35% and the average precision for DCP is 61.05%. Our extensive simulation on Corel database shows that the DCP technique has better time complexity compared to LTrP methods.

Keywords*: *Content-based image retrieval (CBIR)*; *dual-cross patterns(DCP)*; *local tetra patterns (LTrPs)*; *texture*; *average recall.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

The term “content based” of the CBIR refers to analyses of image content in the search, instead of the metadata (e.g. tags, keywords, or descriptions related with the image) The term “content” in this field might allude to colors, shapes, surfaces, or any other data that can be inferred from the image itself. CBIR is alluring since searches that depend simply on metadata are subordinate on annotation quality and completeness. Having people physically annotate on images by entering keyword or metadata in a huge database can be time consuming and might not capture the keyword required to depict the image. Research on image databases and image processing, problems ranging from storage issues to user friendly interfaces are involved in CBIR systems, image retrieval is an context-and application-dependent task.Image retrieval requires change (translation or conversion) of high-level user perceptions to low level image features. It is common practice to represent n features by the numerical values and then represent the whole image or object as a single point in an n-dimensional space. Also Common similarity metrics and multi- dimensional indexing techniques are the factors to be taken into consideration. The indexing structures specifications to boost image retrieval speed and the query specification are some of the main challenges in this domain. In addition to them, query processing also affected by the visual interpretation related cognitive aspects. Data mining, query processing are such other problems that attract researchers to this domain.

1.1.1 Architecture of CBIR Systems

Figure 1.1 depicts a conventional content-based image retrieval system architecture. Two principal functionalities involved are:
1) Data insertion
2) Query processing
Extracting features from the images and putting them into the image database is a part of data insertion module. This process is ordinarily performed off-line. The query processing,

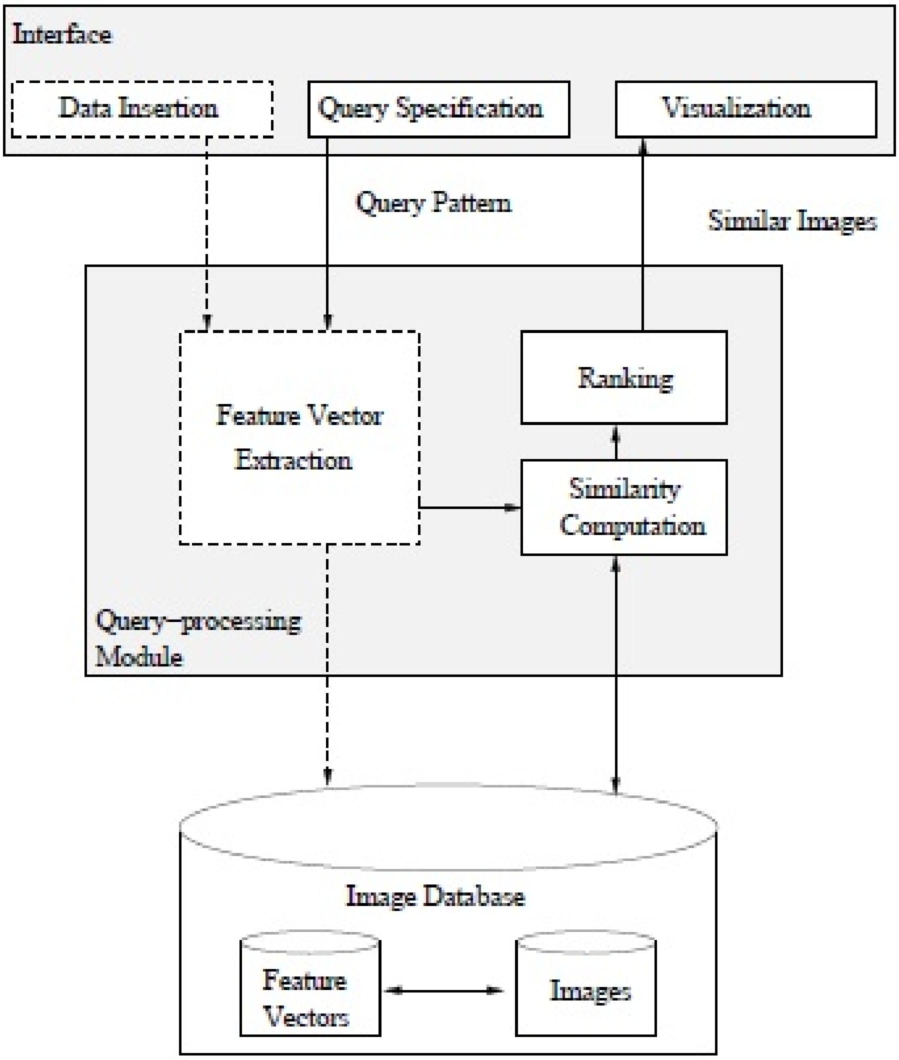


Figure 1.1: Typical architecture of a CBIR.

in turn, is organized as follows: the interface permits a client to specify a query image by implies of a query pattern and to visualize the recovered comparative images. From a query pattern, query-processing module extricates a feature vector and applies a metric (such as the Euclidean distance) to assess the similarity between the query image and the database images. Next, it positions the database images in a decreasing order of likeness to the query image and advances the foremost similar images to the interface module. Feature vector extraction module helps in functionality of both the query processing and data insertion module.

1.2 Motivation

Advanced innovation has led to fast development of computerized media collections, often containing both still images and video recordings. Storage gadgets are filled with terabytes of digital images, making it progressively harder to recover images of interest from such large collections. It is obvious that search capabilities are required for finding what we are searching for in such large image collections, but how can we make such search valuable? Manual comment of images with keywords depicting the image substance can make it simpler to discover images of interest, but this takes a lot of time, making this approach very costly, since we do not have the prior knowledge of searches in advance, it helps to certain extent. Besides, distinctive people are likely to annotate the same picture using different catchphrases, making it difficult to make a reasonable classification and clarify images with the "correct" catchphrases On the ground of this, the use of content-based image retrieval (CBIR) has been advocated.

Content-Based Image Retrieval (CBIR) has a useful role in image retrieval framework. It is a commonly received solution for an efficient and effective method that can look up for the required image from the large database without human interaction. There occurs a need of CBIR because the development digital images, due to widespread capturing of images using web cameras, mobile phones enable with the camera, and digital cameras is rendering the management of image database tedious. It can also be used in other application such as web engines and social media which stores a large number of images and requires fast retrieval of the image selected by the user. The extraction of a feature in CBIR is a noticeable step whose viability is dependent upon the strategy used for feature extraction from given images. These features can be arranged in classes like as histogram, spatial layout, shape, texture, color, etc. The CBIR uses these features to retrieve and index the image database. The objective is to find a unique representation for all different variation because the user captures images in various conditions such as occlusion and varying illumination etc. An all-inclusive and considerable literature study on CBIR is given in [1–4]. The texture is well appropriated for the recognition of items such as marble, slabs, parquet, ceramic tiles, etc., and has drawn wide consideration from texture retrieval industries and recherches. Texture analysis has potential in extracting the outstanding features due to which it is being used in computer vision technology and pattern recognition application. The concept of wavelet correlogram is introduced in Moghaddam et al [5], [6]. Discrete wavelet transform (DWT) [7] have been used for texture classification. DWT generalized Gaussian density plus KullbackLeibler distance, an application of DWT, has given efficient results for retrieval of the texture image [8] and image segmentation [9]. But, DWT can extricate data from horizontal direction, diagonal direction, and vertical direction, in an image.

1.3 Related Work

The local binary pattern (LBP) feature has emerged as a useful operator in the field of texture classification and retrieval. Ojala et al. proposed LBPs [10], which are converted to a rotational invariant version for texture classification [11], [12]. Various extensions of the

LBP [13], completed LBPs [14], dominant LBPs [15], etc., also introduced for rotational invariant and multi resolution gray scale texture classification and other descriptors, LBPs [10] are used for CBIR [12], [11]. In addition to this, the LBP descriptor is used in face representation for the reason of capturing texture information of the face [16]. Zhang et al. proposed local derivative patterns (LDPs) for texture classification, where they considered the LBP as nondirectional first-order local patterns collected from the first-order derivatives and extended the same approach for nth order LDPs [17]. The forms of the LDP and the LBP within the open literature cannot adequately serve with the range of appearance variations such as lighting variations , different facial poses ,changes in facial expression, partial occlusion of face part, the aging problem etc. that commonly occur in unconstrained natural images In order to deal with such problems, the local ternary pattern (LTP) [18] has been proposed for face recognition under different illumination conditions. The LBP, the LDP, and the LTP extricate the information on the bases of distribution of edges, which are coded utilizing as it were two directions (negative direction or positive direction). Therefore, it is evident that the performance evaluation of these techniques can be enhanced by finding the edges in more than two directions. A method using local tetra pattern (LTrP) retrieval algorithm is used for direction limitation. The local tetra patterns (LTrP) method is able to acquire more detailed information by encoding the relationship between the center pixel and its surrounding pixels, with four different values based on the directions that are estimated using the first-order derivatives in horizontal and vertical directions [19], LTRP use for CBIR. Dual-Cross Patterns (DCP) is a novel descriptor, which captures useful information by encoding second order discriminative data in the vertical, horizontal and diagonal image components [20]. Sampling strategy adapted by DCP extracts twice the pixel information as compared to LBP. The sampled pixels are grouped in accordance with maximum joint Shannon entropy, thereby keeping a reasonable feature vector size. A sub- DCP cross encoder (denoted herein as E-1 and E- 2) having exactly the same memory cost and time complexity as LBP, gives a significantly better performance. The DCP have better time complexity, DCP use for CBIR.

1.4 Problem definition

To find novel image indexing and retrieval algorithm using local tetra patterns (LTrPs) and dual-cross pattern (DCP) for Content-based image retrieval (CBIR). There occurs a need of CBIR, because the development digital images, due to widespread capturing of images using web cameras, mobile phones enable with the camera, and digital cameras is rendering the management of image database tedious.

1.5 Organization of the Thesis

Thesis is arrange in the following order.

Chap2: This chapter consists of local patterns(LBP, LTP, LTrPs and DCP) and their example, matlab output of local pattern on test image.

Chap3: This chapter explains two techniques for Content-Base Image Retrivel using LTrP and DCP, experimental setup conducted to evaluate the performance of retrieval framework using LTrP and DCP.

Chap4: This chapter consists of conclusion from the work done and also talks about the future scope of the work.

Chapter 2

Local Patterns and their Examples

2.1 LBPs

The local binary pattern (LBP), an efficient visual texture descriptor, performs a comparison between the center pixel and its neighboring pixel.

$$LBP_{M,R} = \sum_{m=1}^{M} 2^{m-1} f_1(G_M - G_c) \tag{2.1}$$

2

$$f_1(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.2}$$

where G_c is the gray value of the center pixel, G_m is the gray value of its neighbors M , is the number of neighbors, and R is the radius of the neighborhood.

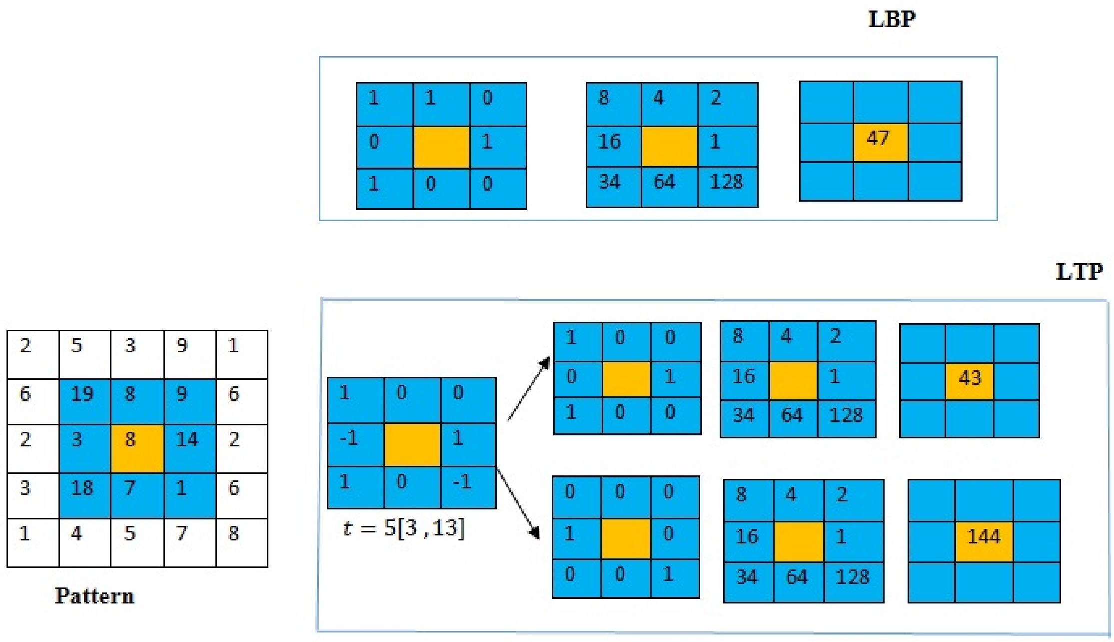


Figure 2.1: Example of the LBP and LTP operators.

2.2 LTPs

As the LBP is sensitive to noise near uniform regions so to improve upon Tan and Triggs [18] extended it to three-valued code and referred as Local Ternary Pattern. Total Number of possible codes in LTP is from *0* to $3^n - 1$ in which pixels with gray level intensity values within the of width $\pm w$ around G_c are quantized to zero, those above (G_c+w) arequantized to +1, and those below (G_c-w) are quantized to -1, $f_1(x)$ operator is changed with three-valued function and is given by the following equation:

$$f_1(x, G_c, w) = \begin{cases} +1, & \text{if } x \geq G_c + w \\ 0, & \text{if } |x - G_c| < w \\ -1, & \text{if } x \leq G_c - w \end{cases} \tag{2.3}$$

where x=G_m

2.3 LTrPs

The LTrP illustrates the spatial structure of the local neighborhood by using the direction of the gray level value of center pixel G_c.

In a given image I, first order derivatives in the direction of 0^0 and 90^0 are represented as $I^1_\theta(G_m)|_{\theta=0^0,90^0}$. Say G_c signify the center pixel in the image, and let G_v and G_h indicate vertical and horizontal neighborhoods of G_c respectively. Then, at the center pixel G_c, the first-order derivatives can be written as:

$$I^1_{0^0}(G_c) = I(G_h) - I(G_c) \tag{2.4}$$

$$I^1_{90^0}(G_c) = I(G_v) - I(G_c) \tag{2.5}$$

and the direction of the G_c can be evaluated as

$$I^1_{Dir.}(G_c) = \begin{cases} 1, & I^1_{0^0}(G_c) \geq 0 \text{ and } I^1_{90^0}(G_c) \geq 0 \\ 2, & I^1_{0^0}(G_c) < 0 \text{ and } I^1_{90^0}(G_c) \geq 0 \\ 3, & I^1_{0^0}(G_c) < 0 \text{ and } I^1_{90^0}(G_c) < 0 \\ 4, & I^1_{0^0}(G_c) \geq 0 \text{ and } I^1_{90^0}(G_c) < 0 \end{cases} \tag{2.6}$$

From (2.6), it is noticeable that the possible direction for each center pixel can be either 1, 2, 3 or 4, and eventually, the image is converted into four values, i.e., directions.

The second-order $LTrPs^2(G_c)$ is defined as

$$LTrPs^2(G_c) = \{f_3(I^1_{Dir.}(G_c), I^1_{Dir.}(G_1)), f_3(I^1_{Dir.}(G_c), I^1_{Dir.}(G_2)), \ldots, f_3(I^1_{Dir.}(G_c), I^1_{Dir.}(G_M))\}|_M = 8 \tag{2.7}$$

$$f_3(I^1_{Dir.}(G_c), I^1_{Dir.}(G_m)) = \begin{cases} 0, & I^1_{Dir.}(G_c) = I^1_{Dir.}(G_m) \\ I^1_{Dir.}(G_m), & \text{else.} \end{cases} \tag{2.8}$$

From (2.7) and (2.8), for each gray level center pixel 8-bit tetra pattern is obtained. Depending upon the direction of center pixel, tetra pattern is separated into four parts. Then, the tetra patterns for each part (direction) are changed to three binary patterns. Let the direction of center pixel ($I^1_{Dir.}(G_c)$) obtained using (2.6) be "1"; $LTrP^2$ then, can be defined by segregating it into three binary patterns as follows:

$$LTrP^2|_{Direction=2,3,4} = \sum_{m=1}^{M} 2^{m-1} f_4(LTrP^2(G_c))|_{Direction=2,3,4} \tag{2.9}$$

$$f_4(LTrP^2(G_c))|_{Direction=\phi} = \begin{cases} 1, & \text{if } LTrP^2(G_c) = \phi \\ 0, & \text{else.} \end{cases} \tag{2.10}$$

where ϕ = 2,3,4.

So also, the other three-tetra patterns for remaining three directions of center pixels are changed over to binary patterns. In this way, we get 12 (4 $\times$ 3) binary patterns.

For texture classification Guo et al.[14] utilizes the magnitude component of the local neighborhood difference operator to propose the magnitude LBP, along side with the sign LBP. They conclude that, despite of the fact that the sign component extricates more useful data as compared with the magnitude component, exploiting the combination of magnitude and sign components can provide superior clues, which are not apparent in any individual component. This concept has propelled us to propose the 13th binary pattern design by utilizing the magnitudes of horizontal and vertical first-order derivatives using

$$Mag_{I^1(G_m)} = \sqrt{(I^1_{0^0}(G_m))^2 + (I^1_{90^0}(G_m))^2} \tag{2.11}$$

$$LP = \sum_{m=1}^{M} 2^{m-1} \times f_1(Mag_{I^1(G_m)} - Mag_{I^1(G_c)})|_{M=8} \tag{2.12}$$

Local pattern with M neighborhoods results in a 2^M possible LBP combination and hence feature vector of length 2^M. The computational cost of this feature vector is very expensive. To lessen the computational cost, the uniform examples [13] are considered. The uniform example alludes to the uniform appearance design that has constrained discontinuities in the circular binary representation. Those patterns that have less than or equal to two discontinuities of 0 to 1 or 1 to 0 in the circular binary representation are referred to as the uniform patterns, and the remaining patterns in which there are more than two transitions from 0 to1 or 1 to 0 are referred to as nonuniform. Thus, for a given query image the different uniform patterns would be $M(M-1)+2$. The possible uniform patterns for M=8 can be seen in [13].

Subsequent to recognizing the local pattern such as the LBP, the LTP or the 13 binary pattern from LTrP, a histogram is built for the entire image using:

$$\begin{aligned} &H_S\left(l\right) = \tfrac{1}{N_1 \times N_2} \sum_{j=1}^{N_1} \sum_{k=1}^{N_2} f_s\left(PTN\left(i,j\right),l\right); \\ &l \in \left[0, M\left(M-1\right)+2\right] \end{aligned} \tag{2.13}$$

$$f_s(x,y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{otherwise} \end{cases} \tag{2.14}$$

where $N_1 \times N_2$ represents the size of the input image.

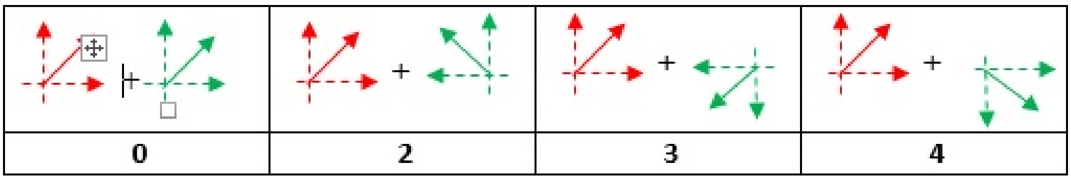
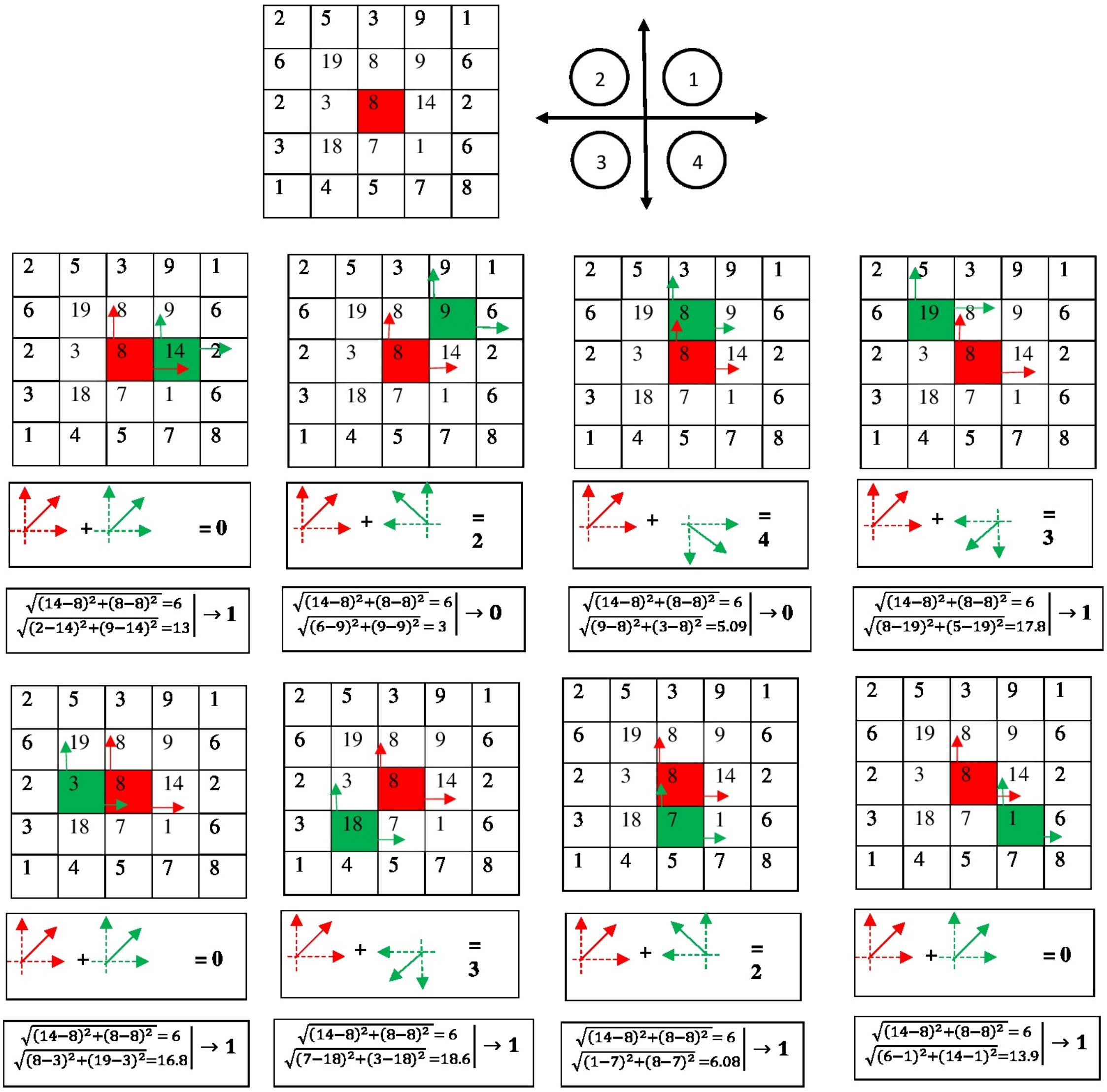


Figure 2.2: determination of tetra pattern for direction 1 of center pixel that results in a LTrP.

For direction 1 of center pixel Fig 2.2 outlines the conceivable local pattern transitions that results in a LTrP. If direction is same as the direction of center pixel then LTrP is coded to 0 else LTrP is coded along the direction of neighborhood pixels. In the similar manner for direction 2,3,4 of the center pixel LTrPs are coded.

Fig 2.3 explains computation of second order LTrP that results in direction 1 of center pixel, it is marked with red. On applying first order derivative in the vertical and horizontal direction of local neighborhood of pixel with value 14, magnitude of 1.3 and direction 1 is obtained. As the direction achieved from neighborhood and the direction of center pixel are similar, 0 should be assigned to the corresponding bit of LTrP. It can seen that magnitude of center pixel is 6. As the magnitude of the neighborhood pixel is less than the magnitude of center pixel 0 will be assigned to corresponding bit of magnitude pattern. Also, the rest of the bits of the LTrP and the magnitude pattern for the remaining seven neighbors are evaluated



Tetra Pattern = 0 2 4 3 0 3 2 0

Pattern1 = 0 1 0 0 0 0 1 0; Pattern2 = 0 0 0 1 0 1 0 0; Pattern3 = 0 0 1 0 0 0 0

Magnitude Pattern = 1 0 0 1 1 1 1 1

Figure 2.3: Example of LTrP to obtain patterns (tetra and magnitude).

bringing about the tetra pattern "0 2 4 3 0 3 2 0" and the magnitude pattern "1 0 0 1 1 1 1 1". On accomplishment of tetra pattern coding, it is partitioned into three binary pattern as follows. Alluding to the created LTrP, first pattern is acquired by keeping "1" where the tetra design esteem is "2" and "0" for other esteem, i.e., "0 1 0 0 0 0 1 0". In the same fashion other two binary patterns examples "0 0 0 1 0 1 0 0" and "0 0 1 0 0 0 0 0" are processed for tetra pattern esteems "3" and "4", individually. Similarly, tetra patterns for center pixel having directions 2, 3, and 4 are processed. Hence, 12 binary patterns are acheived corresponding to four tetra patterns. Using the first order derivative magnitude,13th binary pattern is achieved.

2.4 Dual-Cross Patterns

Image filtering, local sampling, and pattern encoding are the three principle parts in the formation of a DCP image descriptor. The key of DCP is to carry out pattern encoding and local sampling in the most descriptive necessary direction, and this encrypts second-order statistical information.

The sampled points are encoded in two stages. First, encoding of textural details along each of the eight directions is done, which is followed by the combining of the patterns obtained to form the DCP codes. A grouping strategy is used in DCP to reduce the computational complexity but due to the information lost, grouping is done according to joint Shannon entropy to minimize it.

2.4.1 local sampling

Dual-cross pattern encrypts higher order useful information in the direction of image components; vertical, horizontal, and diagonal directions of local neighborhood. In Fig.2.4, a local sampling of the Dual cross pattern is shown. For each center pixel O ,symmetrically sample the points in eight directions (0, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, and $7\pi/4$) in the local neighborhood. In each direction two pixels gets sampled and the final sampled pixels are represented as $\{A_0, B_0; A_1, B_1; ...; A_7, B_7\}$. As shown in Fig.2.4, on the inner

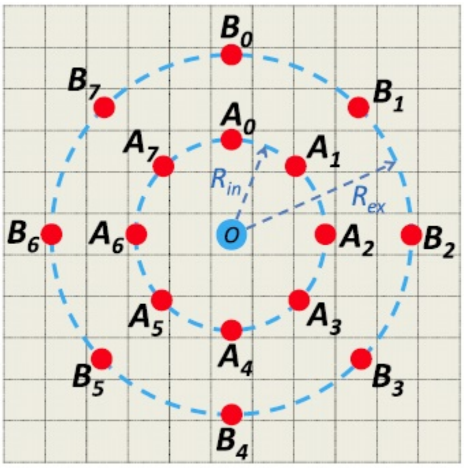


Figure 2.4: Local sampling of DCP. Sixteen pixels are sampled around the local neighborhood central pixel O.

circle having radius R_{in} points $A_0, A_1, ..., A_7$ are uniformly distributed, while $B_0, B_1, ..., B_7$ points are evenly spaced on the outer circle having radius R_{ex}.

2.4.2 Pattern Encoding

Two steps are used to encode the sampled points. First, encoding of textural details along each of the eight directions is done, which is followed by the combining of the patterns obtained to form the DCP codes. A unique decimal number is assigned in each sampling direction to quantize the textural information:

$$Dcp_k = F(A_{A_k} - A_O) \times 2 + F(A_{B_k} - A_{A_k}), 0 \leq k \leq 7 \tag{2.15}$$

where

$$F(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2.16}$$

and A_0, A_{A_k}, and A_{B_k} gray level value of center pixel *O*, and neighboring pixels A_k, and B_k, respectively. Accordingly, second-order statistics along each direction are encoded by the four patterns and each of the four patterns signifies one kind of textural information.

When all eight directions are considered simultaneously the total number of possible dual cross pattern code is 4^8 = 65,536. This value is too high for real-time recognition application, so to reduce computational complexity, eight-direction are divided into two group, and each group called one encoder. By this manner, the accumulative number of local patterns is minimized to $4^4 \times 2$ = 512, which helps in decreasing the computational complexity. Although this technique results in data loss, robustness and compactness of the descriptor are promoted.

2.4.3 Dual-Cross Grouping

The total thirty-five combinations are produced on partitioning eight directions by grouping technique explained in the earlier subsection. Joint Shanon entropy criteria is used for optimal grouping of eight directions to preserve the necessary information required for image retrieval. With the above analysis, Dcp_k $(0 \leq k \leq 7)$ can take one of the four possible values: 0, 1, 2, and 3. The joint Shannon entropy for the set $\{Dcp_0, Dcp_1, Dcp_2, Dcp_3\}$ is represented as

$$H(Dcp_0, Dcp_1, Dcp_2, Dcp_3)$$

$$= -\sum_{dcp_0} ... \sum_{dcp_3} P(dcp_0, ..., dcp_3)\mathrm{log}_2 P(dcp_0, ..., dcp_3) \tag{2.17}$$

where $dcp_0, dcp_1, dcp_2,$ and dcp_3 are particular values of $Dcp_0, Dcp_1, Dcp_2,$ and Dcp_3 respectively. $P(dcp_0, ..., dcp_3)$ is the probability of $dcp_0, dcp_1, dcp_2,$ and dcp_3 values occurring simultaneously. For four variables, joint Shannon entropy is maximum when they have statistical independence.

In images when the pixels are more sparsely scattered they are more independent of each other. Hence, when the sample points are separated by a maximum distance; maximum joint Shannon entropy in each subgroup is obtained. As a result $\{Dcp_0, Dcp_2, Dcp_4, Dcp_6\}$ forms the first subset and $\{Dcp_1, Dcp_3, Dcp_5, Dcp_7\}$ forms the second subset.

2.4.4 DCP Feature Descriptor

These two sub Dcp encoders are called E-1 and E-2, respectively. For each pixel *O*, code generated by sub DCP encoder E-1:

$$E - 1 = \sum_{k=0}^{3} Dcp_{2k} \times 4^k, \tag{2.18}$$

and code generated by the second sub DCP encoder E-2:

$$E - 2 = \sum_{k=0}^{3} Dcp_{2k+1} \times 4^k, \tag{2.19}$$

By combining the codes generated by the two sub DCP encoder E-1 and E-2; final DCP descriptor for pixel *O* is given as:

$$DCP = \left\{ \sum_{k=0}^{3} Dcp_{2k} \times 4^k, \sum_{k=0}^{3} Dcp_{2k+1} \times 4^k \right\}. \tag{2.20}$$

Using the sub DCP encoders, two coded images are generated that are divided into nonoverlapping regions and in each individual region, histogram is computed. To generate the holistic representation all the histograms are combined. By using the similarity metrics such as query matching, a similarity can be measured using this holistic representation between a pair of images.

2.5 Matlab output of local pattern on test image

1	2	3	4	5	6	7	8	9	1	2
2	2	3	9	8	7	6	5	4	3	3
3	4	1	5	7	2	4	3	5	1	4
4	3	6	2	5	5	9	5	6	7	5
5	2	4	6	7	9	1	5	8	3	6
6	5	3	7	3	4	8	2	2	1	7
7	4	2	3	8	3	2	9	4	5	8
8	8	1	8	2	3	2	1	3	2	9
9	7	8	2	4	5	4	3	5	8	1
1	8	1	4	2	3	5	7	1	9	2
2	3	4	5	9	8	7	6	5	4	3

Pattern Encoding $Dcp_0 = 1\,x\,2+1$, $Dcp_1 = 1\,x\,2+1$, $Dcp_2 = 0\,x\,2+0$, $Dcp_3 = 0\,x\,2+1$

$Dcp_4 = 0\,x\,2+1$, $Dcp_5 = 1\,x\,2+1$, $Dcp_6 = 1\,x\,2+0$, $Dcp_7 = 0\,x\,2+0$

Dual-cross grouping First subset {3, 0, 1, 2} second subset {3, 1, 3, 0}

DCP Feature Descriptor E-1 = $Dcp_0\,x\,4^0 + Dcp_2\,x\,4^1 + Dcp_4\,x\,4^2 + Dcp_6\,x\,4^3$

$= 3 \times 1 + 0 \times 4 + 1 \times 16 + 2 \times 64 = 147$

E-2 = $Dcp_1\,x\,4^0 + Dcp_3\,x\,4^1 + Dcp_5 x\,4^2 + Dcp_7\,x\,4^3$

$= 3 \times 1 + 1 \times 4 + 3 \times 16 + 0 \times 64 = 55$

Figure 2.5: Example of DCP.

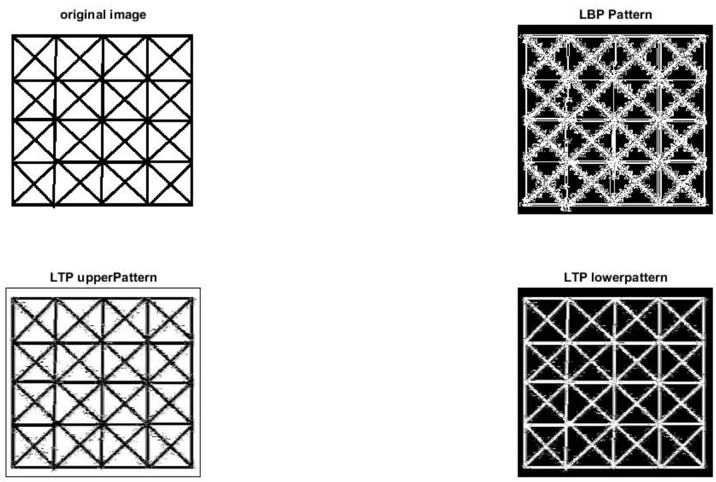


Figure 2.6: Coded output image by using descriptors (LBP and LTP).

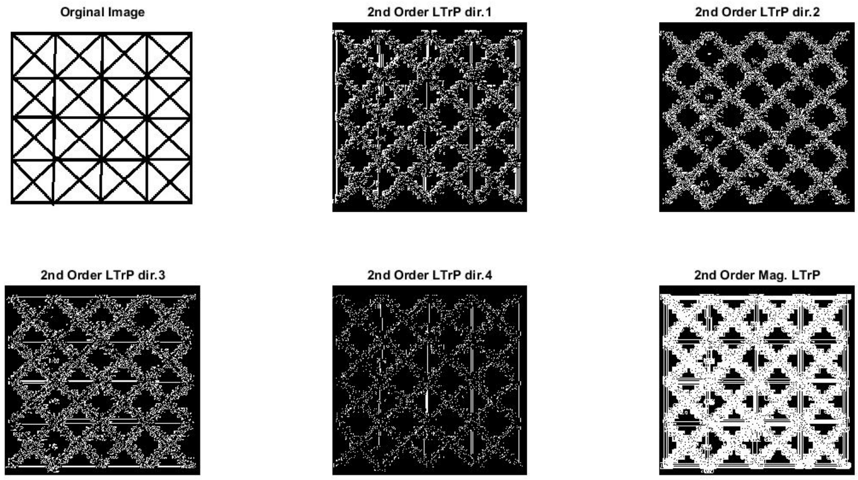


Figure 2.7: Coded output image by using LTrP descriptor.

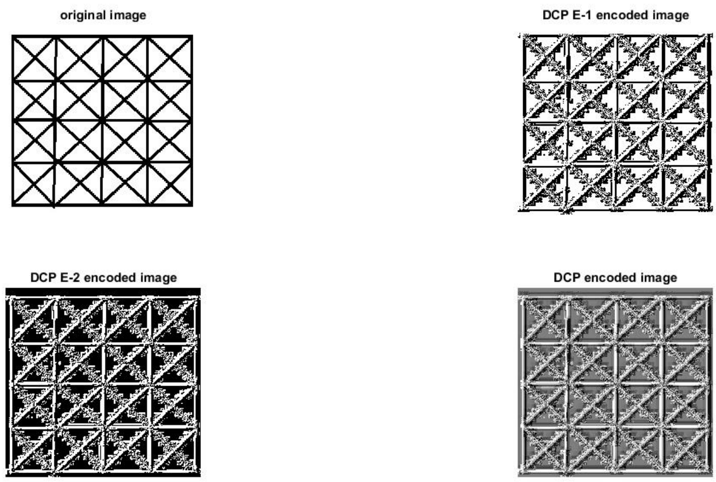


Figure 2.8: Coded output image by using DCP descriptor.

Chapter 3

Techniques for CBIR

3.1 Using local tetra patterns (LTrP)

Fig. 3.1 shows the framework of the technique and an algorithm is explained below.

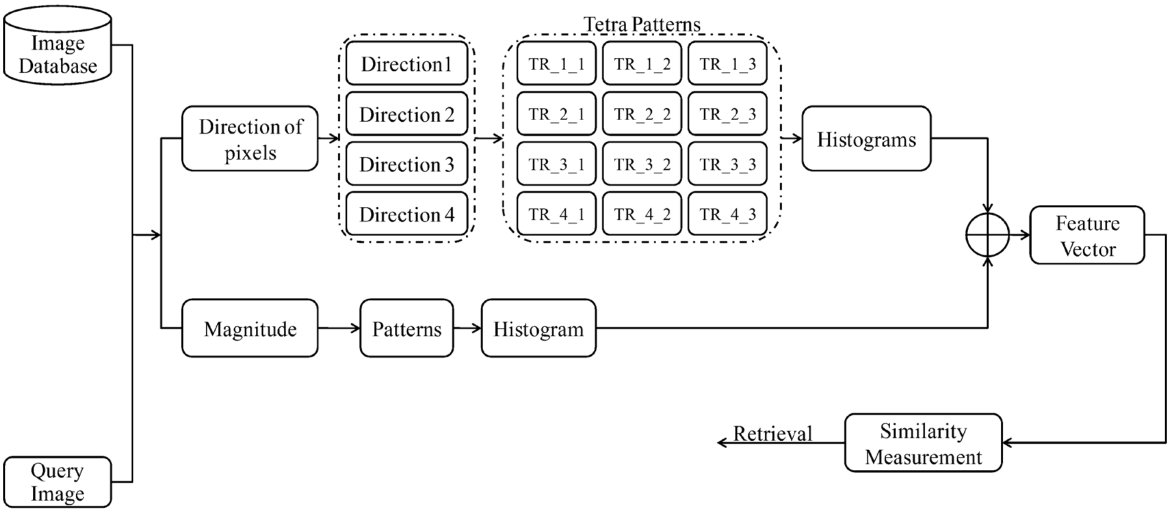


Figure 3.1: Image retrieval system framework.

Algorithm:

I/P: Query image; O/P: Retrieved images

1. Take an image, and change its color model to grayscale model.
2. Find first-order derivatives along vertical and horizontal direction.
3. Direction is evaluated for every pixel.
4. Depending upon the direction of center pixel patterns are partationed in four parts.
5. Find the tetra patterns, and obtained three binary patterns from tetra patterns.
6. From binary patterns build the histograms.

7. Using (2.11), magnitudes of center pixels are computed.
8. Binary patterns are constructed and compute their histogram.
9. Histogram obtained from steps 6 and 8 are combined.
10. Generate the feature vector.
11. The images in given database are compared with the query image using (3.1).
12. Images are retrieved based on the best matches.

3.2 Using Dual-Cross Patterns (DCP)

Fig. 3.2 shows the framework of the technique and an algorithm is explained below.

Algorithm:

I/P: Query image; O/P: Retrieved images

1. Take an image, and change its color model to grayscale model.
2. Perform symmetrical sampling in eight directions in the local neighborhood for each pixel of the image.
3. In each sampling direction, to quantize the textural information pattern encoding is used.
4. Dual-Cross Grouping is performed to minimize information loss.
5. Two cross encoder generates the encoded images and derive the feature vector for each images.
6. Combine both the feature vector.
7. The images in given database are compared with the query image using (3.1).
8. Images are retrieved based on the best matches.

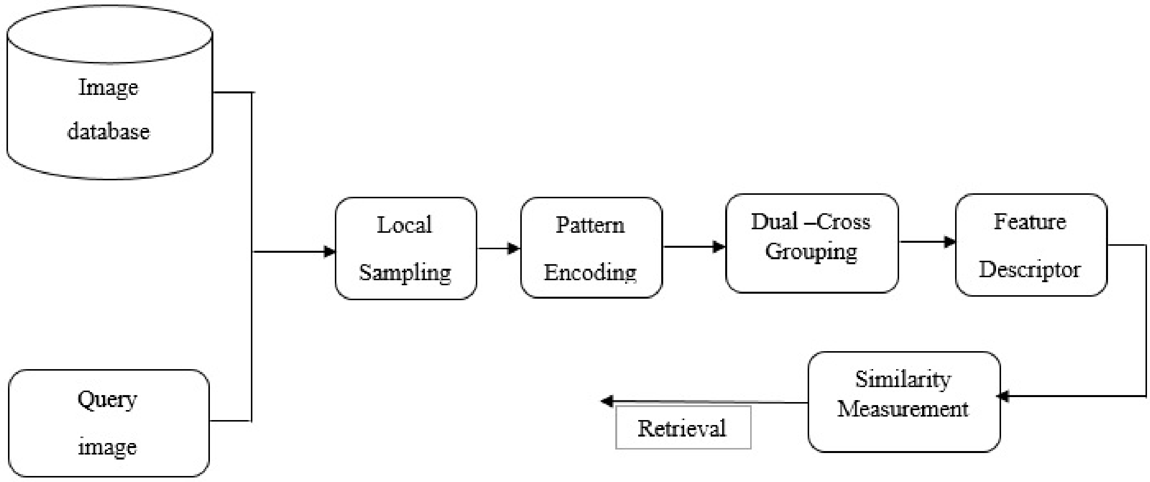


Figure 3.2: Framework of proposed technique.

3.3 Query Matching

F_q is the feature vector obtained from feature extraction for the given query image q, and it is represented as$F_q = \left(F_{q1}, F_{q2}, ..., F_{ql_g}\right)$. Similarly for each image in the database, a feature vector using DCP is derived, and the overall feature vector set F_D of the database $|D|$ is represented as: $F_D = \left(F_{D_{i1}}, F_{D_{i2}}, ..., F_{D_{ilg}}\right)$, $i = 1, 2, ..., |D|$. The objective is to get the *m* top images from the database that have similar categories as that of query image. This is carried out by selecting *m* top-matched images by estimating similarity distance metric d_1 between images in Corel database $|D|$ with the query image using (7).

$$d_1 (q, D) = \sum_{j=1}^{l_g} \left| \frac{F_{D_{ij}} - F_{q_j}}{1 + F_{D_{ij}} + F_{q_j}} \right| \tag{3.1}$$

where $F_{D_{ij}}$ is the *j*th feature of the *i*th image in database $|D|$.

3.4 Experimental Results

3.4.1 Experiment

Simulation is performed in Matlab 8.6 to evaluate the performance of retrieval framework using LTrP and DCP, and Corel database is used for this purpose. It comprises of huge number of different images, contents varying from natural scenes to animals. The database is categorized into multiple categories by domain professionals, each category have 100 images. As it has a collection of diverse images and huge size, researchers have the view that it fulfills all the requirement to assess an image retrieval framework. This experiment used 900 images from Corel database D, and images are collected from nine distinctive domain, namely *Buildings, Elephants, Mountains, Flowers, Buses, Beaches, Food*, and

Figure 3.3: Sample images from database D.

Horses. N_g=100 number of images in each category, with a resolution of either 384×256 or 256×384. Fig.3 shows the samples images from different category of database D. In the simulation work each image is used as a query image from the database D. Framework should retrieve those m images $Y= (y_1, y_2, ..., y_m)$ from the database, which has small and sufficient matching distance from the query image, evaluated using (3.1). Framework retrieval accuracy depends upon the category of the query image and retrieval image y_i=1,2,...,*m* if the category is same then expected image is suitably recognized by it otherwise it has failed to do so.

The simulation performance of presented technique is evaluated in terms of average recall and average precision. The precision for query image Aq is given as follows:

$$Pr(A_q, m) = \frac{1}{m} \sum_{j=1}^{|D|} |\theta\left(\psi\left(A_j\right), \psi\left(A_q\right)\right)| \, Rank\left(A_j, A_q\right) \leq m \tag{3.2}$$

where $\psi(y)$ is the category of "y", "m"represents the number of retrieved images, $|D|$ is the total number of images in the database and *Rank(*A_i, A_q*)* returns the rank of query image A_i among all images of database $|D|$, and

$$\theta\left(\psi\left(A_j\right), \psi\left(A_q\right)\right) = \begin{cases} 1, & \phi(A_j) = \psi(A_q) \\ 0, & else. \end{cases} \tag{3.3}$$

Recall is given as

$$Re(A_q, m) = \frac{1}{N_g} \sum_{j=1}^{|D|} |\theta\left(\psi\left(A_j\right), \psi\left(A_q\right)\right)| \, Rank\left(A_j, A_q\right) \leq m \tag{3.4}$$

The average precision of the *i*th similar category of the reference image database is given by

$$Pr^{i}_{ave}(m) = \frac{1}{N_g} \sum_{j \in N_g} Pr(A_i, m) \tag{3.5}$$

In the similar manner, average recall can be defined.

3.4.2 Result

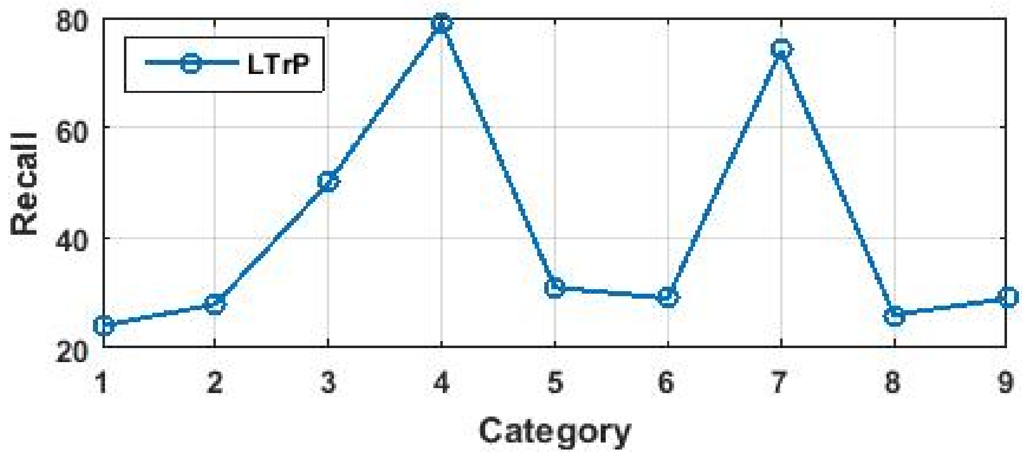


Figure 3.4: Performance of LTrP method on database D, Recall vs. Category

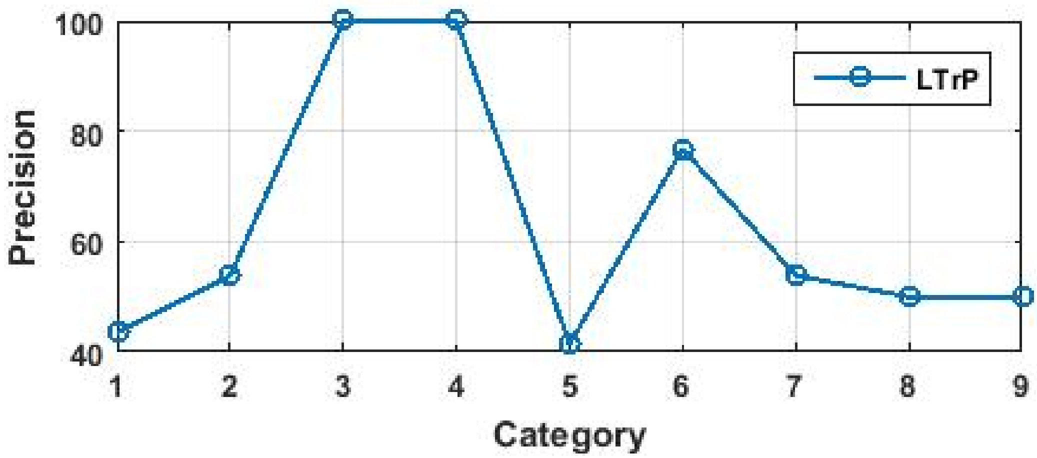


Figure 3.5: Performance the LTrP method on database D, Precision vs. Category.

For Corel database, Fig. 3.3 and 3.4 shows category wise recall and precision performance for the LTrP. Recall and precision respectively values are the useful parameter for quantitative performance measurement of CBIR technique. Performance is supposed to be good if these parameters have high value. Fig. 3.5 and 3.6 shows category wise recall and precision performance respectively for the DCP.

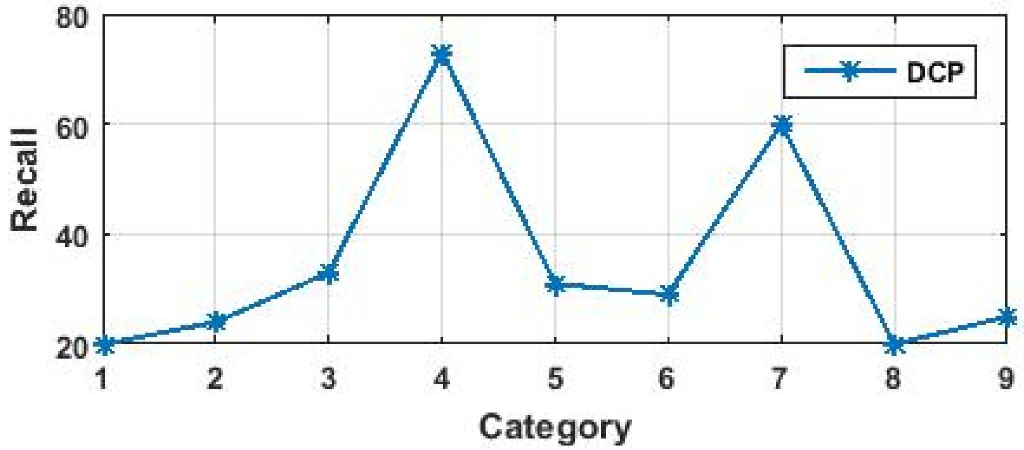


Figure 3.6: Performance of the DCP method on database D, Recall vs. Category .

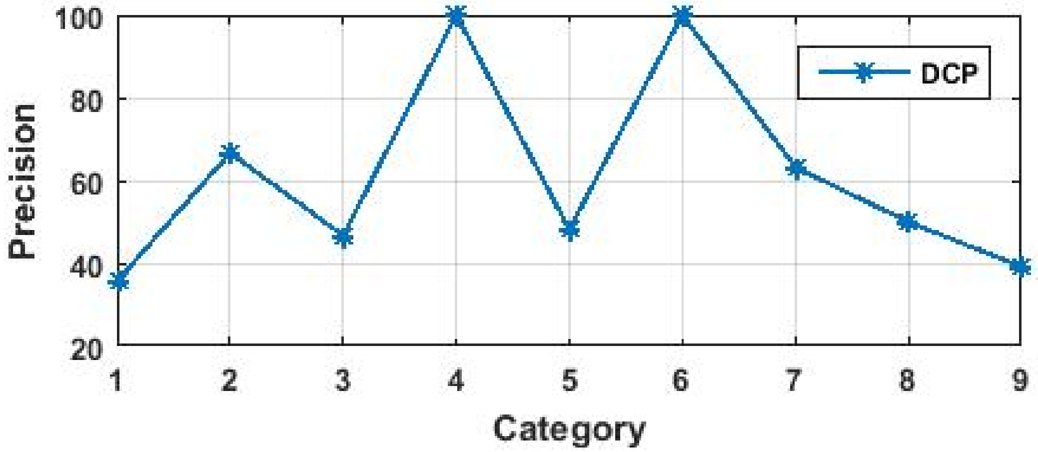


Figure 3.7: Performance of the DCP method on database D, Precision vs. Category .

Fig. 3,7 and 3,8 shows the performance of the DCP, LTrP techniques which is estimated in terms of average recall percentage and average precision percentage respectively on the Corel database. The performance of LTrP technique gives better result compared to DCP technique. Table 3.1 shows computational complexity of LTrP and DCP method as we can see DCP has better computational complexity compared to LTrP. Table 3.2 shows time complexity of LTrP and DCP methods as we can see DCP required very less time compared to other method i.e. DCP has good time complexity.

Table 3.1: Computation Complexity

Method	No of Addition for a pixel	No of Multiplication for a pixel
LTrP	39	54
DCP	14	22

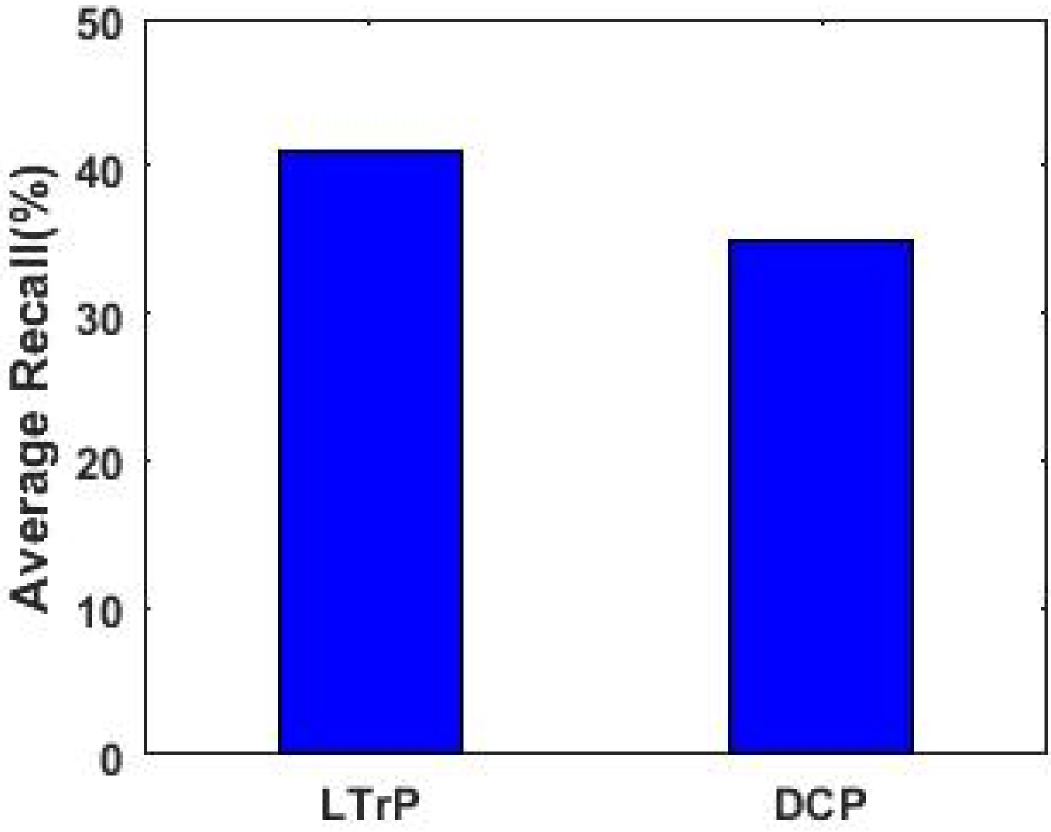


Figure 3.8: Performance comparison of the DCP with LTrP methods on database D, Recall vs. Category .

Table 3.2: Time Complexity

Method	Time for one image
LTrP	1926.48 sec
DCP	34.75 sec

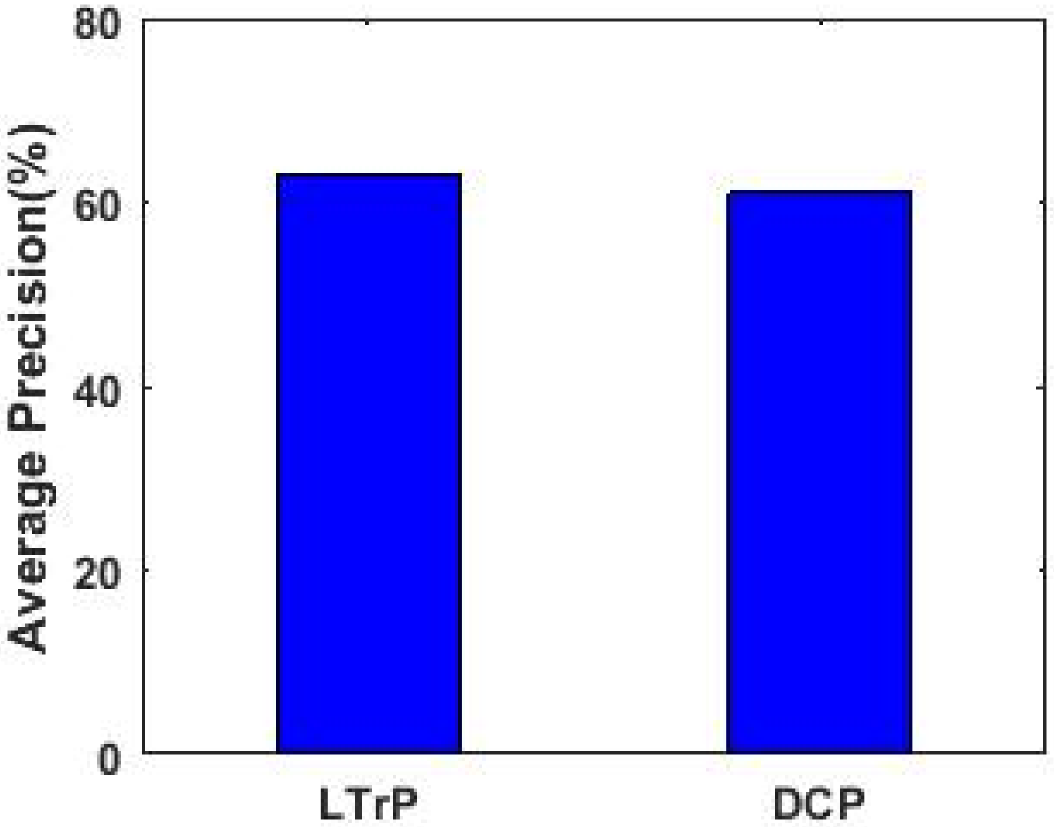


Figure 3.9: Performance comparison of the DCP with LTrP methods on database D, Precision vs. Category .

Chapter 4

Conclusion and Future Scope

In this thesis we presented a novel technique for fast CBIR and evaluate LTrP on CBIR. The essence of dcp is to carry out pattern encoding and local sampling in the most descriptive direction within the image. DCP encrypts second order information using two cross encoders. Local tetra patterns (LTrP) is a method which acquires more detailed information by using four possible directions of every center pixel in an image, and is calculated from first order derivatives in horizontal and vertical directions. The significant improvement of the proposed technique is in Computation Complexity as compared to LTrP methods is shown in Table 1. The average recall for LTrP is 42% and the average precision for DCP is 63% The average recall for DCP is 35% and the average precision for DCP is 61.05%.

As proposed technique has encouraging results, it can be applied to other pattern classification applications like fingerprint recognition, bioinformatics, etc. In the future ,we will try to implement our proposed technique with combination of other descriptors.

References

[1] Rui, Y., Huang, T. S., and Chang, S.-F., 1999. “Image retrieval: Current techniques, promising directions, and open issues”. *Journal of visual communication and image representation,* **10**(1), pp. 39–62.

[2] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., and Jain, R., 2000. “Content-based image retrieval at the end of the early years”. *IEEE Transactions on pattern analysis and machine intelligence,* **22**(12), pp. 1349–1380.

[3] Kokare, M., Chatterji, B., and Biswas, P., 2002. “A survey on current content based image retrieval methods”. *IETE Journal of Research,* **48**(3-4), pp. 261–271.

[4] Liu, Y., Zhang, D., Lu, G., and Ma, W.-Y., 2007. “A survey of content-based image retrieval with high-level semantics”. *Pattern recognition,* **40**(1), pp. 262–282.

[5] Moghaddam, H. A., Khajoie, T. T., and Rouhi, A. H., 2003. “A new algorithm for image indexing and retrieval using wavelet correlogram”. In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, Vol. 3, IEEE, pp. III–497.

[6] Moghaddam, H. A., and Saadatmand-Tarzjan, M., 2006. “Gabor wavelet correlogram algorithm for image indexing and retrieval”. In Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, Vol. 2, IEEE, pp. 925–928.

[7] Ahmadian, A., and Mostafa, A., 2003. “An efficient texture classification algorithm using gabor wavelet”. In Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, Vol. 1, IEEE, pp. 930–933.

[8] Do, M. N., and Vetterli, M., 2002. “Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance”. *IEEE transactions on image processing,* **11**(2), pp. 146–158.

[9] Laine, A., and Fan, J., 1993. “Texture classification by wavelet packet signatures”. *IEEE Transactions on pattern analysis and machine intelligence,* **15**(11), pp. 1186–1191.

[10] Ojala, T., Pietikäinen, M., and Harwood, D., 1996. “A comparative study of texture measures with classification based on featured distributions”. *Pattern recognition,* **29**(1), pp. 51–59.

[11] Ojala, T., Pietikainen, M., and Maenpaa, T., 2002. “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”. *IEEE Transactions on pattern analysis and machine intelligence,* **24**(7), pp. 971–987.

[12] Pietikäinen, M., Ojala, T., and Xu, Z., 2000. “Rotation-invariant texture classification using feature distributions”. *Pattern Recognition,* **33**(1), pp. 43–52.

[13] Guo, Z., Zhang, L., and Zhang, D., 2010. “Rotation invariant texture classification using lbp variance (lbpv) with global matching”. *Pattern recognition,* **43**(3), pp. 706–719.

[14] Guo, Z., Zhang, L., and Zhang, D., 2010. "A completed modeling of local binary pattern operator for texture classification". *IEEE Transactions on Image Processing,* **19**(6), pp. 1657–1663.

[15] Liao, S., Law, M. W., and Chung, A. C., 2009. "Dominant local binary patterns for texture classification". *IEEE transactions on image processing,* **18**(5), pp. 1107–1118.

[16] Li, X., Hu, W., Zhang, Z., and Wang, H., 2010. "Heat kernel based local binary pattern for face representation". *IEEE Signal Processing Letters,* **17**(3), pp. 308–311.

[17] Zhang, B., Gao, Y., Zhao, S., and Liu, J., 2010. "Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor". *IEEE transactions on image processing,* **19**(2), pp. 533–544.

[18] Tan, X., and Triggs, B., 2010. "Enhanced local texture feature sets for face recognition under difficult lighting conditions". *IEEE transactions on image processing,* **19**(6), pp. 1635–1650.

[19] Murala, S., Maheshwari, R., and Balasubramanian, R., 2012. "Local tetra patterns: a new feature descriptor for content-based image retrieval". *IEEE Transactions on Image Processing,* **21**(5), pp. 2874–2886.

[20] Ding, C., Choi, J., Tao, D., and Davis, L. S., 2016. "Multi-directional multi-level dual-cross patterns for robust face recognition". *IEEE transactions on pattern analysis and machine intelligence,* **38**(3), pp. 518–531.

[0]This reference format follows ASME style. You are advised to follow one reference format of any dominant journal of your field.

Dissemination

Conferences [1]

1. **Bhavana Kakde** and Manish Okade, "A Novel Technique for Fast Content-Based Image Retrieval using Dual-Cross Patterns."in *Proc. IEEE-3rd International Conference for Convergence in Technology (I2CT)*,Bombay Section Pune,2018.