FEATURE EXTRACTION TECHNIQUES AND DESCRIPTORS USED IN IMAGE RETRIEVAL SYSTEM- A SURVEY

[1]UJWAL R.HARODE,[2]Dr.S.A.PATIL

[1]Assistant Professor
[1]Electronics Engineering, Department
[1]Pillai College of Engineering, New Panvel , India.

Abstract: In the field of Image Processing Content extraction is one of the most important step. Identification means the preparation and delivery of content that matches with the available resources in an optimal way. The content extraction method can be classified as perceptual because it detects the prominent regions in images. In many applications image retrieval is mostly used to find or to get the image for particular purpose. To retrieve the more similar content from the large image dataset numbers of techniques are available.
The main motivation for extracting the content of image is the approachability problem. A problem that is even more relevant for dynamic multimedia data, which also have to be access. While content extraction for retrieval techniques is developed for text, there is a long way to go in extracting features and the implementation of more suitable and effective processing model not only for images but for videos also. Description of the content of multimedia information is required for classification purpose. In this paper we try to explore the growth of content extraction techniques and different descriptors used in image retrieval field.

***Index Terms - Image Processing, Content Extraction, Feature Extraction, Image Descriptors, Image Retrieval*, similarity measures, performance measure metrics.**

I. INTRODUCTION

With the rapid advances in real-time image processing and large capacity data processing, there is an increasing demand for safe, effective, substantial and economic image classification and retrieval process system. In these systems, an image is typically represented by numerous low-level features and high level features. These features are nothing but the information which is useful for computing a retrieval task. The main features to be detected in an image are ridges, edges, blobs and corners. Many detectors for the same like Canny, Sobel, Prewitt, Harris, LoG, DoG etc are used. Feature detection, description and matching are essential elements of computer vision applications; therefore the meaningful content extraction is utmost important.

The image Representation with multiple features from multiple perspective approaches is a better idea for achieving better result. Here the main issue is to organize the information appropriately so that the expected outcomes can be realized.

1.1General Purpose Systems-

So many general-purpose systems are available for CBIR. The main focus of such system is on the feature information extraction which used to match the query image with the stored images from database. Some of such systems are discussed below:

1.1.1 QBIC

IBM Almaden Research Center, San Jose (1995) had developed Query by Image Content (QBIC) system. In this system sketches, texture, shape, color are used for the image representation. In this system user defined color, texture patterns, sketches, drawings, camera, object, motion based queries are used on large images and video databases.

1.1.2 NETRA

This system had been developed by Electrical and Computer Engineering department of California University, Santa Barbara (1997). Three feature vectors are use for image representation in this system. Color histogram is used as the first vector for image color feature representation. From the Gabor Wavelet Transform, standard deviation and normalized mean are derived which represents the image texture feature. These image texture features are used as a second vector. The shape feature are derived from the curvature function of the contour and used as a third vector. From these three vectors similarity and dissimilarity of data objects is measured by using the Euclidean distance tool.

1.1.3KIWI

France based INSA Lyon had been developed Key-point Indexing Web Interface (KIWI), in 2001.Wavelet-based salient point detector is used for the extraction of the key points in the query image. From R, G, and B color component and Gabor Filter, Color histograms are computed and use for shape description.

1.1.4 Photo book

This system had been developed in 1997 by Vision and Modeling Group at MIT Media Lab. In this system images are represented using three different methods based on image content type i.e. face, shape and texture. It uses three sub-books- Appearance Photo book, Shape Photo book and Texture Photo book. Users can select any one from face, shape and texture features to query an image or a combination of these features along with a text.

All these systems discussed above are formulating the special semantic features for query based image retrieval. But due to the variety of visual features in image retrieval applications, these semantic features descriptions have not been perfect. For example, all of these systems can't be useful for some special images (like medical images for example) because these systems uses some simple feature extraction methods which may provide unwanted results and they are not that much precise.

2 Various CBIR Techniques-Literature surveys

Early in 1980's almost all learning methods were based on the linear decision surfaces. These methods have nice theoretical properties. Moreover the keyword based image retrieval was the conventional method to retrieve the images. This method leads to semantic gap which is nothing but the difference formed when the same object is represented by two different notations. This needs the tool for translating the natural language into a low level computer language. Moreover the examination of this translational capacity was examined by Chomsky hierarchy and found that beyond a particular level of expressional power the translation is not possible. This lead to the development of a system for image retrieval where retrieval is based on the various contents in the images like color, shape, texture etc. The first microcomputer based image database image retrieval is developed by Massachusetts Institute of Technology in 1990's. This paper got the inspiration from the Romaric Pighetti paper on Hybrid Model of CBIR combining SVM with Multi Objective Genetic Algorithm [1].

The Relevance feedback mechanism as stated in M.Arevalillo-Herrez Distance Based relevance feedback using hybrid interactive genetic algorithm [2] gives the inspiration to the paper on RF loop. The nearest neighbours are gathered using nn algorithms and further precise classification is done using the SVM. The GA is adopted in the single objective sense wherein the optimum is achieved for that particular objective.

Lots of search techniques are available with the wide range of applications, but it is difficult to understand the user interest. To make these techniques successful, relevance feedback _can be use,_ where the user can add feedback to the search results by marking images in the results as "relevant", "not relevant", or "neutral" and then repeating the search with the new parameters[17] [18].

E.Chang, S.Tong, K.Goh and C.Chang, Support Vector Machine concept – dependent active learning for image retrieval, introduced a method where annotations sets should lead us to higher classification rate [3]. The basis for a learning algorithm is that it must learn from less number of annotations. In order it to be effective it has to learn the user's query concept quickly and accurately simultaneously asking the user to annotate less number of images. The complexity of the concept is characterized by the three measures namely hit rate, isolation and diversity. These annotations are obtained from the user through the relevance feedback loop. This method made an improvement from the previous methods but the solution space was not wide. The main drawback of this method is local minima. In the empirical study of the datasets it is clear concept independent - relevance feedback and traditional passive learning is completely outperformed by the active learning relevance feedback algorithm.

The SVM is a machine learning concept associated with other learning algorithms to study the images. The SVM is a supervised learning model that categorises the classes. It is a non-probabilistic binary linear classifier.The SVM model is formed using the radial basis kernel function. Here there are many features considered for the high dimensional feature space. The hyper plane which is nothing but the high dimensionality feature space gives a good gap between the two classes and hence reduces the generalization error of the classifier.

M.Cruciano, D.Estevez paper on hyper plane queries [4] gives vast information on SVM's hyper planes.M.Crucianu, D.Estevez, V.Oria and J.P.Tarel designed a system using the feature space m tree for speeding up the active learning. The main factor in this method which bridges the semantic gap is by the interaction with the user. The decision frontier might be quite complex but the initial space is mapped into the high dimensional feature space by the specific kernel functions. In this method the main concept consists in performing kNN hyper plane queries with the M tree built up in the feature space. Hyper plane queries are formed using the SVM classification based on kernel function. The relevance feedback is a noticeable method to reduce the semantic gap. Though the retrieval time is ok, the issue is with the scalability. In this method the results shows that considerable increase in the retrieval time is achieved by increasing fewer rounds in the feedback loop. The results also point out that approximate search can be optimize in challenging conditions.

C. D. Ferreira, J. A. Santos, R. da S. Torres, M. A. Gonc¸alves, R. C. Rezende, and W. Fan.[5] had proposed a method for genetic programming for image retrieval based on RF loop using heuristic approach and genetic algorithm. Color hologram filter, wavelet texture filter and spatial filter are used in sequence to find out the similarity or dissimilarity of images to a query image.

The Genetic Algorithm is particularly to have a kind of natural selection of the images. To optimise the search the GA is used

The Genetic Algorithm is multi point search methods. GA is not only very easy to use but also an effective optimisation tool. Each string of the search space forms the candidate solution to the problem called chromosome. The objective function value of the chromosome is called the fitness value.

P.Gosselin and M.Cord. Active learning methods for interactive image retrieval [6] lead to arriving at fewer annotations sets for classification. Still the same difficulties of small solution space are limiting the usage of this system. The decision trees and NNs allowed efficient learning of nonlinear decision surfaces but this suffered from the local minima.

The Semi-supervised SVM batch active mode active learning for image retrieval based method is compared with neural networks, logistic regression and has shown good results comparatively. Kernelized locality-sensitive hashing for scalable image search as suggested in Kulis-Grauman[7] paper gives as system based on the kernel functions used for nonlinear svm.

GA has been widely used in conjunction with other Artificial Intelligent techniques. Only a fewer has been made with combining SVM with GA though it has very useful application in this area. The work of Lai et al. present an IGA [8],[9] to address CBIR context from whom this framework is inspired, but the preference is to combine GA with a SVM to retrieve the images of interest more precisely. The results show that this combination outperforms that of pure SVM and NN algorithms. This method simultaneously optimises the parameter and feature selection without degrading the classification accuracy of SVM. This significantly improves the classification accuracy and has fewer input features for SVM.

In Romeric Pighetti, Dennis Pallez, Frederic Precioso paper "The hybrid content based image retrieval combining Multi Objective Interactive Genetic Algorithm and Support Vector Machine"[10] the main focus is on attaining relevant images from very large databases in less number of iterations in RF loop and converging to global optima. The global optima are attained by the application of the GA. In previous paper single objective GA is applied and the optimum of that objective is achieved. In this paper more than one objective is considered and the Pareto optimal solution of those objectives is considered.

In [11],[12]and[13] the SVM is studied in detail for statistical learning. Moreover the feature subset selection considered is discussed in the above papers. Finally the Support Vector Machine is to mainly maximise the margin so that the separability called the Winston terminology the street separating hyper plane between the two classes is well defined. The SVM is trained with initially the query image. The SVM gives fitness values for the images retrieved using low level descriptors based on the minimum distance from the decision boundary. With these fitness values the GA acts the selection, crossover and mutation eventually. The project has been implemented till this stage.

Usama Sayed, Mofaddel, Elhafiez and Abdel Gawad[14] have proposed method of object extraction of gray scale image using curve let transform. The morphological gradients are used to extract object from image. Wavelet transform are efficiently used in image processing but randomly generated edges cannot be handle by wavelets.

It is state by [15], that the saliency of content i.e. object, a person, or a pixel, is the quality by which it stands out relative to its neighbors. The contrasts between image contents of interest and their neighborhood typically considered as Saliency.

The visual saliency models can be use as a new evaluation method in CBIR[16].It is acts as a filter for the key points used by the recognition system. The visual attention models are based on Mean Average Precision (MAP).

3 CBIR Systems

In this section we display the different features and the strategies used for their extraction in CBIR System.In order to provide more efficient content extraction for any specific application, it is necessary to extract meaningful content from scenes.

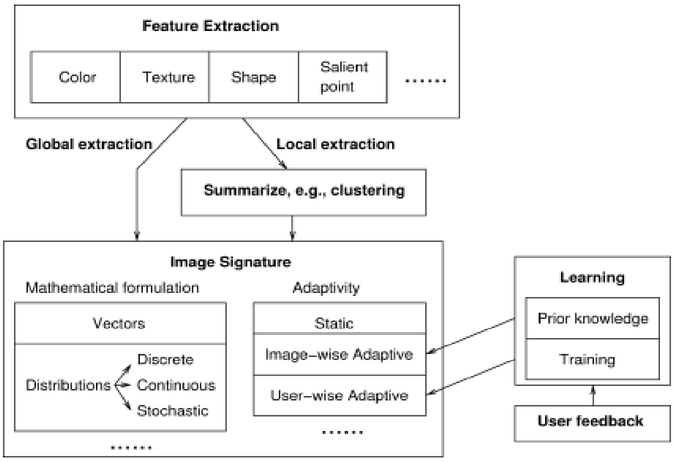

Fig1 CBIR Block diagram

3.1 Query Methods and Processing

In the field of image retrieval, an important parameter is user-system interaction complexity. From a user perspective, different methods can be used as a query to a system. The characteristics of different querying methods and the system support requirement are described below.

Keywords- The most popular and simple way for image searching is word or bigram as use in the Google and Yahoo! image search engines.

Free-Text- Here a sentence, question, complex phrase or story can be used as a query for image search.

Image-The most efficient way of querying is to use image of interest as query and search for similar images in stored databases.

Graphic- This presents a query by a hand-drawn or computer-generated picture.

Composite-Two or more methods can be use for querying a system such as in relevance feedback system.

All query classes make use of Text, Sketch, Shape, Volume, Color, Texture, Spatial constraints, Browsing, Objective and Subjective attributes, Motion and Domain concepts for retrieval of images.

3.2 Feature Extractions in CBIR

In image retrieval the content of the image are described by the low-level and high-level features such as shape, texture, color etc. The visual information for each pixel is extracts from the image and generally saves those as feature value (a set of feature value called a feature vector) in a database. These extracted features of the image are used to compare the query image with the other images from database for retrieval purpose.

The extracted features are represented in terms of vectors like –Vi and Vq Where

$V_i = [w'_1 w'_2 .... w'_d]$ Is a vector of i image of the database and $V_q = [q_1 q_2 ..... q_d]$ Is a vector of the query image q.

From these feature vectors some kind of similarity matching can be computed which give the similarity measure between two images.

Many other metrics can also be used to compute the query-image and database images similarity measure. Symmetrical and asymmetrical distances can also serve the above purpose. Example of non-symmetrical distance includes the K-L divergence distance. The alternative methods can also be use to for the vector representation which includes logical-based reasoning. These methods can handle complex queries with several attributes. For example, the user may demand images which are similar to image query 1 in shape and query 2 in texture.

3.2.1Color Feature

This method uses the perceptual color changes occur in color space as the boundaries. The RGB color space is divided into sub-spaces also called as color categories. The perceptual color of a pixel can be specified by a color category. The segmentation of the image according to the perceived color involves mapping of all pixels in to their respective categories or color space and then these pixels are grouped according to their color category.

3.2.2Texture Feature

The similar color areas in images can be differentiated by measuring the texture similarity (such as sky and sea, or leaves and grass etc.) by using variety of techniques. Generally the second-order statistics gives the relative brightness of selected pairs of pixels from query and stored images. From these it is possible to calculate measures of image texture such as the degree of contrast, coarseness, directionality and regularity. Alternatively the Gabor filters and fractals are also use in texture analysis for retrieval purpose. Texture queries can be treated same as color queries. The desired texture of query image is selected first and then the system retrieves images with most similar texture measures in value to the query.

The Structural and Statistical methods are used for texture representation. Structural methods include morphological operator and adjacency graph. In this method, the texture is described by identifying structural primitives and their placement rules. The Statistical methods uses Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), Markov random field, fractal model, and multi-resolution filtering techniques such as Gabor and wavelet transform to characterize the texture of the image.

3.2.3 Shape

The shape is not the image shape but it is related to the query. Shapes can be determined by segmentation or edge detection methods or by using some filters to identify given shapes of an image [20]. Shape descriptors should be invariant to translation, rotation, and scaling. Fourier transforms, Moment Invariants can be used as shape descriptors [19].

3.2.4 Segmentation

Image segmentation is simply a meaningful representation of an image for analysis purpose. Image segmentation is helpful in locating the objects and boundaries (lines, curves, etc.) in images which share some common characteristics.

After segmentation, a set of points or contours are extracted from adjacent regions of the image. These set of points or contours are significantly different with respect to some characteristics [26].

3.2.5 Content levels

The multiple content levels are present in images like luminance and color are considered as low-level content, physical objects are regarded as high-level content. Texture and patterns, generally counted as mid-level content. Though there is no sharp cut-off between the levels of content which can be perceived by a human [27].

3.3Query levels

Query is the visual content of user interest used to search the images from large scale images database. These visual content may be described as the feature such as color, texture, shape or the spatial location of image elements or query by example or

Some high level content attributes. Depending upon these features, the queries are classified in different levels.

Level 1: The primary features like color, texture, shape or the spatial location of image elements are use for retrieval. For example, 'find pictures like this'.

Level 2: Some derived features are use for retrieval of objects of given type identified with some degree of logical inference. For example 'find pictures of a flower' and

Level 3: A high-level reasoning about the purpose of the objects or scenes is use for Retrieval. This includes retrieval of pictures with some significance. For example, 'find pictures of joyful people' [31].

3.4Similarity Measure

Similarity or dissimilarity is calculated by using some functions between two images on the basis of features, [1]. Lots of images are stored in database with extracted features. Search parameters from an image are compared with stored one and similarity measure is computed by using some function. There are different classes of similarity measures like color similarity, texture similarity, shape similarity, and object and relationship similarity.

3.5Taxonomy of Images

Basically images are broadly divided as intensity images, indexed images, scaled images, binary images and Intensity Images. An image is represented as a matrix where every element has a value of intensity level at the corresponding position.

3.5.1Indexed Images

Two matrices are used to store indexed images. The first matrix contains one number for each pixel. The size of this matrix is same as the original image size. The second matrix is mapping of image pixels in color space, which use for storage. Its size may be different from the image size, [29].

3.5.2Scaled indexed images

A scaled indexed image uses matrix values which are linearly scaled to form lookup table indices. To display a matrix as a scaled indexed image, use the MATLAB image display function imagesc, [30].

3.5.3Binary Images

This image format also stores an image as a matrix where 0 for black and a 1 for white, [28].

3.5.4Intensity Images

In an image matrix, every element has a value corresponding to brightness/darkness of the pixel at the corresponding position. The brightness of the pixel is generally represented in two ways-the double class and uint8 (or data type). The double class assigns a floating number ("a number with decimals") between 0 and 1 to each pixel and uint8 assigns an integer between 0 and 255 to represent the brightness of a pixel, [28].

3.6. Descriptor

Image descriptors are characteristic features used for describing local image regions. The vector of image pixels is a simplest descriptor. Similarity between two regions of the image can be measure by using Cross-correlation. But the computational complexity of this technique is very high due to high dimensionality. This computational complexity can be reduce by sub-sampled point neighborhood,[32].

Basically descriptors are classified as distribution-based descriptor, non- parametric transformations, spatial-frequency techniques, differential techniques, Scale Invariant Feature Transform (SIFT) detector and Visual Descriptors.

3.6.1Distribution-based descriptors

Histogram is a simplest distribution-based descriptor that gives the pixel intensities distribution in image. Johnson and Hebert in the context of 3D object recognition introduced a representation (spin image), which is generated using a histogram, [33].

3.6.2Non-parametric transformations

Zabih and Woodfill developed an approach which uses non- parametric statistics (local transforms). It is based on the information about ordering and reciprocal relations between the data, rather than the data values [34].

3.6.3Spatial-frequency techniques

The Fourier transform describes the image content into the basic functions in frequency domain. In this method the spatial relations between points are not explicit and the basic functions are infinite, therefore difficult to adapt to a local approach. These problems can be overcome by using the Gabor transform. The description of small changes in frequency and orientation is high dimensional. Gabor filter and wavelets are frequently used in texture classification [35].

3.6.4Differential descriptors

A point neighborhood can be estimated by computing the set of image derivatives up to a given order. Koenderink investigated local derivatives (local jet) properties while differential invariants derived by Florack.The steerable filters are developed by Freeman and Adelson.These filters steer the derivatives in a particular direction to estimate the components of the local jet. The derivatives construct by Gaussian convolution will be used to construct discrete operators which give the salient differential geometry in the scene [33].

3.6.5Scale Invariant Feature Transform (SIFT) detector

Key points are extracted by the SIFT detector and their descriptors are computed by the SIFT descriptor [26]. A key point (or interest point) is defined by some particular image, intensities around it, such as a corner. A key point can be used for deriving a descriptor. The SIFT detector is invariant to translation, rotations, and rescaling of the image.

3.6.6GIST Descriptor

GIST is a global feature descriptor and it saves the time of computation than local descriptors. The inherent issues with massive image databases are long time interval to read all the image information (Long interval particularly with a per-pixel analysis). To avoid these types of issues, descriptors are used. Descriptor is a priory available data that's used to identify objects in retrieval system. It saves image features that are required for feature search, comparison and decision making. GIST typically means quintessence. It provides essence of an image. it was created for recognition of comparable scenes, like mountains, tall buildings, streets. GIST defines the features that separates a scene from the rest and represent dominant spatial structure of a scene [38].

3.6.7Visual Descriptors

Visual descriptors are general information descriptors which give a description about color, shape, regions, and motion which is useful to extract the information about scene content. A best example is the face recognition application [37].

3.6.8General information descriptors

General information descriptors consist of a set of descriptors that describe the elementary characteristics such as color, texture, shape, motion, location and others. By using signal processing these descriptions can be derived [34].

3.6.8.1Color Descriptor

It gives the notable characteristics of scene content. Generally tools like Dominant Color Descriptor (DCD), Scalable Color Descriptor (SCD), Color Structure Descriptor (CSD), Color Layout Descriptor (CLD), and Group of frame (GoF) or Group-of-pictures (GoP) etc. are used to describe the color, [35].

3.6.8.2Texture Descriptor

Texture is one of the most important qualities in image representation. Texture descriptors are derived by computing the region homogeneity and the histograms of region borders. A set of descriptors is formed by Homogeneous Texture Descriptor (HTD), Texture Browsing Descriptor (TBD), and Edge Histogram Descriptor (EHD), [33].

3.6.8.3Shape Descriptor

Human can recognize the objects through their shape. Shape contains important semantic information of an object. Human visual system implements segmentation to extract the information from the scene. Such a segmentation system like human visual system is not available yet. However a series of algorithms serve this purpose. These descriptors describe regions, contours and shapes for 2D images and for 3D volumes. The shape descriptors are the defined by three descriptor of: Region-based Shape Descriptor (RSD), Contour-based Shape Descriptor (CSD), and 3-D Shape Descriptor (3-D SD), [36].

3.6.8.4Motion Descriptors

This descriptor is use to describe the motion activity like camera motion and the object motion in the sequence. The motion descriptors describe motion in video sequence. Motion Activity Descriptor (MAD), Camera Motion Descriptor (CMD), Motion Trajectory Descriptor (MTD), and Warping and Parametric Motion Descriptor (WMD and PMD), are the examples of motion descriptors [37].

3.6.8.5Location Descriptor

It gives the location in spatial domain. The elements can also be located in the temporal domain also. The descriptor set are categories as: Region Locator Descriptor (RLD), and Spatio Temporal Locator Descriptor (STLD), [34].

3.7. Similarity and performance measurement parameters

Performance measurement means the estimation of the parameters under which the system is reaching the targeted results.

3.7.1Similarity Feature Extraction

The extracted features from query image are compared on the basis of similarity with the features of images from database. For this comparison three types of measurements are used [14].

3.7.1.1Euclidean Distance

It is the displacement of a pixel from the nearest background point. The equation of Euclidean distance is given as [14]

$$D(p,q) = \sqrt{\sum_{i=1}^{n}[(q_i - p_i)]^2} \quad (1)$$

3.7.1.2Chi Square Distance

Chi Square Distance is useful for histogram matching. It is used as a kernel in SVM for image classification. It has the distributional equivalence property. It ensures that the distances between rows and columns are invariant when two columns (or two rows) with identical profiles are aggregated [14].Greater differences between expected and actual data produce a larger Chi-square value

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \qquad (2)$$

3.7.1.3Weighted Euclidean Distance

Multiply squared differences by corresponding weights are known as Weighted Euclidean Distance [14].The weighted Euclidean distance-based approach (WEDBA) is based on the weighted distance of alternatives from the most and least favorable situations, respectively.

$$d(A,B) = \sqrt{\sum_i w_i (A_i - B_i)^2} \qquad (3)$$

Where A_i is i^{th} feature of A and W_i is the weight given to feature i.

3.7.2Performance Parameter

The retrieval process Evaluation is a crucial issue in CBIR. So many methods are available for retrieval system performance measurement. Precision and Recall are the most common performance parameters.

3.7.2.1Precision

Precision rate is defined as a ratio of number of retrieve relevant images similar to the query to the total number of retrieved images in response to query [21][22[23].

$$precision = \frac{[\{relevant\_documents\} \cap \{retrieved\_documents\}]}{\{retrieved\_documents\}}$$

3.7.2.2Recall

Recall rate is defined as a ratio of number of retrieve relevant images similar to the query to the total number of relevant images available in the database [21][22[23].

$$recall = \frac{[\{relevant\_documents\} \cap \{retrieved\_documents\}]}{\{relevant\_documents\}}$$

It can be looked at as the probability that a relevant document is retrieved by the query. Recall alone is not enough but one needs to measure the number of non-relevant documents also, for example by computing the precision.

3.7.2.3Fall-out

The proportion of non-relevant documents that are retrieved out of all non relevant documents available-

$$fall\_out = \frac{[\{non\_relevant\_documents\} \cap \{retrieved\_documents\}]}{\{non\_relevant\_documents\}}$$

It can be looked at as the probability that a non-relevant document is retrieved by the query.

3.7.2.4 F-measure

It is the weighted harmonic mean of precision and recall. The traditional F-measure or balanced F-score is given as:

$$F = \frac{2 \times precision \times recall}{(precision + recall)}$$

The general formula for non-negative real β is:

$$F_\beta = \frac{(1+\beta^2).(\mathrm{precision} \times \mathrm{recall})}{(\beta^2.precision + recall)}$$

Two other commonly used measures are the measure F_2, which weights recall twice as much as precision and the $F_{0.5}$ measure, which weights precision twice as much as recall. The F-measure was derived by (Rijsbergen, 1979) so that Fβ "measures the effectiveness of retrieval with respect to a user who attaches β times as much importance to recall as precision". It is based on Rigsbergen's effectiveness measure.

$$E = 1 - \frac{1}{\frac{\alpha}{p} + \frac{1-\alpha}{R}}$$

Their relationship is

$$F_\beta = 1 - \mathrm{E}$$

Where

$$\alpha = \frac{1}{1+\beta^2}$$

3.7.2.5 R-Precision

It is defined as the precision after R images have been retrieved by the system where R is also the total number of retrieved images for the given query. This measure is highly correlated to Average Precision.

3.7.2.6Mean average precision

It is the mean of the average precision scores for each query. Where Q is the number of queries

$$MAP = \frac{\sum_{q=1}^{Q} Ave[P(q)]}{Q}$$

3.7.2.7Ground-truth

It is a term related to location based information collection in remote sensing. The collection of such information is useful in remote-sensing data calibration. A team will be collect detailed samples from the sites. Scientists can use this data to verify and update the existing data and maps [24].

4. Conclusion

Content-based image retrieval is very much popular among various applications. Accuracy and performance time are main concern for CBIR. In this paper, a survey has been done on some techniques of feature extraction and the different descriptors used in CBIR. Some of discussed techniques got few drawbacks like some important feature information loss, more processing time, computational complexity etc. It is found that if more time required completing the retrieval task, the effectiveness of overall system degrades. Thus, an effective, relevant image feature extraction is needed to get better results efficiently. Therefore, in our future work is going to design a new better approach for better result optimization of content-based image retrieval system.
The future work is to concentrate on different profound learning systems and to satisfy the semantic gap between machine observation and human discernment for image retrieval. The retrieval of images can be further extended to more number of images of different categories along with other combination of different features.

References-

[1] Romari Pighetti, Denis Pallez and Frederic Precioso. “Hybrid Content Based Image Retrieval combining Multi Objevtive Genetic Algorithm and SVM”.*21st International Conference on Pattern Recognition (ICPR 2012)*, 2849-2952, Nov 11-15, 2012.

[2]M. Arevalillo-Herrez, F. J. Ferri, and S. Moreno-Picot. “Distance-based relevance feedback using a hybrid interactive genetic algorithm for image retrieval”.*Applied Soft Computing*, 11(2):1782 – 1791, 2011.

[3] E. Chang, S. Tong, K. Goh, and C. Chang. “Support vector machine concept-dependent active learning for image retrieval”.*IEEE Trans. on Multimedia*, 2, 2005.

[4] M. Crucianu, D. Estevez, V. Oria, and J.-P. Tarel. “Hyperplane queries in a feature-space m-tree for speeding up active learning”. In *BDA*, 2007.

[5] C. D. Ferreira, J. A. Santos, R. da S. Torres, M. A. Gonc¸alves, R. C. Rezende, and W. Fan.” Relevance feedback based on genetic programming for image retrieval.*Pattern Recogn”. Lett*., 32(1):27–37, Jan 2011.

[6] P. Gosselin and M. Cord. “Active learning methods for interactive image retrieval”. *Image Processing, IEEE Transactions on*, 17(7):1200–1211, july 2008.

[7] B. Kulis and K. Grauman. “Kernelized locality-sensitive hashing for scalable image search”. In *IEEE ICCV*, pages 2130 – 2137, 29 2009-oct. 2 2009.

[8] C.-C. Lai and Y.-C. Chen. “A user-oriented image retrieval system based on interactive genetic algorithm”.*Instrumentation and Measurement, IEEE Transactions on*,60(10):3318 –3325, oct. 2011.

[9]S.-F.Wang, X.-F.Wang, and J. Xue. “An improved interactive genetic algorithm incorporating relevant feedback”. In*IEEE ICMLC*, volume 5, pages 2996–3001,aug.2005. 267, 2006.

[10] Romari Pighetti, Denis Pallez and Frederic Precioso. “Hybrid Content Based Image Retrieval combining Multi Objevtive Genetic Algorithm and SVM”.*21st International Conference on Pattern Recognition (ICPR 2012)*,2849-2952,Nov 11-15,2012.

[11]M. Arevalillo-Herrez, F. J. Ferri, and S. Moreno-Picot. “Distance-based relevance feedback using a hybrid interactive genetic algorithm for image retrieval”. *Applied Soft Computing*, 11(2):1782 – 1791, 2011.

[12] E. Chang, S. Tong, K. Goh, and C. Chang. “Support vector machine concept-dependent active learning for image retrieval”.*IEEE Trans. on Multimedia*, 2, 2005.

[13] M. Crucianu, D. Estevez, V. Oria, and J.-P. Tarel. “Hyperplane queries in a feature-space m-tree for speeding up active learning”. In *BDA*, 2007.

[14] Usama Sayed, M. A. Mofaddel, W. M. Abd-Elhafiez and M. M. Abdel-Gawad “Image Object Extraction Based on Curvelet Transform” Natural Sciences Publishing Cor. Appl. Math. Inf. Sci. 7 No. 1, 133-138 ,2013.

[15] Li, J., & Levine, M. (2012).” Visual Saliency Based on Scale-Space Analysis in the Frequency Domain”. *IEE Trans Pattern*, *35*(4).

[16] Awad, D., Mancas, M., Richie, N., Courboulay, V., & Revel, A. (2015). “A CBIR-Based Evaluation Framework for Visual Attention Models”.

[17] *Datta, Ritendra; Dhiraj Joshi; Jia Li; James Z. Wang (2008). "Image Retrieval: Ideas, Influences, and Trends of the New Age". ACM Computing Surveys.* ***40*** *(2): 1–60. doi:10.1145/1348246.1348248*

[18]Bird, C.L.; P.J. Elliott, Griffiths (1996). "User interfaces for content-based image retrieval".

[19]Rui, Yong; Huang, Thomas S.; Chang, Shih-Fu (1999). “Image Retrieval: Current Techniques, Promising Directions, and Open Issues”. *Journal of Visual Communication and Image Representation*. **10**: 39–62.

[20]Tushabe, F.; M.H.F. Wilkinson (2008).” Content-based Image Retrieval Using Combined 2D Attribute Pattern Spectra”. Springer Lecture Notes in Computer Science. Lecture Notes in Computer Science. 5152. pp. 554–561. doi:10.1007/978-3-540-85760-0_69. ISBN 978-3-540-85759-4

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

[22] Suman Lata and Parul Preet Singh,”A Review on Content Based Image Retrieval System”, International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), ISSN: 2277 128X, Volume 4, Issue 5, May 2014.

[23] Amanbir Sandhu and Aarti Kochhar , “Content Based Image Retrieval Using Texture, Color And Shape for Image Analysis”, International Journal of Computers & Technology (IJCT),Volume 3, No. 1, AUG, 2012.

[24]Ground truth. (2019, August 01). In Wikipedia, the free encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Ground_truth

[25]Zhu, M. (2004). Recall, Precision and Average Precision.

[26]Stockman, G. C., & Shapiro, L. G. (2001). Compter Vision (2001st ed., pp. 279–325). New Jersey: Prentice- Hall.

[27]Black Jr., J. A., Fahmy, G., & Panchanathan, S. (2002). "A method for evaluating the performance of CBIR systems". *Arizona State University*.

[28]Gonzalez, R. C., & Woods, R. . (2010). *Digital Image Processing* (3rd Ed). India: Prentice-Hall.

[29]Palm, W. (2015). Courseware based on MATLAB and Simulink. Retrieved August, 2019, fromhttps://in.mathworks.com/help/images/indexed-images.html?searchHighlight=indexed%20images&s_tid=doc_srchtitle

[30]Danish, M., Rawat, R., & Sharma, R. (2013). "A Survey: Content Based Image Retrieval Based On Color, Texture, Shape & Neuro Fuzzy". *Int. Journal Of Engineering Research And Application*, *3*(5), 839–844.

[31]Liu, Y. (2009). "A survey of content-based image retrieval with high-level semantics". *Elsevier Science Direcct*.

[32]Mikolajezk, K., & Schmid, C. (2010)." A performance evaluation of local descriptors".

[33]Hu, R., & Collommosse, J. (2013)." A Performance Evaluation of Gradient Field HOG Descriptor for Sketch Based Image Retrieval".

[34]Gibbon, D. C., Huang, Q., Liu, Z., Rosenberg, A. E., & Shahraray, B. (2012, March)."System and method for automated multimedia content indexing and retrieval".

[35]Mikolajezk, K., & Schmid, C. (2010)." A performance evaluation of local descriptors".

[36]Manjunath, B. ., Salembier, P., & Sikora, T. (2002). Visual Descriptors. *Wiley and Sons*.

[37]Wu, J., & M, J. (2010). CENTRIST: "A Visual Descriptor for Scene Categorization". *IEEE Transactions on Pattern Analysis and Ma*, *33*(8), 1489 – 1501.

[38]A. Oliva and A. Torralba." Modeling the shape of the scene: A holistic representation of the spatial envelope". Int. J. Comput. Vision, 42(3):145–175, May 2001.