Abstract—Paper presents a survey on the technical achievements of Content Based Image Retrieval (CBIR), this area has been active from past two decades. The survey includes the information of local patterns covering the research aspects of image feature extraction which is the most fundamental and prominent base of CBIR. Furthermore, from current technology available and demand from the real world applications such as face recognition, medical diagnosis etc., and some research directions are identified and suggested.

Keywords—CBIR(Content Based Image Retrieval); Local Patterns;

I. INTRODUCTION

These days the internet culture gaining the observable recognition due to dramatic enhancement of the digital image databases, which makes tedious task to manage the huge databases using human annotations. Image processing and retrieval plays a significant role in important fields for the society such as medical diagnosis, education, entertainment, security, military applications etc. areas due to large awareness of internet to the people which creates bulk image data like online shopping, uploading, graphics and entertainment etc. so that maintaining that bulk data is difficult. To overcome these problems an automatic, efficient, smart system is needed named content based image retrieval. The prominent methods of content based image retrieval are feature extraction, similarity measurement, relevance feedback and image acquisition etc. Content based image retrieval works based on the content of an image like color, texture, shape, spatial layout, faces, biometric etc[3][4][5]. but user may take the photographs in different conditions and premises (view, angle, illumination, poses, and background etc.) so general features are not sufficient to retrieve the relevant images. Furthermore, the features are classified into two categories first, low – level features and high – level features. Low level features are related to the basic contents of an image, same color different images may get the common histogram, so it demands the high level semantic conceptual techniques such as neural networks, fuzzy logic and machine algorithms etc [6][7] are required to recognize and retrieve the addition and exact features like emotional expressions etc. it is a challenging task for content based image retrieval.

Early on years due to the appearance of enormous data repositories some difficulties are placed by human interpretation. To overcome those problems Content Based Image Retrieval was proposed. Instead of human interpretation by text based images were indexed based on their content such as color, texture, shape, faces etc. for past two decades several techniques have been implemented for image retrieval for different applications but still need a enhanced expert system to retrieve the effective images from huge repository. So there is a need of survey on what techniques has made-up and what are the research instructions is lead to different applications. The motivation is to get the best technique to be used in further image retrieval inventions and applications.

II. RELATED WORK

A. Feature Extraction

In Content Based Image Retrieval, there are three essential steps are involved such as feature extraction, indexing and retrieval. Feature extraction is the basic and most vital step in image retrieval. In feature extraction, features can be retrieved based on the either text based interpretation or content based as color, texture, shape, faces etc... in indexing algorithms and techniques are required to search an image from huge amount of repository faster and retrieval of image is done by applying the feature extraction methods by comparing the query image with the database images using the similarity measurement techniques.

B. Color

The color symphony of image can turn out to be a vibrant feature for the purpose of content based image retrieval(CBIR). Katro [1] used the color of every corresponding pixel in two images for assessment and the number of equivalent pixels having the same color to verify the similarity between them. Swain and Ballard[2] proposed the first color histogram, which solves this sensitivity problem. In that color histograms were extracted and histogram intersection method was utilized for comparing two images. Since this method is quite easy to apply and gives realistic results especially in small to medium sized
repositories, several other histogram based Color Structure Descriptor (CSD) [3] is also based on color histogram, but it provides a exact color description by recognizing contained color distributions of each color. Furthermore color structure in visual scenery is vigorous to sound, image degradation changes in size, resolution and direction. Eventually most of the existing CBIR systems use various color descriptors in order to retrieve related images however their retrieval performance is usually restricted especially on large repositories due to lack of favoritism of such color descriptors. Unlike the conventional color histograms, CSD is extracted by accumulating from table sized structuring window. The image scanned and CSD counts the number of times a particular color is appeared within the structuring window. An excellent review and an proficient illustration of color histogram based on Karheumen-Loeve-Transform(KLT) which can be found in Tran and Lenz[4]. The primary feature of such histogram based color descriptors are RGB or HSV is to cluster the pixels into fixed color pitcher, which are quantizing the entire color space using a predefined color pitcher. This twofold approach clustering all the features having similar color and reducing the color level from bulk via quantization, is the main reason behind the inadequate success thus the color histograms achieved since the both operations are indeed small steps for obtaining the perceivable elements. Their performance is still quite limited and usually degrades drastically in large repositories.

Content Based Image Retrieval is the main source on untrue retrievals, which makes accidental matches between images with similar inclusive color properties but different in the color distribution model. There are numerous approaches are there to address those drawbacks, Segmentation methods may be an alternative among many, however, they are not practicable in mainly automatic segmentation is an ailing problem; therefore, it is not consistent and vigorous for applications on large repositories. Pass [5] introduced color coherence vectors (CCV), CCV partitions each histogram pitchter into two types, logical, if it belongs to a large uniformly colored region or illogical, if does not. Huang used a color characteristic called color correlogram [6] which characterizes not only the color distribution pixels, but also spatial association of pair of colors. Sticker used first three innermost moments called mean, standard deviation, variance of each color for image retrieval[7]. Lu proposed color feature based on vector quantized indexed histograms in the Discrete Cosine transform (DCT) field they computed 12 histograms, 4 for each color component from 12 index sequences[8]. Aura proposed the color similarity extraction by quantizing on color space. Urdiales [9] proposed a novel algorithm to extract a color graph from a actual image. This proposal had two main innovations. In color field the leading colors are extracted along with their large-scale properties and tree symmetry to partition the image to characterize the spatial color distribution. A new Principle Component Analysis(PCA) based color feature vector has been defined.

C. Texture
Texture retrieval is the stem of texture analysis is one of the most important characteristic of an image. Texture analysis has been broadly used in CBIR systems due to its prospective value. Texture analysis and retrieval has gained extensive attention in the field of medical, industrial, document analysis and several more. Various techniques have been proposed for texture study, such as, automated binary texture feature, the feature vectors are constructing using spatial area information. Another possibility is the use of distorted domain data to extract some higher-level features [10]. Wavelet oriented methods, which provide space–frequency decomposition of the image, have been used. Daubechies’ wavelets are the most regularly used in CBIR for their fast working out and reliability. Wavelets are used to extract feature vectors. Although common wavelet methods such as SIMPLicity[11] or wavelet correlogram [33] allow for a multi resolution decomposition, they have restricted oriental selectivity and are not able to retain capricious directional information. To overcome this shortcoming, introduced other multi resolution and multidirectional image disintegration techniques such as 2D Gabor transform [11, 12], discrete contour let transform [12], steerable pyramid [13], ridgelet transform [14], curvelet transform [15] etc.. Kokare et al. [16] proposed a texture image indexing retrieval method using two-dimensional rotating wavelet filters, which could improve classification of diagonally oriented textures. Among these techniques, the 2D Gabor wavelet has been shown to provide indexing features with analogous or better average retrieval performance with respect to other multidirectional wavelet decompositions. The features were making use of mean and variance of the wavelet coefficients and no post-processing has been made before using them in retrieval.

D. Shape
Shape is a well defined term in CBIR, which supports for additional and accurate feature extraction. This is feature which logically distinguish the images. There are two major features of the shape: Global features (like aspect ratio) and local features (like boundary segments). Shape of an image can be plotted using area, perimeter, radiuses, skeleton, statistical moments, form signature, Fourier and Hough contour signature etc. Shape feature extraction is based on mainly boundary based, region based and contour based. In region based shape extraction regional boundaries which are invariant to formations as shape descriptor, this technique allows the segmentation on object and apply the techniques such as Angular Radial Transformation (ART), shape indexing methods etc. on sub regions. In contour based extraction after segmentation each sub region has a separate description has involves the highest peak and ‘x’ and ‘y’ positions of the sub objects.

III. LOCAL PATTERNS
A. LBP
The Local Binary Pattern (LBP) operator was commenced by Ojala et al. [17] for texture classification. Given a center pixel in an image, the LBP value is calculated by comparing its gray scale color values with its neighbors. LBP is a texture based feature extraction and also applied to the face recognition, medical diagnosis and biometric fields.

\[
LBP_{P, R} = \sum_{i=1}^{P} 2^i \times f_1(g_i - g_c)
\]
f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)

where \( g_e \) is the gray scale value of the center pixel, \( g_p \) is the gray scale value of its neighbor pixels, ‘P’ is the number of neighbor pixels and R is the radius of the neighborhood pixels.

If we set \( P = 8; R = 1 \), we obtain LBP\(_{8,1} \), which is similar to the LBP operator proposed in [17]. The two differences between LBP\(_{8,1} \) and LBP are: 1) The pixels in the neighbor set are categorized that they form a circular chain according to his proposal and 2) the gray values of the diagonal pixels are resolute by exclamation. Both modifications are necessary to obtain the circularly equivalent neighbor set, which allows for deriving a rotation invariant version of LBP\(_{P,R} \).

Subrahmanyam [18] have proposed the Local Maximum Edge Binary Patterns (LMEBP) for an image extraction and object tracking applications. In proposed Local Maximum Edge Binary Patterns for a given image the first most greatest edge is obtained by the degree of local difference between the center pixel and its eight neighborhood pixels. Local maximum binary patterns are following

\[ d(g_p) = d(g_e) - d(g_p) \] where \( p = 1, 2, \ldots, 8 \) \quad (3)

Snehase, Mukherjee, proposed magnitude and directions are used as weighted votes into local orientation histograms[17].

**B. LTP**

Local Ternary Patterns, in which gray color levels in a region of width +t or -t around are quantized to zero, greater is quantized to +1 and lower is quantized to -1, i.e. the indicator is replaced by a 3-valued function:

\[ f(m, p_c, t) = \begin{cases} 1, & m \geq p_c + t \\ 0, & |m - p_c| < t \\ -1, & m \leq p_c - t \end{cases} \quad (4) \]

In this manner, pixels are having three values will take one of those for each threshold. Neighboring pixels are collective after threshold in a ternary pattern [19]. Compute the histogram of these ternary values will result in a great collection, then the ternary pattern is split into two binary patterns by substitute with 1 with 1′s, -1 with 1′s and 0′s are leave as 0′s only in pattern. Histograms are combined to build the feature descriptor twofold over the number of LBP.

**C. LEP**

Local Extrema Patterns (LEPs) in [20] has proposed by subrahmanyam murala, LEP depicts the spatial layout of the local texture using the local extrema of center gray pixel \( c_g \). Specified a center pixel in the 3x3 pattern, LEP value is computed by match up to its gray scale values with its neighborhood pixels in 0º, 45º, 90º, and 135º directions.

**D. LTrP**

The Local Tetra Pattern (LTrP) concept proposed in [21] by Subrahmanyam murala describes the spatial local structure of the local texture using the direction of the center grayscale color pixel. For image, the first-order derivatives along with axis 0º (x - axis) and axis 90º (y - axis) directions are denoted as \( I_x(g_e) = 0º, 90º \). Let \( g_c \) designated as the center pixel in, and let \( g_x \) and \( g_y \) designated as the horizontal and vertical neighborhood pixels of center pixel \( g_e \) respectively. Then, the first-order derivatives at the center pixel \( g_e \).

\[ I_{0x}(g_e) = I(g_x) - I(g_e) \quad (5) \]

\[ I_{90x}(g_e) = I(g_y) - I(g_e) \quad (6) \]

Local pattern orientations are determined based on four directional planes such as (x, y), (-x, y), (-x, -y), (x, -y) areas based on center pixel with respect to horizontal and vertical pixels for neighbor pixels initially, consequent in an LTrP for direction ‘1’ of the center pixel. The LTrP is encoded to ‘0’ either it is equal to the direction of center pixel, otherwise implied in the direction of neighbor pixel. Using the same resemblance, similarly LTrPs are computed for center pixels which are having directions 2, 3, and 4 respectively.

Additionally, the remaining bits of the LTrP for seven neighbors are computed ensuing the tetra pattern for example “4 3 4 0 3 2 0” after encoding the tetra pattern, we split it into three binary patterns as follows. Pass on to the built LTrP, the first pattern is obtained by substituting “1” where the tetra pattern value is “2” and “0” for other values, as “0 0 0 0 1 0.” likewise, the other two binary patterns “1 0 1 0 1 0 0” and “11010000” are calculated for tetra pattern values “3” and “4” respectively. In the same approach, for directions 2, 3, and 4 also tetra patterns for center pixels are computed. Therefore, with four tetra patterns, 12 (4 * 3) in four directions and three binary patterns are gained. Histograms are combined to generate a feature descriptor twelve times (12 X LBP) in excess of the number of LBP patterns.

**E. LDP**

The DBC (Directional Binary Code) is proposed to encode the directional edge values are in a 8 neighborhood pixel positions. Given an image I, it designate the first order derivatives by the direction of 0º, 45º, 90º, and 135º directions as \( I_{d,1} \), where \( \theta = 0º, 45º, 90º, \) and 135º and ‘1’ is the distance between the selected pixel and its neighborhood pixel values. For example, the distance between the center pixel and its four directional neighbor pixels is ‘1’ i.e. in four directions. Let \( x_{ij} \), be a point in I, then the four directional derivatives at \( x_{ij} \).

**F. LGMEPOP**

LGMEPOP utilizes the first three most important (maximum) edge positions in an octal code production. Then, these three greatest edge positions are determined into three-eights octal numbers to produce the LGMEPOP. Further, the LGMEPOP is categorized into two categories which are named as Sign Maximum Edge Position Octal Patterns (SMEPOP) and Magnitude Maximum Edge Position Octal Patterns (MMEPOP). The SMEPOP and MMEPOP are implied based on the sign and magnitudes of leading edges respectively.

\[ LGMEPOP = SMEPOP / MMEPOP \]
G. Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>Local Binary Pattern</td>
</tr>
<tr>
<td>LTP</td>
<td>Local Ternary Pattern</td>
</tr>
<tr>
<td>LEP</td>
<td>Local Extrema Pattern</td>
</tr>
<tr>
<td>LTrP</td>
<td>Local Tetra Pattern</td>
</tr>
<tr>
<td>LDP</td>
<td>Local Directional Pattern</td>
</tr>
<tr>
<td>LGMEPOP</td>
<td>Local Gabor Maximum Edge Position Octal Pattern</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>256</td>
</tr>
<tr>
<td>LTP</td>
<td>2*256</td>
</tr>
<tr>
<td>LEP</td>
<td>13*256</td>
</tr>
<tr>
<td>LTrP</td>
<td>24*501</td>
</tr>
</tbody>
</table>

REFERENCES


