An Appraisal of Content-Based Image Retrieval (CBIR) Methods

ABSTRACT

Content Based Image Retrieval (CBIR) is an aspect of computer vision and image processing that finds images that are similar to a given query image in a large scale database using the visual contents of images such as colour, texture, shape, and spatial arrangement of regions of interest (ROIs) rather than manually annotated textual keywords. A CBIR system represents an image as a feature vector and measures the similarity between the image and other images in the database for the purpose of retrieving similar images with minimal human intervention. The CBIR system has been deployed in several fields such as fingerprint identification, biodiversity information systems, digital libraries, Architectural and Engineering design**,** crime prevention, historical research and medicine. There are several steps involved in the development of CBIR systems. Typical examples of these steps include feature extraction and selection, indexing and similarity measurement. However, each of these steps has its own method. Nevertheless, there is no universally acceptable method for retrieving similar images in CBIR.

Aim: Hence, this study examines the diverse methods used in CBIR systems. This is with the aim of revealing the strengths and weakness of each of these methods.

Methodology: Literatures that are related to the subject matter were sought in three scientific electronic databases namely CiteseerX, Science Direct and Google scholar. The Google search engine was used to search for documents and WebPages that are appropriate to the study

Results: The result of the study revealed that three main features are usually extracted during CBIR. These features include colour, shape and text. The study also revealed that diverse methods that can be used for extracting each of the features in CBIR. For instance, colour space, colour histogram, colour moments, geometric moment as well as colour correlogram can be used for extracting colour features. The commonly used methods for texture feature extraction include statistical, model-based, and transform-based methods while the edge method, Fourier transform and Zernike methods can be used for extracting shape features.

Conclusion: Each of the CBIR methods has their own advantages and disadvantages. However, there is a need for a further work that will validate the reliability and efficiency of each of the method.

Keywords: CBIR, features, feature extraction, Image

1. INTRODUCTION

CBIR is a term that was first introduced by Kato in 1992 [1]. Content-based image retrieval (CBIR) is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR). The major aim of a CBIR is to find images of interest from a large image database using the visual content of the images. CBIR is however entirely different from other classical information retrieval systems because they are highly unstructured. This is because digitized images consist purely of arrays of pixel intensities, with no inherent meaning [2]. They are however cheap, fast and efficient when compared with the text based image search method [3]. CBIR draws its methods from the field of image processing and computer vision. Generally, a Content based image retrieval (CBIR) is a term that is used to describe a retrieval technique which involves the use of visual information or contents called low level features to search and retrieve images from a large scale image database according to the requests of the user which is provided in the form of a query image. Nevertheless, image content may include semantic content [4]. The visual information is usually in form of colours, textures, shapes and spatial arrangement of region of interest. A CBIR retrieves relevant images by comparing the features of the images in the database with a given query image as well as finding the images that are similar to the queried image [5]. Thus, a CBIR can be viewed as an image search technique that is intended to search images that are almost similar in terms of colour, shape and text to a given query. Hence, the principal goal of a CBIR is to represent each image as a feature vector and to measure the similarity between the queried image and the images in database and also to retrieve similar images based on the features and not on textual annotations. The general architecture of a CBIR system is as shown in figure 1.

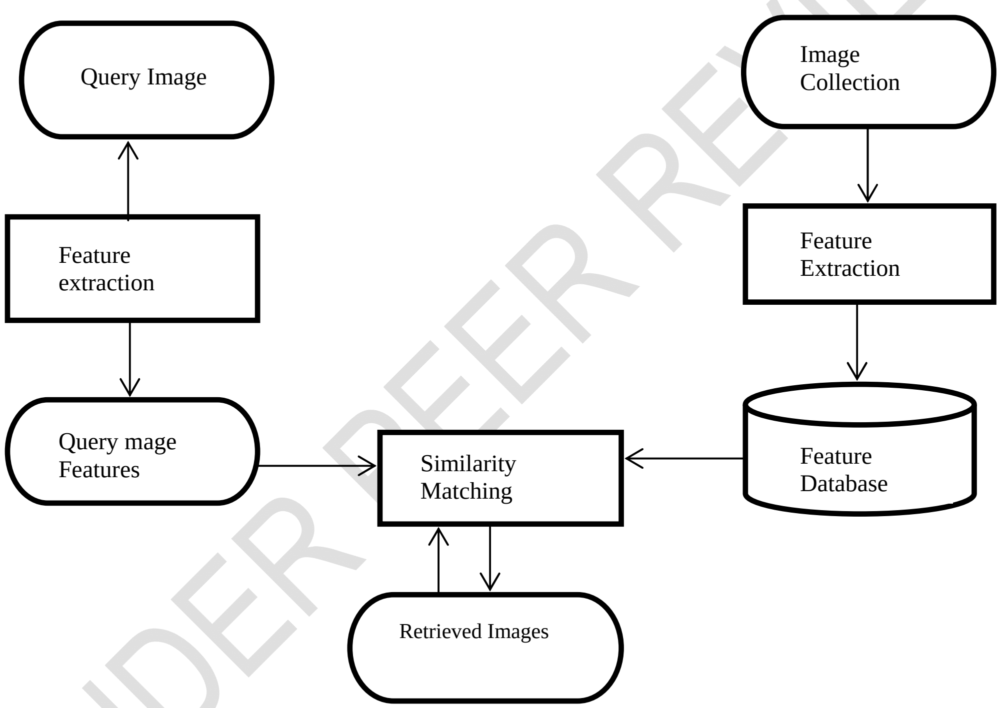

Figure 1: The General Architecture of a CBIR [3]

CBIR is performed usually in two steps. These include indexing and searching. During indexing the contents or features of both the queried image and the images in the image database are extracted and stored in the form of a feature vector in a feature database. This process is called the feature extraction. There are several methods that are used for extracting features in CBIR. Examples of features that can be extracted in CBIR include colour, texture and shape. Each of these characteristic features has diverse extraction methods. For colour extraction methods, colour space, colour histogram and colour moments are usually deployed.The commonly used methods for textural feature extraction are described by Manjunath and Ma [6] as statistical, model-based, and transform-based methods. One of the most widely used shape feature extraction method is the Edge method. In the searching step, a user query image feature vector is constructed and compared with all feature vectors in the database for similarity in order to retrieve the most similar images to the query image from the database [7, 8]. This process is referred to as similarity measurement. Again, there are diverse methods for computing similarity between a queried image and the images in the database. Typical examples of the methods used for similarity measurement include Sum of Absolute Difference (SAD), Sum of squared absolute Difference (SSAD), City Block Distance Canberra Distance and Euclidean Distance. Nevertheless, there is no universally acceptable method for extracting features and retrieving similar images in CBIR. Hence, this study examines the diverse methods used in CBIR systems. This is with the aim of revealing the strengths and weakness of each of these methods.

The paper is as organized as follows: section 2 is the methodology, section 3 reviews the features in CBIR, section 4 examines feature extraction methods in CBIR while section 5 examines the methods for computing similarity between a queried image and the images in a database. The paper concludes in section 6.

2. METHODOLOGY

Literatures that are related to the subject matter were sought in three scientific electronic databases namely CiteseerX, Science Direct and Google scholar. The Google search engine was used to search for documents and WebPages that are appropriate to the study

3. FEATURES IN CBIR

Features are observable patterns in the image that contain relevant information of an image. Pradeep et al. [9] also viewed a feature as a piece of information that is relevant for solving a computational task related to a certain application. Features describe and define the content of an image. They are described as the characteristics or the properties of the image. Features are usually used in image processing for searching, retrieval, and storage in order to achieve a high classification rate. The main goal of feature extraction is to obtain the most relevant information from the original data and represent that information in a lower dimensionality space .A good feature set contains discriminating information, which can distinguish one object from other objects [10]. One of the easiest ways of comparing images is through their features. This is because the direct method of comparing images by their pixels is not feasible and it is also time consuming for thousands of images stored in databases. Examples of features that can be extracted from images include colour, texture and shape. However, no particular feature is most suitable for retrieving all types of images.

3.1 Colour Features

Colour is the sensation caused by the light as it interacts with the human eyes and brain. Colours simplify objects identification. Color is one of the most widely used low-level visual features used in CBIR. However, different images can have the same colour distribution. Unfortunately, the retrieval of images with colour features only does not give accurate result because in many cases, images with similar colors do not have similar content.

3.2 Textural Features

There is no standard definition for texture. Haralick et al. [11] defines texture as a characteristic of an image that provides a higher-order description of the image and includes information about the spatial distribution of tonal variations or gray tones. Texture according to Hiremath and Pujari [12] is an innate property of virtually all surfaces, including clouds, trees, bricks, hair and fabric. Texture contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment [12]. Texture can also be defined as the pattern of information or arrangement of the structure of an image. An image can have more than one texture.

3.3 Shape Features

The shape of an object can be defined as the characteristic surface configuration of the object as defined by the outline or contour [3]. Shape features provide information for image retrieval, because humans can recognize objects solely from their shapes and shape carries semantic information about an object. Shape is highly significant in CBIR because it corresponds to region of interests in the images. There are two types of shape features. These include boundary-based and region-based features. The boundary based features extracts features based on the outer boundary of a region while the region-based features extracts features based on the entire region of an object [12].

4. FEATURE EXTRACTION IN CBIR

According to Nithya and Santhi [13], feature extraction in CBIR is a method of capturing the visual content of an image. Feature extraction can also be described as the process of extracting information that is semantically meaningful from images. The objective of feature extraction is to represent a raw image in a reduced form in order to facilitate decision making process. Hence, Kayode [14] views feature extraction as a special form of dimensionality reduction which takes place when the input data to an algorithm is too large to be processed and it is suspected to be redundant.

4.1 Colour Extraction Methods

There are diverse methods for extracting colour features in an image. These include colour space, colour histogram and colour moments.

4.1.1 Colour space

Colours are usually defined in three-dimensional color space so as to facilitate the specification of colours in an acceptable way. The RGB colour space is the most widely used color space. RGB stands for Red, Green, and Blue. RGB colour space combines the three colors in different ratio to create other colors. One of the major disadvantages of RGB colour space as emphasized by Mikhraq [3] that the RGB colour space is not uniform. The HSx color space is commonly used in digital image processing to convert the color space of an image from RGB color space to one of the HSx color spaces. HSx color comes in diverse forms. These include the HSI, HSV and HSB color spaces. The H and S in these colour space represents Hue and Saturation while the I, V, and B stand for Intensity, Value, and Brightness respectively. HSV color space is however the most commonly used colour space [15].

4.1.2 Color histograms

A color histogram according to Sivakamasundari and Seenivasagam [16] is a type of bar graph, where the height of each bar represents an amount of particular colour of the colour space being used in the image. The bars in a colour histogram are referred to as bins and they represent the x-axis and the number of bins relies on the number of colours in the image. The number of pixels in each bin is represented by the y-axis of the bar graph. There are two basic methods of obtaining a colour histogram. These include the global color histogram (GCH) and the local color histogram (LCH). GCH method takes the histogram of the image and computes the distance between two images by measuring the distance between their colour histograms. The drawback of the GCH as emphasized by Mikhraq [3] is that this method does not include information about all image regions. An LCH on the other hand divides an image into fixed blocks or regions, and takes the colour histogram of each of those blocks individually [3]. The similarity between two images is compared using LCH by computing the distance between the blocks of the images in the same location. The advantage of the LCH over the GCH is that the LCH is more efficient for image retrieval. However Mikhraq [3] stated that the LCH is computationally expensive and it does not work well when images are translated or rotated.

4.1.3 Colour moments

Color moments are measures that are used to compute the similarity of images based on their colour features. In color moments, the colour distribution of an image is seen as a probability distribution which is characterized by unique moments which include mean, standard deviation and skewness. The mean is defined as the average colour value in the image; the standard deviation is the square root of the variance of the distribution while the skewness is a measure of the degree of asymmetry in the distribution [3].

4.1.4 Geometric moment

This feature uses one value for the feature vector, thus, when the size of the image becomes relatively large, the computation of the feature vector will require a lot of time. The advantage of this method is that it produces a better result when combined with other feature extraction methods [17]. The drawback of geometric moments is that higher order moments are difficult to construct.

4.1.5 Colour correlogram

Colour correlogram is used for encoding the color information of an image [17]. A colour correlogram is a three-dimensional table indexed by colour and distance between pixels which expresses how the spatial correlation of colour changes with distance in a stored image. The colour correlogram may be used to distinguish an image from other images in a database. To create a colour correlagram, the colours in the image are quantized into m colour values $c_1 \ldots c_{m.}$ The advantage of this method is that it can be used to describe the global distribution of local spatial correlation of colours. It is also simple to compute.

4.1.6 Average RGB

The color average is described by Sharma and Sighn [18] in the RGB color space by X, as shown in equation 1.

$$X = (R(avg), G(avg), B(avh)\ )^t \qquad (1)$$

where R(avg), G(avg), and B(avg) are red, green and blue images average value

Sharma and Sighn [18] emphasized that this feature is used to filter out images with larger distance at first stage when multiple feature queries are involved.

4.1.7 Dominant Colour Descriptor (DCD)

This method is based on colour histogram. DCD chooses a small number of colors from the highest bins of a histogram. The number of bins chosen depends on the threshold value of the bin height [18].

4.1.8 Colour Coherence Vector

The colour coherence method is also based on the colour histogram. According to Sharma and Sighn [18], the colour coherence divides a histogram into two components namely coherent and non-coherent components. In coherent component, the pixels are spatially connected while in non- coherent component the pixels are isolated.

The advantages and disadvantages of the colour based extraction methods are summarized in Table 1.

Table 1. Pros and Cons of Different Colour Extraction Methods

Methods	Advantages	Disadvantages
Colour Space	RGB is easy to implement. HSV represents a wealth of similar colours; they are also device dependent	RGB colour space is not uniform.
Colour Histogram	comparison of histogram features is faster and more efficient than other methods	It can be potentially identical for two images with different colour content
Colour Moments	There is no need to store the complete colour distribution; hence image retrieval is faster with this technique. It also involves the comparison of less features	They cannot handle occlusion successfully
Geometric Moments	produces a better result when combined with other feature extraction methods	higher order moments are difficult to construct, thus they are difficult to compute
Colour Correlogram	It is simple to compute and may be used to distinguish an image from images in a database. It takes into cognizance the local colour spatial correlation. Effective for CBIR from a large image database.	they have high computational complexity and low retrieval accuracy
Average RGB	Its computation cost is not high	it is less accurate if not combined with other feature extraction methods
DCD	It is scalable.accurate if compact	it does not give spatial information of the image
Colour Coherence Vector	it gives spatial information about the image	it has high computational cost

4.1 Texture Extraction Techniques

The commonly used methods for texture feature extraction are described by Manjunath and Ma [12] as statistical, model-based, and transform-based methods.

4.2.1 Statistical Approaches

In statistical methods, the spatial distribution of grey values is computed by finding the local features at each point in the image, and deriving a set of statistics from the distribution of the local features. Typical examples of statistical approaches include gray level co-occurrence matrix (GLCM) and Tamura features.

4.2.1.1 Gray Level Co-Occurrence Matrix (GLCM)

Gray Level Co-Occurrence Matrix (GLCM) is usually computed to discriminate different textures. The GLCM is a function of an angular relationship between two pixels with corresponding gray level, *i* and *j* and a function of the distance between them which characterizes the spatial distribution of gray levels between them. An element in the GLCM, $P_{d,\Theta}$ (*i,j*), represents the frequency of occurrence of the pair of gray levels (*i,j*), separated by a distance *d* at a direction *Θ*. When each entry in the matrix is divided by the total number of neighbouring pixels *R*, a normalized GLCM is obtained, and the sum of its elements is equal to 1.

The notations in equations (2)-(5) are used to describe the various textual features in GLCM.

$$\mathrm{p(i,j)} = \text{the } \mathrm{(i,j)^{th}} \text{ entry in a normalized GLCM given by } {}^{\mathrm{P_{d,\theta}(i,j)}}/_{\mathrm{R}} \tag{2}$$

$$\mathrm{N_g} = \text{the number of distinct gray levels in quantized image} \tag{3}$$

$$\mathrm{px(i)} = \sum_{\mathrm{j=1}}^{\mathrm{N_g}} \mathrm{p(i,j)}\text{, the } \mathrm{i^{th}} \text{ entry in the marginal probability matrix} \tag{4}$$

obtained by summing the rows of $\mathrm{p(i,j)}$and $\mathrm{px(i)}$is the $\mathrm{i^{th}}$ entry of row i

$$\mathrm{py(j)} = \sum_{\mathrm{j=1}}^{\mathrm{N_g}} \mathrm{p(i,j)} \text{ the } \mathrm{j^{th}} \text{ entry in the marginal probability matrix} \tag{5}$$

obtained by summing the rows of $\mathrm{p(i,j)}$and $\mathrm{px(j)}$is the $\mathrm{j^{th}}$ entry of column j

Hence, the general equations for the Haralick features are given in equations (6) and (7) respectively.

$$px+y(k) = \sum_{i=1}^{N_g}(i,j)\sum_{\substack{j=1\\ i+j=k}}^{N_g} p(i,j)\ ,k=2,3,\dots\dots\dots,2N_g \tag{6}$$

$$px+y(k) = \sum_{i=1}^{N_g}(i,j)\sum_{\substack{j=1\\ |i-j|=k}}^{N_g} p(i,j)\ ,k=0,1,\dots\dots\dots,N_g-1 \tag{7}$$

Based on the notations above, Haralicks et al. [11] proposed 13 common statistical features known as the Haralicks textual features. The 13 Haralick features are given in equations (8) to (25)

i. Energy: This is also known as the angular second moment. It measures the textual uniformity of an image. Energy is as given in equation (2.20).

$$energy = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(p(i,j))^2 \tag{8}$$

ii. Contrast: This is a measure of intensity or gray-level variations between the reference pixel and its neighbor. Contrast is as given in equation (9).

$$contrast = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \mid i-j \mid^2 p(i-j) \tag{9}$$

iii. Correlation: This presents how a reference pixel is related to its neighbour. Correlation is expressed in equation (10) as follows:

$$correlation = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i,j)p(i,j)-\mu_x\mu_y}{\sigma x\sigma y} \tag{10}$$

Where$\mu_x, \mu_y, \sigma x \; and \; \sigma y$ are the means and standard deviations of *p(x)* and *p(y)* respectively.

iv. Homogeneity: This is also known as Inverse Difference Moment. It measures image homogeneity. Heterogeneity is as given in equation (11).

$$Homogeneity = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{1}{1+(i-j)^2}p(i,j) \tag{11}$$

v. Entropy: This measures the disorder or complexity of an image. Entropy is as given in equation (12).

$$Entropy = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}p(i,j)\log\left(p(i,j)\right) \tag{12}$$

vi. Variance: This is also referred to as the sum of squares. It measures the dispersion of the difference between the reference and the neighbour pixel in a window. Variance is as given in equation (13).

$$Variance = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu)^2 p(i,j) \tag{13}$$

Where μ is the mean gray level of the image

vii. Sum Average: The sum average is given as shown in equation (14).

$$Sum\ Average\ (SA) = \sum_{k=2}^{2N_g}k.p_{x+y}(k) \tag{14}$$

viii. Sum Entropy: Sum Entropy is given as shown in equation (15).

$$Sum\ Entropy, SE = -\sum_{k=2}^{2N_g}p_{x+y}(k)log\left(p_{x+y}(k)\right) \tag{15}$$

ix. Sum Variance: This is given in equation (16).

$$Sum\ Variance\ (SV) = \sum_{k=2}^{2N_g}(k-SE)^2 p_{x+y}(k) \tag{16}$$

x. Difference Variance: This is expressed in equation (17).

$$Difference\ Varianve\ (DV) = \sum_{k=0}^{N_g-1}k^2\, p_{x-y}(k) \tag{17}$$

xi. Difference Entropy: The difference entropy is as given below in equation (18)

$$Difference\ Entropy\ (DE) = -\sum_{k=0}^{N_g-1}p_{x-y}\,(k)\log\left(p_{x-y}(k)\right) \tag{18}$$

xii. Information Entropy of Correlation (IEC): This is expressed in equation (19).

$$IEC = \frac{HXY-HXY1}{\max\,(HX,HY)} \tag{19}$$

Where

$$HXY = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log(p(i,j)) \quad (20)$$

$$HXY1 = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\log(px(i)py(j)) \quad (21)$$

HX and HXY1 are the entropies of *px* and *py* respectively

xiii. Information Measure of Correlation (IMC): This is as expressed in equations (22) and (23) respectively.

$$IMC = (1 - \exp(-2(HXY2 - HXY)))^{1/2} \quad (22)$$

Where

$$HXY2 = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} px(i)py(j)\log(px(i)py(j)) \quad (23)$$

4.2.1. 2 Tamura Features Extraction Method

Tamura *et al.* [19] also proposed texture representations that were based on psychological studies of human perception, and these representations consists of six statistical features, including coarseness, contrast, directionality, regularity, line-likeness, roughness to describe various texture properties.

i. Contrast: Contrast measures the distribution of gray levels that varies in an image and to what extent its distribution is biased to black or white. The second order and normalized fourth–order central moments of the gray levels are used to define the contrast. Contrast is given as in equations (24) and (25) respectively.

$$\text{Contrast} = \sigma/(\alpha 4) \quad (24)$$

$$\alpha 4 = \mu 4/\sigma 4 \quad (25)$$

$\mu 4$ is the fourth moment about the mean and σ is the variance.

ii. Directionality: Directionality of an image is measured by the frequency distribution of oriented local edges against their directional angles. This texture feature given by Tamura does not differentiate between orientations or patterns but measures the total degree of directionality in an image. Directionality is expressed as depicted in equation (26).

$$\text{Directionality} = 1 - \text{rnpeaks}\sum_{p=1}^{npeaks}\sum a \in w_p\,(\text{a} - \text{ap})^2\text{Hdirectionality(a)} \quad (26)$$

where npeaks is the number of peaks, a_p, is the position of the peak, w_p is the range of the angles attributed to the P^{th} peak, r denotes a normalizing factor related to quantizing levels of the angles a, and a denotes quantized directional angle, $HDirectionality$, is the histogram of quantized direction values, a is constructed by counting number of the edge pixels with the corresponding directional angels.

iii. Line-Likeness: Line-Likeness in an image is the average coincidence of direction of edges that co-occurred in the pairs of pixels separated by a distance along the edge direction in every pixel.

iv. Regularity: Regularity measures a regular or similar pattern that occurred in an image. Regularity is defined in equation (27) as follows:

$$\text{Regularity} = 1 - \text{r}(\text{Scrs} + \text{Scon} + \text{Sdir} + \text{Slin}) \quad (27)$$

Where *Scrs*, *Scon, Sdir* and *Slin* are similar coarseness, contrast, directionality and line-likeness in an image respectively.

v. Roughness: Roughness is the summation of contrast and coarseness measures. Roughness is as shown in equation (28).

$$\text{Roughness} = \text{Contrast} + \text{Coarseness} \tag{28}$$

vi. Coarseness: Coarseness basically relates to the distance in gray levels of spatial variations, which is implicitly related to the size of primitive elements forming the texture. It has the direct relationship to scale and repetition rates and most fundamental texture features. Coarseness is expressed as shown in equation (29).

$$\mathrm{A}_k(\mathrm{x},\mathrm{y}) = \sum_{i=x-2k-1}^{x+2k-1}\sum_{j=y-2k-1}^{y+2k-1}\frac{\mathrm{f(i,j)}}{2^{2k}} \tag{29}$$

Where 2^{2k} size is the average of neighborhood.

In most cases, coarseness, contrast and directionality are commonly used for CBIR systems because they capture high-level perceptual attributes of a texture and are also useful for browsing of images [20].

4.2.2 Model Based Approaches

Model-based texture methods is used to compute the process that generated the texture. The model based approach is generated by computing a random field as stated by Mikhraq [3] as follows:

Assuming an image is modeled as a function $f\ (r,)$, where r is the position vector representing the pixel location in the 2-D space and ω is a random parameter. For a given value of r, $f\ (r,)$ is a random variable because ω is a random variable. Once a specific texture ω is selected, $f\ (r,)$ is an image, which is a function over the two-dimensional grid indexed by r. Function $f\ (r,)$ is called a a random field. A typical example of the model based approach is the Markov random fields.

4.2.3 Transform-Based Methods

Typical examples of transform based methods include Fourier transform and wavelet analysis.

4.2.3.1 Fourier transform

In Fourier transform, the image signal is broken into sine waves of various frequencies. A variant of the Fourier transform is the Fast Fourier Transform (FFT). The FFT according to Shukla and Vania [17] refers to a class of algorithms for efficiently computing the Discrete Fourier Transform (DFT). Hence, Shukla and Vania [17] emphasized that FFT is not an approximation of the DFT, but rather it is the DFT with a reduced number of computations. One of the disadvantages of the FT is that it does not capture the objects locations in an image [17].

4.2.3.2 Discrete wavelet transform

Discrete Wavelet Transform involves the decomposition of an image into basic functions obtained through translation and dilation of a special function. The Discrete Wavelet Transform is very effective in image analysis and compression [17].

4.2.3.3 Ranklet transform

The Ranklet Transform belongs to a family of non-parametric, orientation-selective, and multi-resolution features. This method has three main properties. First, it is nonparametric because it deals with the relative order of pixels instead of their intensity values. Second, it is orientation selective because it is modeled on Haar wavelets. Lastly, it is multi-resolution. This implies that the Ranklet Transform can be calculated at different resolutions using Haar wavelet supports. The Ranklet Transform performs better than the pixel-based and wavelet-based image representations.

4.2.3.4 Steerable pyramid

This technique generates a multi-scale, multidirectional representation of the image [17]. It involves the decomposition of the image into low-pass sub-band and high-pass subband. However, the decomposition is iterated in the low-pass subband [20].

The advantages and disadvantages of the textural based extraction methods are summarized in Table 2.

Table 2. Pros and Cons of Different Textural Extraction Methods

Methods	Class	Advantages	Disadvantages
GLCM	Statistical	GLCM produces better results than other texture discrimination methods. It also enhances the details of an image and gives the interpretation. It reduces image compression time. It is a good discriminator when studying images. Hence, it is a widely used textural extraction method	Characterized by high development cost and time
Tamura	Statistical	It motivates human visual perception	works only on homogenous texture images and performs poor on generic images
Fourier Transform	Transform based	It improves signal to noise ratio	it does not capture the objects locations in an image
Discrete Wavelet	Transform based	Discrete Wavelet Transform is very effective in image analysis and compression	It has poor directionality.
Ranklet Transform	Transform based	It performs better than the pixel-based and wavelet-based image representation. They are robust in detecting outliers	It has high computational cost
Steerable Pyramid	Transform based	it allows the independent representation of scale and orientation of image structure	space-domain implementation is not perfect

4.3 Shape Extraction Techniques

Examples of shape extraction methods include the edge method, Fourier descriptor and Zernike method. These methods are briefly described below.

4.3.1 Edge method

One of the most widely used shape feature extraction method is the Edge method. Edge is used to capture the information about the shape of an object. A typical variance of edge is the edge histogram. It is used to represent the relative frequency of occurrence of five types of edges in each local area called a sub image or an image block. The sub-image is obtained by dividing the image space into a 4x4 non-overlapping blocks. Thus, the image partition always yields 16 equal-sized sub-images regardless of the size of the original image [3].

4.3.2 Fourier descriptors

This technique involves the application of Fourier transform on the shape boundary of an image. The Fourier transformed coefficients are usually referred to as the Fourier descriptors (FD) of the shape. They are robust and easy to derive [17]. Fourier descriptors are not affected by noise [17].

4.3.3 Zernike moments

This method allows independent moment invariants to be constructed to an arbitrarily high order [17]. It is suitable for complex shape representation and does not need to know boundary information of the image [17].

The advantages and disadvantages of the shape extraction methods are summarized in Table 3.

Table 3. Pros and Cons of Different Shape Extraction Methods

Methods	Advantages	Disadvantages
Edge Method	captures the information about the shape of an object. It also locates sharp discontinuities in an image	The discontinuities abrupt changes in pixel intensity scene.
Fourier descriptors	They are robust and easy to derive	for Fourier descriptors to be accurate, their values are usually calculated and stored in float numbers
Zarnike Method	suitable for complex shape representation and does not need to know boundary information of the image	Computational complexity is high

5. Similarity Measures in CBIR

The similarity between two images, represented by their features values, is defined by a similarity measure [3]. In similarity measurement, the query image is compared with the images in the database. Similarity measure is usually computed by finding the similarity between the query image and the database images. This is usually done by computing the difference between the query feature vector and the database feature vectors. Typical examples of the distance metrics used in CBIR include the following:

<u>5.1 Sum of absolute difference (SAD)</u>

The sum of absolute difference (SAD) is extensively used for computing the distance between the images in CBIR to get the similarity. In this metric, the sum of the differences of the absolute values of the two feature vectors, Q_i and D_i is calculated. This distance metric according to Selvarajah and Kodituwakku [21] can be calculated as shown in equation (30).

$$\Delta d = \sum_{i=1}^{n}(|Q_i| - |D_i|) \qquad (30)$$

where n is the number of features, $i = 1, 2. . ., n$. Both images are the same for d= 0 and the small value of Δd shows the relevant image to the query image.
SAD is simple when the query image and the image in the database are similar [22]

5.3 Sum of squared absolute Difference (SSAD)

In this metric, the sum of the squared differences of absolute values of the two feature vectors is calculated. This distance metric according to Selvarajah and Kodiyuwakku [21] can be calculated as shown in equation (31).

$$\Delta d = \sum_{i=1}^{n}(|Q_i| - |D_i|)^2 \qquad (31)$$

SSAD is more computationally complex than SAD.

5.4 Euclidean distance

This distance metric is the most commonly used for similarity measurement in image retrieval because of its efficiency and effectiveness [9]. It measures the distance between two vectors of images by calculating the square root of the sum of the squared absolute differences and it can be calculated as shown in equation (32).

$$\Delta d = \sqrt{\sum_{i=1}^{n}(|Q_i| - |D_i|)^2} \qquad (32)$$

5.5 City block distance

This distance metric is also called the Manhattan distance. The city block distance metric has robustness to outliers. This distance metric according to Szabolcs [23] is computed by the sum of absolute differences between two feature vectors of images and can be calculated as shown in equation (33).

$$\Delta d = \sum_{i=1}^{n}(|Q_i| - |D_i|) \qquad (33)$$

The city block distance metric gives a large value for the two similar images which create dissimilarity between similar images.

5.6 Canberra distance

This metric is used for numerical measurement of the distance between the query and database feature vectors. The value of this method is arranged in ascending order such that the top most shows high similarity. It has similarity with city block distance metric [24].Canberra distance is computed as shown in equation (34).

$$\Delta d = \sum_{i=1}^{n}\frac{|Q_i - D_i|}{|Q_i + D_i|} \qquad (34)$$

The advantages and disadvantages of the diverse similarity methods are summarized in Table 4.

Table 4. Pros and cons of different methods for similarity measurements in CBIR

Methods	Advantages	Disadvantages
Sum of Absolute Difference (SAD)	It is simple when the query image and the image in the database are similar	it is sensitive to background issues of images such as variations in size, color, illumination and direction of light
Sum of squared absolute Difference (SSAD)	It be used in both pixels and transformed domains but in the transformed domain	SSAD is more computationally complex than SAD
Euclidean Distance	it is the most efficient and effective similarity measure	It assumes that the sample points are distributed about the sample mean in a spherical manner
City Block Distance	has robustness to outliers	gives a large value for two similar images which create dissimilarity between similar images
Canberra Distance	It is good for data that are spread about the origin	It can only be used for positive values.

6. CONCLUSION

This study examines different techniques used in CBIR systems. The study reviewed diverse literatures that are related to CBIR. The study found out that there are three basic features that can be extracted in CBIR. These include colour, texture and shape. The study also revealed that each of these features has different extraction methods. For instance, colour can be extracted in images using colour histogram, geometric moments, colour sp[ace and colour moments. The study revealed the strengths and weaknesses of each of these techniques. For instance, the colour space method is easy to implement but it is not uniform while the colour histogram is faster and more efficient than other colour extraction methods. It can however be identical for two images with different colours. The study also reveals that the GLCM, Tamura, Fourier transform, Ranklet transform and discrete wavelets are typical examples of textural extraction methods. Similarly, the edge method, Fourier descriptors and Zernike method were the shape extraction methods revealed in this study. Furthermore, the study investigated the techniques for computing the similarity between a query image and the images in the database. The result of the study showed that examples of similarity measures used in CBIR include sum of absolute difference, sum of the squared differences of absolute values and city block distance are

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Lew M, Sebe N, Djeraba C, Jain R. Content-based multimedia information retrieval: state of the art and challenges.ACM Transactions on Multimedia Computing, Communications and Applications. 2006; 2(1): 1-19.
2. Sukhdeep K, Deepak A . Image content based retrieval system using cosine similarity for skin disease images.technique-survey. International Journal of Computer Science and Information Technologies. 2013; 5 (4):89-95.
3. Mikhraq AK Content-based image retrieval (cbir) system based on the clustering and genetic algorithm. A Master's Thesis in Bioengineering Computer Engineering Department, Faculty of Engineering Deanery of Higher Studies, Islamic University Gaza Palestine. 2013.
4. Singh N, Dubey S, Dixit P, Gupta J. Semantic image retrieval by combining colour, texture and shape features. International Conference on Computing Sciences. 2012.
5. Hiremath P, Pujari J. Content based image retrieval using colour, texture and shape features.15th International Conference on Advanced Computing and Communications. 2008.
6. Manjunath B, Ma W. Texture features and learning similarity. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'96), San Francisco, California. 1998; 425–430.
7. Nezamabadi-pour, Hossein, Saryazdi, Saeid. Object-based image indexing and retrieval in DCT domain using clustering techniques. World Acad. Sci. Eng. Technol. 2005; 3–6.
8. Veltkamp RC, Tanase M, Sent D. Content-based image retrieval systems: A survey. Utrecht University, Department of Computing Science, Utrecht, The Netherlands, 2002; 97-124
9. Pradeep N, Girisha H, Sreepathi B, Karibasappa K. Feature extraction of mammograms. International Journal of Bioinformatics Research. 2012; 4(1): 241-244.
10. Venkat GN, Vijay RV. Content-Based Image Retrieval Systems. IEEE Computer Society, Insistute of Electrical and Electronics Engineering. 1995;28(9):18-22.

11. Haralick RM, Shanmugan KS, Dunstein I. Textural features for image classification. IEEE Trans. Syst., Man, Cybern. 1973; 3(6):610–621.
12. Hiremath P, Pujari J. Content based image retrieval using colour, texture and shape features.15th International Conference on Advanced Computing and Communications. 2008.
13. Nithya R, Santhi B. Comparative study on feature extraction method for breast cancer classification. Journal of Theoretical and Applied Information Technology. 2011; 33(2): 220-226.
14. Kayode AA. Development of a computational model for classification of mammographic images. Ph.D Thesis, Computer Science Department, ObafemiAwolowo University Ile-Ife, Osun State, Nigeria.2014
15. Shih TK., Huang J, Wang C, Hung JC, Kao C. An intelligent content-based image retrieval system based on colour, shape and spatial relations. Proc. Natl.Sci.Counc. ROC(A). 2001; .25(4): 232-243.
16. Sivakamasundari G, Seenivasagam V. Different relevance feedback techniques in CBIR: a survey and comparative study. International Conference on Computing, Electronics and Electrical Technologies. 2012.
17. Shukla J, Vania J. A survey on CBIR features extraction techniques. International Journal Of Engineering And Computer Science. 2014; 3(12): 9555-9559.
18. Sharma, R. Singh, N. Comparative study of different low level feature extraction techniques. International Journal of Engineering Research and Technology. 2014; 3(4):1454-1460.

19. Tamura H, Mori S, Yamawaki, T. Texture features corresponding to visual perception. IEEE Trans. on Systems, Man, and Cybernetics. 1978: 8(6);460-473
20. Bagri N, Johari P. A comparative study on feature extraction using texture and shape for content based image. Proceedings of the SPIE Medical Imaging: PACS and Imaging Informatics. 2016: 5748: 134 -143.
21. Selvarajah S, Kodituwakku SR. Analysis and comparison of texture features for content based image retrieval. International Journal of Latest Trends in Computing. 2011;2:108-113.
22. Malik F, Baharudin, B. Analysis of distance metrics in content-based image retrieval using statistical quantized features related to breast density patterns. Journal of Digital Image Processing. 2013; 20(2): 172–190.
23. Szabolcs S. Colour Histogram Features Based Image Classification in Content Based Image Retrieval System. A comprehensive review on recent relevance feedback techniques in content based image retrieval. International Journal of Computer Science and Information Technologies. 2008; 3(4); 4818-4822.
24. Schulz, J Algorithms-Similarity. 2007. Assessed 15 September 2017. Available: http://www.code10.info/index.php?option=com_content&view=article&id=49:article_canberra-distance&catid=38:cat_coding_algorithms_data-similarity&Itemid=57