Volume 5, Issue 5, May 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Review on Content Based Image Retrieval

Harpreet Kaur, Lakhwinder Kaur
Dept. of Computer Engineering
Punjabi University
Patiala, India

Abstract— CBIR is Content Based Image Retrieval which retrieves the image based on their content rather than keywords. This paper introduces CBIR and its techniques. It gives an overview of currently available literature on CBIR. Different feature extraction techniques like DCT and LBP and different distance metrics like euclidean distance, manhattan distance to find the best match are introduced.

Index Terms— CBIR, Manhattan Distance, Euclidean Distance, LBP, DCT, SVM.

I. INTRODUCTION

CBIR stands for Content Based Image Retrieval. It is the process of searching image from large database using its content. Content can be color, shape, texture or any other spatial features. CBIR method is originated in 1992 by T. Kato. T. Kato used the term CBIR to describe the process of retrieving images based on their content [1]. T. Kato retrieves the image from database using content of the image rather than keywords. Content can be color, shape or texture.

It is growing interest in various fields because it provides appropriate results. CBIR is also known as QBIC (Query By Image Content) because it matches the content of images in the database [3]. Searching the image by its content is efficient method as compared to searching the image using keywords. Although searching the image by keyword is easy but it does no provide efficient results. Content Based Image Retrieval Method provides fast and accurate results as compared to any other retrieval method [1].

Many researchers have developed CBIR techniques. In the early 1990s, because of the emergence of large-scale image collections, the difficulties faced by the manual annotation approach became more and more acute. To overcome difficulties, content-based image retrieval was proposed. One of the most fascinating studies was done on similarity assists image browsing by Rodden in 2001 [2]. The users were asked to illustrate a set of destination guide articles for a travel website. The similarity by visual content view was compared with a text caption similarity view. Other interesting studies have been done on the process of managing personal photograph collections by Rodden and Wood in 2003. In 2004, Worring, et al. gave useful insights into how users apply the steps of indexing, filtering, browsing, and ranking in video retrieval [2].

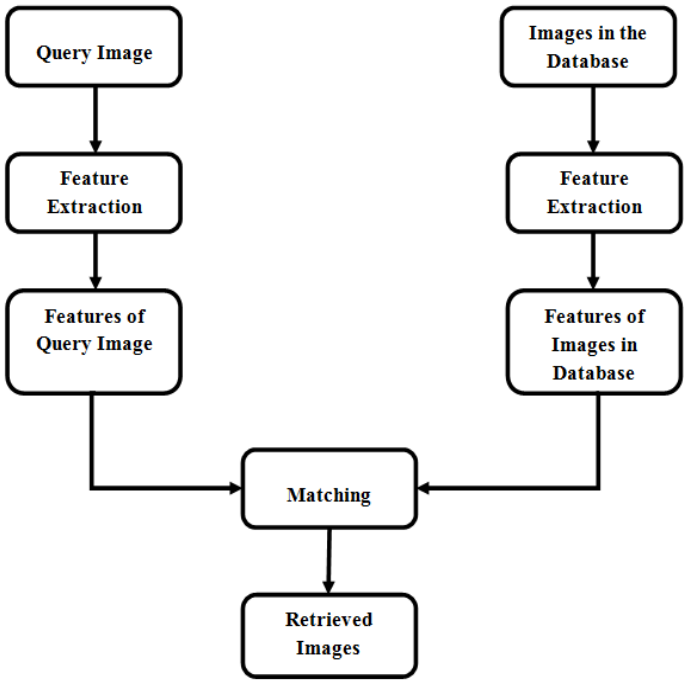

Fig. 1: Block diagram Of CBIR [3].

Steps of retrieving images using the content [4]:

1. Query image is provided by the user in order to retrieve related images from the database.
2. Features are extracted from the query image.
3. Features are also extracted from images in the database.
4. Features of query image are compared with the features of other image present in the database using different distance measures.
5. Best match is provided to the user.

II. IMAGE RETRIEVAL

Image Retrieval is process of retrieving images from the database. There are two different methods by which images can be retrieved: TBIR and CBIR [7].

A. TBIR

TBIR stands for Text Based Image Retrieval. This method is most commonly used method. In TBIR, image is searched using the keywords. User provides the keyword or tags related to the image which he/she wants to search from the database.

Keywords or tags that describe the data are used to search the image that's why this method is also known as metadata.

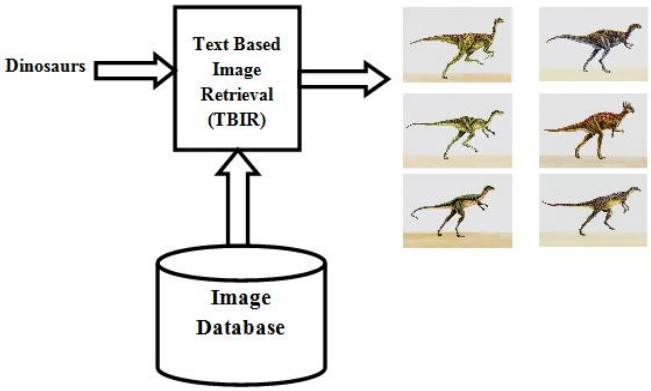

Fig. 2: Text Based Image Retrieval [3]

Sometimes it doesn't provide appropriate results. So it is inefficient method of image retrieval.

Picture Archiving and Communication System (PACS) uses the concept of TBIR. Most Hospital uses PACS that helps to retrieve and organize the medical images [3].

There is need to manually assign keywords to every image present in the database. As there is difference in human perception, so it may lead to inaccuracies while retrieving the image. It also provides a lot of garbage in the result. If database is large, there is also difficultly in manually assigning the keywords. TBIR is inefficient method for retrieving the images from large database [3].

B. CBIR

With the increase in digital images, there is need of appropriate image retrieval method. Content Based Image Retrieval (CBIR) method is process of retrieving the images based on their content.

As Text Based Image Retrieval (TBIR) is proved to be inefficient and provides inaccurate results, CBIR is growing interest in many application areas [3]. CBIR is process of retrieving the images from database using content of image rather than the keywords. There is no need to manually assign keywords in CBIR as content of image is extracted as the feature and match with content of images in the database. Content of the image can be color, shape or texture. CBIR provides efficient results as compared to TBIR. There is growing interest in CBIR because of limitation in TBIR.

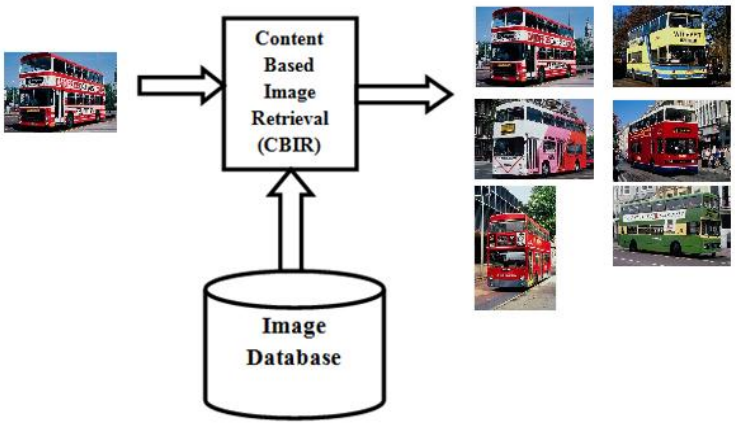

Fig. 3: Content Based Image Retrieval [3]

Query image is provided by the user. Features are extracted from the query image. Features are also extracted from images in the database. Features of query image are matched with features of images in the database. Best match is provided as result to the user. In this way, CBIR provides the result. CBIR provides the fast and accurate results as compared to any other method. As CBIR provides efficient results, it is growing interest in many fields like medical, search engines, crime prevention, and many more [3, 7].

CBIR is also known as QBIC and CBVIR. QBIC stands for Query By Image Content and CBVIR stands for Content Based Visual Information Retrieval [3]. Content Based Image retrieval is application of computer vision to image retrieval problem. It is process of searching the images from the database by passing the query image. Based on the content of query image, images are retrieved from the large database. Search will analyze the content of the image. Content can be color, shape or texture.

CBIR plays an important role in Image Processing techniques. CBIR provides better indexing and efficient results as compared to Text Based Image Retrieval (TBIR). Traditional methods of image indexing have proven to be insufficient and extremely time consuming that's the reasons for the development of CBIR in many large image databases [3]. These old methods of image indexing, ranging from storing an image in the database and assigning it with a keyword or number, to associating it with a categorized description, have become obsolete. It is not done in case of CBIR.

Several CBIR systems currently exist. Some of them are Query By Image Content (QBIC), VIR image engine, Visual SEEK and Web SEEK, NeTra, Multimedia Analysis and Retrieval System (MARS), Visual Information Processing for Enhanced Retrieval (VIPER) [1].

1) Methods of CBIR:

CBIR is the process of searching image by its content like color, shape or texture [4].

a) Color

Method of retrieving the image on the basis of its color is widely used. In this method, color histogram of each image in the database is computed. At the time of searching, color histogram of query image is also computed. Color histogram of query image is match with the color histogram of images in the database. Then those images are selected whose color histogram matches most closely with the histogram of query image. In this way, images are retrieved by calculating color histogram.

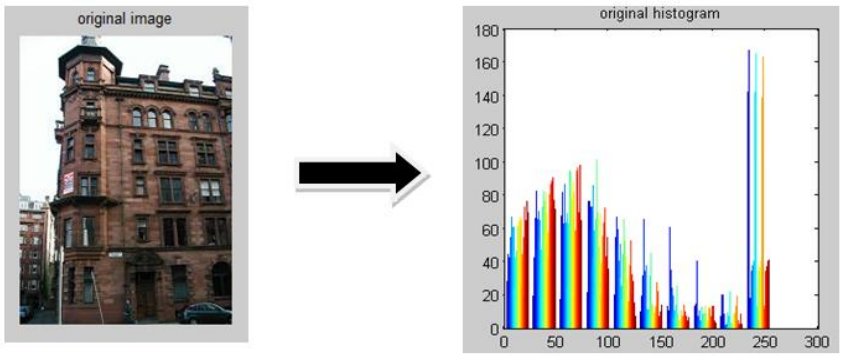

Fig. 5: Color Histogram of image

Color features method is widely used method for image retrieval. This method is easier as compared to shape and texture because color information is easy to extract as compared to texture and shape information. Extracting color features can be completed without any concern to image size or orientation. Color difference histogram is also used to retrieve the images from the database.

b) Shape

Shape does not refer to shape of image but it refers to shape of particular region. Shape can be determined by applying segmentation or edge detection. Shape filters are used by many methods to identify shapes. Some shape descriptors are Fourier transform and Moment Invariant.

There are two types of shape features. First feature is global feature such as aspect ratio. Aspect ratio can be defined as proportional relationship between width and height. For example, if aspect ratio is 16:9 then width is 16 and height is 9. Second feature is local features such as set of consecutive boundary segment. Shape representation can be divided into two categories. Categories are boundary based and region based.

Fig. 6: Boundary Based shape representation [5].

In boundary based, shape can be represented only by outer boundary. Boundary based describes the region using external characteristics such as pixel along outer boundary.

Fig. 7: Region Based Shape Representation [5].

In region based, entire shape region is represented using internal characteristics such as pixel contained in that region.

c) Texture

To find visual patterns in the image, texture method is used. This method helps to distinguish between images having same color such as sky and sea. Important information about the structural arrangement of the surface, like clouds, leaves, bricks, fabric, etc. can be extracted from texture. Texture also provides the information that describes relationship of surface to surrounding environment.

Fig. 8: Different textures [6].

Texel is used to represent the texture. Texels are placed into no. of sets which also tells about the where in the image texture is located. Many techniques are based on texture for image retrieval. Measure of texture is computed such as degree of contrast, coarseness, directionality and regularity. Gabor filter is used for texture.

III. APPLICATIONS

CBIR is growing interest in many fields. Application areas of CBIR are as follows [7]:

A. Crime Prevention

Records of criminals like photograph, fingerprint with other information is maintained by the police and investigating teams. These records are then used for face recognition and fingerprint matching. To identify any person or to find if person is having the criminal record in past or not, these records are very useful and helpful. Using CBIR, new records are matched with previous records to find the information related to criminals. In UK, this method is commonly in practice. Even Metropolitan Police Force in London involved with a project which is setting up an international database of images of stolen objects.

B. Medical Diagnosis

In medical field, visual information such as X- Rays, ultrasound, and scanned images with other details of patient is stored in the database. These records or images are helpful in finding the treatment with help of previous cases. These records are also useful in teaching and research. With the help of these records, researchers can even find the treatment of new diseases. That's why CBIR is growing interest in medical field also.

C. Military

Identification of targets from satellite photographs, recognition of enemy aircraft from radar screens, identification of targets from satellite photographs and guidance systems provision are military applications of imaging technology.

D. Search Engines

Nowadays Search engines like Google also uses the concept of CBIR for searching the image.

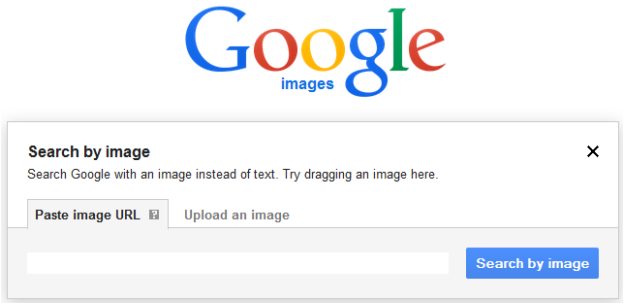

Fig. 4: Google image search [8].

E. Fashions and Graphic Design

Images play an important role in fashion and graphic design. Designer used images of previous design to finalize the final product. Designer also used it to present the ideas to clients and colleagues.

F. Cultural heritage

There are many visual objects in the Museums and art galleries. Ability to identify the objects that share visual similarities can be useful to researchers as well as art lovers and that can be done with the help of concept of CBIR.

G. Video Retrieval

Large videos can be breakup into individual shots. We can extract the features from the individual shots and then we can search for video clips containing specified types of movement. That can be done with the help of CBIR.

H. Publishing and Advertising

Many newspaper publishers maintain their photograph library. Photographs can be retrieved using CBIR.

IV. DISTANCE MEASURES

After feature extraction, features are matched with those in the database to find the nearest match. Distance measures used to find match would be Manhattan distance or Euclidean Distance [9].

A. Manhattan Distance

Manhattan distance is represented by L1. It is simple sum of horizontal and vertical components. To calculate the manhattan distance between two points, grid like path is followed from one data point to another. Manhattan distance between two points is sum of difference of their corresponding components [9, 17].

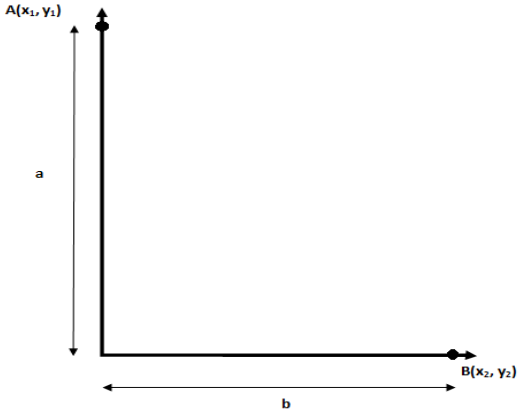

Fig. 9: Manhattan Distance [9].

Manhattan distance between two points A and B is equal to sum of a and b.

The formula for manhattan distance between a point ***X***= (***X1***, ***X2***, ***X3*** etc.) and a point ***Y***= (***Y1***, ***Y2***, ***Y3***, etc.) is:

$$d = \sum_{i=1}^{n} |x_i - y_i|$$

Where *Xi* and *Yi* are the values of the i[th] variable, at points *X* and *Y* respectively and n is the number of variables [17].

If u = (x_1, y_1) and v = (x_2, y_2) are two points, then the Manhattan Distance between u and v is given by

$$MH(u, v) = |x_1 - x_2| + |y_1 - y_2|$$

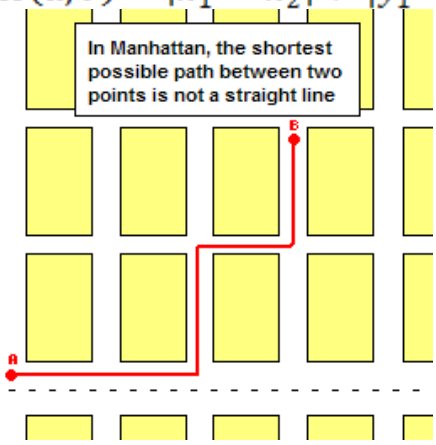

Fig. 10: Representation of Manhattan Distance [10].

Manhattan distance is also known as City Block Distance. Distance is calculated as horizontal plus vertical component which is same as the way you move in the city like manhattan where you have to move around the building instead of going straights [17].

B. Euclidean Distance

Euclidean distance is represented by L2. Euclid states that shortest distance between two points on a line is a straight line. Euclidean distance is a non-parametric classifier [17].

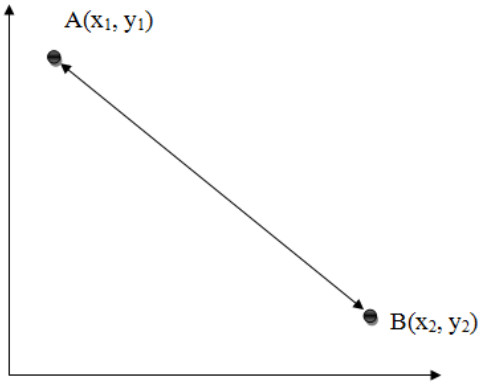

Fig. 11: Euclidean Distance [17]

Euclidean distance between two points A and B is

$$EU(u,v) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Shortest distance between two points is Euclidean Distance [17].

V. CLASSIFICATION METHOD

A. SVM

SVM stands for Support Vector Machine. SVM was invented by V. Vapnik and his co-workers in 1970s in Russia and became known to west in 1992 [11]. SVM is a linear classifier that finds the hyperplanes to separate two classes of data i.e. positive and negative. SVM is supervised learning technique in which learning is based on learn by results. SVM analyze the data and identify patterns used for classification [12].

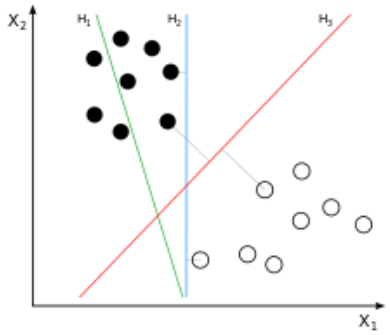

Fig. 12: Different hyperplanes [11].

H_1, H_2 and H_3 are three different hyperplanes. In thefigure, Hyperplane H_1 does not separate the classes. Hyperplane H_2 separate the classes but only with a small margin. Hyperplane H_3 separates them with the maximum margin [12]. If hyperplane is having largest distance from the nearest point then good separation is achieved by those hyperplanes. Here good separation means larger the separation between two hyper planes gives lower generalization error. That's by it is called maximum margin classifier. Two parallel hyper planes are constructed on each side of the hyper planes that separate the data. If geometric gap between hyperplanes is high than classification error is low.

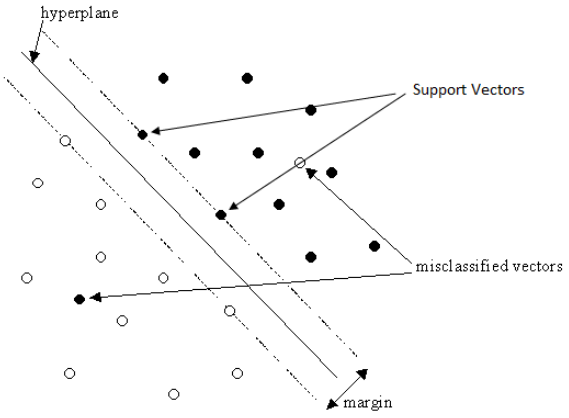

Fig. 13: Hyperplanes, Support vectors [11].

SVM is used as a classifier for classification of images. Classification process is applied to features that are extracted from feature extraction step. It is done to find maximum margin hyperplanes in high dimensional feature space.With the help of SVM, we get much more performance and optimal results as compared to traditional methods. In searching phase, given feature database is represented as two set of vertices in n-dimensional space and constructs a hyperplane that max margin between image relevant to query and images non-relevant to query [9].

Consider a given set of points in the form of training data which separate two classes of pattern based on given training set:

$$S= \{(x_1, y_1), (x_2, y_2), (x_3, y_3) \dots (x_i, y_i)\}$$

Where x_i is a p-dimension real vectors, $y_i = \{-1, +1\}$ and n is a number of sample.

According to Vapnik's formula

$$y_i = \langle w.x \rangle + b$$
$$= \sum w_i x_i + b$$

Where w is normal vector to the Hyperplane and b is constant or scalar [11].

The parameter $\frac{b}{\|w\|}$ determines the offset of the hyperplane from the origin along the normal vector w.

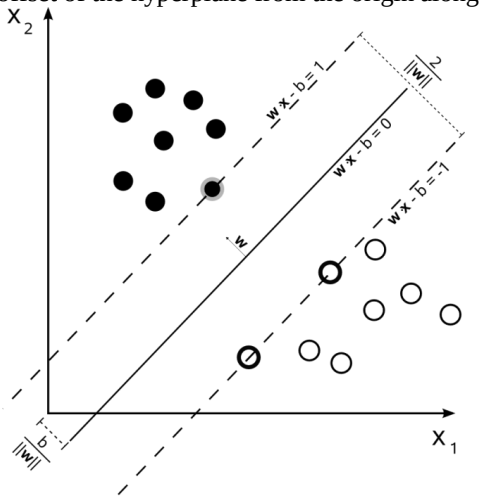

Fig. 14: Hyperplanes. [12]

By adding a scalar value b it increases the margin between hyperplanes and in the absence of b hyperplanes is forced to pass through the origin. So in SVM we always use parallel hyperplanes which maintain distance between them.

VI. FEATURE EXTRACTION METHOD

A. DCT

DCT stands for Discrete Cosine Transform. DCT is defined as a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. The discrete cosine transform (DCT) helps separate the image into parts with respect to the image's visual quality. The DCT is similar to the discrete Fourier transform (DFT). When DCT is applied, it transforms a signal/image from the spatial domain to the frequency domain. Discrete Cosine Transform is a Fourier-related transform. It is similar to the discrete Fourier transform (DFT), but DCT use only real numbers [16].

Discrete Cosine Transform is defined as [19]:

$$F(u,v) = \frac{1}{\sqrt{MN}} \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \times \cos\left(\frac{(2x+1)u\pi}{2M}\right)$$
$$\times \cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad u = 0, 1, \dots, M, \quad v = 0, 1, \dots, N$$

Where,

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}, & u = 0 \\ \sqrt{\frac{2}{M}}, & u = 1, 2, \cdots, M-1 \end{cases}$$
$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}, & v = 0 \\ \sqrt{\frac{2}{N}}, & v = 1, 2, \cdots, N-1. \end{cases}$$

Discrete Cosine Transform is used in signal processing, image processing and in lossy data compression.

The DCT coefficients of an image represent the regularity, complexity and some texture features of an image and it can be directly applied to image data in the compressed domain. As DCT is applied directly to image data in compressed domain, it may solve the large storage space problem and the computational complexity of the existing methods [16].

Steps to apply DCT on image [19]:

1. Extract Red, Green and Blue components of the color image and divide each into block of size 8*8.
2. Apply the DCT Transform on each block of individual color planes of image to extract feature vector.
3. The left topmost component (DC term) of each block of each color components i.e. R, G and B are taken as an image feature and stored in feature vector.
4. The Euclidian distances between the feature vectors of query image and the feature vectors of images in the database are calculated.

B. LBP

LBP stands for Local Binary Pattern. A local binary pattern is a type of feature used for classification in computer vision. It is the particular case of the Texture Spectrum model. Model was proposed in 1990. LBP was first described in 1994. Local Binary Pattern is a powerful feature for texture classification [15].

The local binary pattern operator is an image operator which transforms an image into an array or image of integer labels describing small-scale appearance of the image. These labels or their statistics, most commonly the histogram, are then used for further image analysis. It operated with eight neighboring pixels using the center as a threshold. The final LBP code was then produced by multiplying the threshold values by weights given by powers of two and adding the results [22].

It divides the query image and database images into equally sized blocks. After diving images into block, LBP histograms are extracted. Then the block histograms are compared.

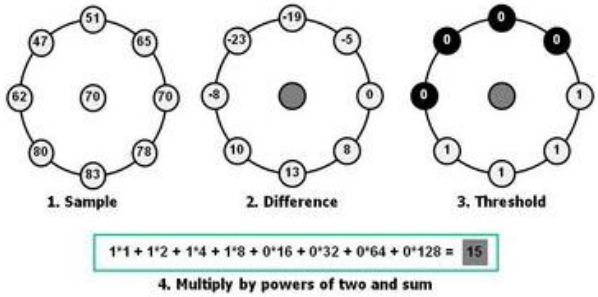

Fig 15: Local Binary Pattern [13]

Steps to create Local Binary Pattern [15]:

1. Image is divided to blocks. For example: 16*16 pixels for each block.
2. For each pixel in block, compare pixel with its 8 neighbors.
3. When the center pixel's value is greater than that of neighbor's pixel value, write 1. Otherwise, write 0.
4. After that we get 8-digit binary number.
5. Compute the histogram of the frequency of each number occurring over that block.
6. Normalize the histogram.
7. Concatenate histograms of all blocks. Then we get feature vector for the image.

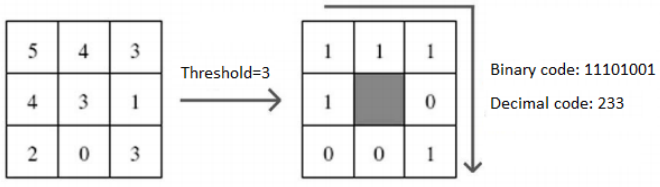

Fig. 16: Local Binary Pattern

VII. CONCLUSION

This paper introduces CBIR and its techniques. CBIR is an important research topic covering a large no. of research domains. In future, we will try to implement these techniques and compare the results of these techniques with each other.

REFERENCES

[1] Wikipedia, "Content Based Image Retrieval", http://en.wikipedia.org/wiki/Content-based_image_retrieval.

[2] Michael Lew, et al., "Content-based Multimedia Information Retrieval: State of the Art and Challenges", ACM Transactions on Multimedia Computing, Communications, and Applications, pp. 1–19, 2006.

[3] M.NARAYANA**,** SUBHASH KULKARNI, "Comparison between Euclidean Distance Metric and SVM for CBIR using Level Set Features", International Journal of Engineering Science and Technology, Vol. 4, No.01, January 2012.

[4] Eakins, John; Graham,Margaret. "Content-based Image Retrieval". University of Northumbria at Newcastle. Retrieved 2014-03-10.

[5] Symmetry-Based Shape Representations, http://www.lems.brown.edu/vision/Presentations/Kimia/IBM-Oct-99/talk.html.

[6] Google Images, "Texture", https://www.google.co.in/search?site=imghp&tbm=isch&source=hp&biw=1280&bih=710&q=texture&oq=texture&gs_l=img.3..0l10.3120.5109.0.5338.7.7.0.0.0.0.315.1074.2-3j1.4.0.msedr...0...1ac.1.61.img..3.4.1071.YseltdAg9x4.

[7] Jagpal Singh, Jashanbir Singh Kaleka and Reecha Sharma, " Different Approaches of CBIR Techniques", International Journal of Computers & Distributed Systems, Volume 1, Issue 2, August, 2012.

[8] Google, "Google Images", https://www.google.co.in/imghp.

[9] Meenakshi Madugunki, Dr. D.S.Bormane, Sonali Bhadoria and Dr. C. G. Dethe, "Comparison of Different CBIR Techniques", IEEE, 2011.

[10] Joachim Despland, "Pathfinding for the Mammoth Project", http://www.joachimdespland.com/mammoth.html.

[11] Wikipedia, "Support Vector Machine", http://en.wikipedia.org/wiki/Support_vector_machine.

[12] Sonali Jain, "A Machine Learning Approach, SVM for Image Classification in CBIR", International Journal of Application or Innovation in Engineering & Management, Volume 2, Issue 4, April 2013.

[13] Scholarpedia, "Local Binary Patterns", http://www.scholarpedia.org/article/Local_Binary_Patterns.

[14] Yogita Mistry, Dr. D.T. Ingole, "Survey on Content Based Image Retrieval Systems", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 1, Issue 8, October 2013.

[15] Wikipedia, "Local Binary Patterns", http://en.wikipedia.org/wiki/Local_binary_patterns.

[16] Wikipedia, "Discrete Cosine Transform", http://en.wikipedia.org/wiki/Discrete_cosine_transform.

[17] Gunjan Khosla, Dr. Navin Rajpal and Jasvinder Singh, "Evaluation of Euclidean and Manhanttan Metrics In Content Based Image Retrieval System", Int. Journal of Engineering Research and Applications, Vol. 4, Issue 9(Version 1), September 2014, pp.43-49.

[18] Dr. A. Kannan, "Content Based Image Retrieval Using Nearest Neighbour and Hybrid NNSVM Methods to Diagnose MR Images", International Journal of Innovative Research in Science, Engineering and Technology, Volume 3, Special Issue 3, March 2014.

[19] Vibha Bhandari, Sandeep B.Patil, "CBIR Using DCT for Feature Vector Generation", International Journal of Application or Innovation in Engineering & Management, Volume 1, Issue 2, October 2012.

[20] K. Ashok Kumar, Y. V. Bhaskar Reddy, "Content Based Image Retrieval Using SVM Algorithm", International Journal of Electrical and Electronics Engineering (*IJEEE*), Vol-1, Iss-3, 2012.

[21] C. S. Sastry, Saurabh Jain, and Ashish Mishra, "Application of *l*1-norm minimization technique to image retrieval", International Science Index, Volume 3, No, 8, 2009.

[22] Valtteri Takala, Timo Ahonen, and Matti Pietik¨ainen, "Block-Based Methods for Image Retrieval Using Local Binary Patterns", Springer-Verlag Berlin Heidelberg, 2005.

[23] Oana Astrid V AT AMANU, Mihaela IONESCU, Mirela FRANDES, Simona APOSTOL, "Content-Based Image Retrieval using Local Binary Pattern, Intensity Histogram and Color Coherence Vector", IEEE, 2013.

[24] C.Callins Christiyana, V. Rajamani, "Comparison of Local Binary Pattern Variants for Ultrasound Kidney Image Retrieval", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 10, October 2012.

