Image Mining in the Context of Content Based Image Retrieval: A Perspective

Nishchol Mishra[1], Dr. Sanjay Silakari[2]

**[1] School of I.T.
Rajiv Gandhi Proudyogiki Vishwavidyalaya,
Bhopal, India**

**[2] Professor& Dean, Comp. Sc. & Engg.,
Rajiv Gandhi Proudyogiki Vishwavidyalaya ,
Bhopal, India**

Abstract

The emergence and proliferation of social network sites such as Facebook, Twitter and Linkedin and other multimedia networks such as Flickr has been one of the major events of this century. These networks have acquired immense popularity and have become a part of the daily lives of millions of people. Many of these network sites are thus extremely rich in content, and contain a tremendous amount of multimedia content waiting to be mined and analyzed. Analyzing this huge amount of multimedia data to discover useful knowledge is a challenging task. It has opened up opportunities for research in Multimedia Data Mining (MDM). Multimedia Data Mining can be defined as the process of finding interesting patterns from media data such as audio, video, image and text that are not ordinarily accessible by basic queries and associated results. This paper mainly focuses on Image Mining techniques and how Content-based Image Retrieval can be helpful for Image mining.

Keywords: *Multimedia Data Mining, Image Mining, Content Based Image Retrieval, Multimedia Information retrieval.*

1. Introduction

Recent explosion in the quantity of multimedia data stored in social networks and other multimedia sites has engendered the need for new and better techniques for accessing data. Indexing and retrieval are at the core of multimedia system design— the knowledge potential of huge quantities of multimedia data may lay unexplored in the absence of effective tools for easy access to the collected information. Once collected, the data must be organized efficiently. The object of the retrieval process is to obtain limited information to meet the needs of a user at a particular time, within a particular domain application. Often it is extremely difficult to achieve this objective in actual practice. A major challenge, therefore, lies in developing techniques that can "interpret" the multimedia contents in large data collections so efficiently as to extract all information items relevant to the user query [6]. Multimedia mining deals with the extraction of implicit knowledge; in other words, it looks for multimedia data relationships or other patterns that are not explicitly stored in multimedia files. Multimedia mining is more than just an extension of data mining; it is an interdisciplinary endeavor that draws upon expertise in computer vision, multimedia processing, multimedia retrieval, data mining, machine learning, database and artificial intelligence. Rapid progress in digital data acquisition and storage technology has led to enormous and fast-growing volumes of data. Valuable information hidden in this data cannot be extracted without the aid of powerful tools because of the overwhelming size and volume of the data. Multimedia mining systems that can automatically extract semantically meaningful information (knowledge) from multimedia files are increasingly in demand. Generally, multimedia database systems store and manage a large and varied collection of multimedia objects such as image, video, audio and hypertext data. Knowledge discovery from multimedia documents thus involves the analysis of non-structured information [3]. To achieve it, we need tools for discovering relationships between objects or segments within multimedia document components; e.g. classifying images based on their content, extracting patterns in sound, categorizing speech and music, and recognizing and tracking objects in video streams. In general, the multimedia files from a database must be first preprocessed to improve their quality. Thereafter they undergo various transformations and a process of features extraction to generate important features from the multimedia files. With the generated features, mining can be carried out using data mining techniques to discover significant patterns. These resulting patterns are then evaluated and interpreted in order to obtain knowledge for the final application.

2. The Characteristics of Multimedia Data Mining

Multimedia database systems store and manage large quantities of multimedia data sets such as audio, video, images, graphics, sounds, text, documents and hypertext data. This multi-media information does not have a uniform structure and a unified approach. Multimedia data is usually multi-dimensional and unstructured or semi-structured, with each medium having its own characteristics, its own way of presentation of information [1]. Apart from presenting information independently, each medium can also express a different characteristic of the same event, so that the media taken together describe the existence, development and result of an event in its entirety. So, there have to be features of information, attributes and the relationships in multimedia data sets that are not within our intuitive grasp. Multimedia data mining involves intelligent data analysis, aimed at finding these features, attributes and relationships in order to construct models for making decisions, taking countermeasures and achieving fusion analysis. "Based on the data stored in them, multimedia databases are used in content-based image retrieval, sound delivery system, video on demand system, World Wide Web and identifying the password command voice based user interface, etc. Multimedia Mining focuses on the following five fronts: Image Mining, Video Mining, Audio Mining, Web Mining and Multi-Media Integrated Mining. Image mining involves the introduction of data mining technology into the image field of study, to discover the information and knowledge hidden in a large quantity of image data. It is the process of identifying hidden, valid, novel, potentially useful, and ultimately understandable semantics of information and knowledge from extensive image data"[1].

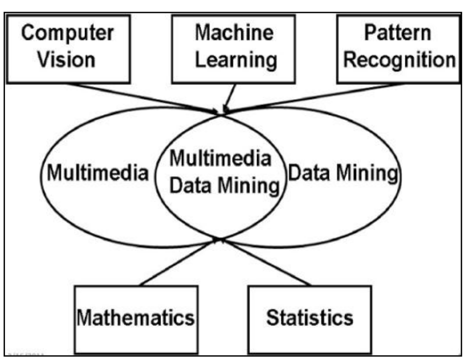

Fig. 1Multimedia Data Mining

The typical data mining process consists of several stages and the overall process is inherently interactive and iterative. The main stages of the data mining process are (1) Domain understanding; (2) Data selection; (3) Data preprocessing, cleaning and transformation; (4) Discovering patterns; (5) Interpretation; and (6) Reporting and using discovered knowledge [2] [15].

3. Image Mining

Over the past ten years or so, significant progress has been made in making computers learn to understand, index, and annotate pictures representing a wide range of concepts. Image mining deals with the extraction of implicit knowledge, that is, image data relationship or other patterns not explicitly stored in the images. Image mining is more than just an extension of data mining to the image domain. The cardinal role of image mining is to discover the means of an effective processing of low-level pixel representations, contained in a raw image or image sequence, to arrive at high-level spatial objects and relationships [2][4]. The focus of image mining is on the extraction of patterns from a large collection of images. While there seems to be some overlap between image mining and content-based retrieval (since both deal with large collections of images), image mining goes beyond the problem of retrieving relevant images. In image mining, the goal is to discover image patterns that are significant in a given collection of images and the related alphanumeric data [2]. The fundamental challenge in image mining is to reveal out how low-level pixel representation enclosed in a raw image or image sequence can be processed to recognize high-level image objects and relationships [3][12][13]. Fig. 2 illustrates the typical Image Mining Process.

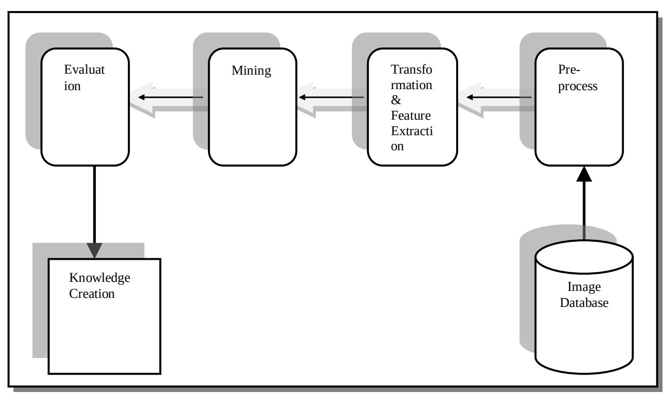

Fig.2 Typical Image Mining Process

3.1 Preprocessing:

In image data, the spatial segmentation can be done at region and/or edge level based on the requirements of the application. It can be automatic or manual and should be approximate enough to yield features that can reasonably capture the image content.

3.2 Feature Extraction and Transformation:
Color, edges, shape, and texture are the common image attributes that are used to extract features for mining. Feature extraction based on these attributes may be performed at the global or the local level. There are obvious trade-offs between global and local descriptors. Global descriptors are generally easy to compute, provide a compact representation, and are less prone to segmentation errors; but they tend to integrate and, therefore, are often unable to discover subtle patterns or changes in shape. Local descriptors, on the other hand, tend to generate more elaborate representations and can yield useful results even when part of the underlying attribute.

3.3 Image Mining Techniques:
a**) Classification:** Classification as a technique for analyzing multimedia data relies on the properties of a given set of multimedia data divided into predefined class labels. Intelligently classifying images by content is an important way to mine valuable information from large image collections. There are two major types of classifiers - parametric classifier and non-parametric classifier. Data classification can be achieved through the following two-step process: (1) Establishing classifiers or describing the predefined data types or concept sets, (2) Using models to classify data [1]. Commonly used classification tools are decision tree classification method, rule-based classification, neural networks, support vector machines, Naïve Bayes classification etc.

b) Clustering: In image clustering (also called unsupervised classification) images are grouped into meaningful clusters on the basis of similarity and not on the basis of known structures or labels. The problem here is to find groups and structures which are similar, without *a priori* knowledge of predefined data types (hence the name 'unsupervised classification'). Cluster analysis makes a data object decompose or divide into multiple classes, or clusters, so that the same class of data objects has a high similarity, but is as different as possible from other types of data. A cluster is also a collection of data objects for analysis, but clustering, unlike classification, does not make use of predefined data types derived from known class labels of training data sets [1].

3.4 Discussion
Image mining applications started with the use of data mining methods on low-level image features. The limitations of the approach soon became apparent. Firstly, this approach is normally used to classify or cluster only a small number of image categories and Secondly, it is often difficult to generalize results obtained by methods using low-level visual features and apply them to additional image data beyond the training set [1].

Thus recent research is focused on developing image mining frameworks that can handle heterogeneous types of images on a large scale automatically.

CBIR: An Introduction & Survey
R. Datta et al. [12] defines CBIR as : Content based image retrieval is a technology that in principle helps to organise digital pictures archives by their visual content, by this definition anything ranging from image similarity function to a robust image annotation engine falls under the purview of Content based image retrieval. Smeulders et al. [14] in their paper define problem with all current approaches is the semantic gap, between low level content and higher level concepts. A. Yoshitaka et al. [11] in their paper used content based retrieval for multimedia databases. In this paper they mention CBR is not an individual entity but relies on underlying data model, a priori knowledge of the area of interest and the scheme for representing queries. M. S. Lew et al. [13] in their paper presented a comprehensive survey of the latest CBIR techniques; they proposed prevalent research topics which have potential for improving Multimedia Retrieval by bridging the semantic gap are as: Human centred computing, New Features, New Media, Browsing and Summarization, and Evaluation & Benchmarking. S. Nandgopalan et al. [8] proposed a novel approach for generalized image retrieval based on semantic concepts like color, texture and edge histogram descriptor. S. Silakari et al. [7] Proposed framework focuses on color as feature, Color Moment and Block Truncation Coding (BTC) are used to extract features for image dataset. K.K Pandey et al. [16] proposed a new matching technique to find the similar value between query color image and database color image using histogram, spatiogram and bins, their method uses RGB and HSV color space. T.N. Manjunath et al. [5] conducted survey on current state of Multimedia data mining and Knowledge discovery approaches techniques for mining multimedia data.

4. Image Mining using CBIR Concepts:
This section mainly focuses on how low level features of image can be extracted and can be used effectively for Image mining purpose. In subsequent sections extraction techniques are discuss in details and how Indexing and Retrieval can be improved. To overcome the disadvantages inherent in a text-based retrieval system, content-based image retrieval (CBIR) was introduced in the early 1980s. Content-based image retrieval (CBIR) is the application of computer vision to the image retrieval problem and deals with the problem of searching for digital images in large databases. CBIR generally works on the basis of querying using an example image or a part of an image. Various types of algorithm are developed and these algorithms may vary depending on the application, but result images should all share common elements with

the provided example. Content-based image retrieval uses the visual contents of an image such as color, shape, texture, and spatial layout to represent and index the image. In a Generalized CBIR system (Figure 3), the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities /distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image database [9]. All the mathematical expressions defined in subsequent sections and their descriptions and nomenclature are taken from F. Long et al. [9].

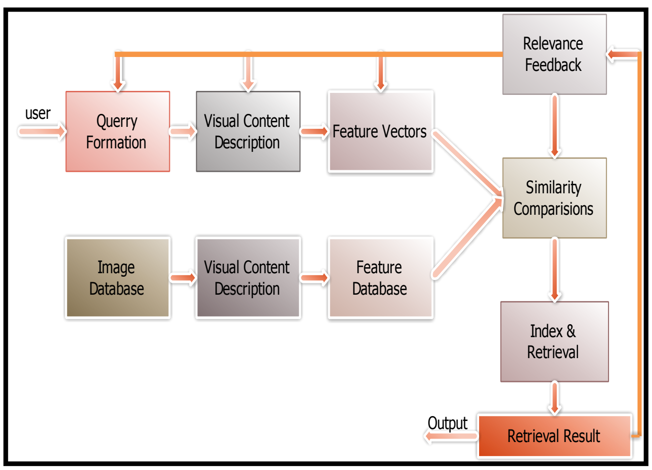

Fig. 3 Generalized CBIR Systems

4.1 Image Content Descriptors:

Image content may include both visual and semantic content. Visual content can be very general or domain specific. General visual content include color, texture, shape, spatial relationship etc. Domain specific visual content, such as human faces, is application dependent and may involve domain knowledge. Semantic content is obtained either by textual annotation or by complex inference procedures based on visual content. A visual content descriptor can be either global or local. A global descriptor uses the visual features of the whole image, whereas a local descriptor uses the visual features of regions or objects to describe the image content [9].

Color

Color property is one of the most widely used visual features in content-based image retrieval (CBIR) systems. Researches in this field fall in three main subareas: (a) definition of adequate color space for a given target application, (b) proposal of appropriate extraction algorithms, and (c) study/evaluation of similarity measures.

Color Histogram

It is the most commonly used descriptor in image retrieval. The color histogram is easy to compute and effective in characterizing both the global and the local distribution of colors in an image. The color histogram extraction algorithm involves three steps: partition of the color space into cells, association of each cell to a histogram bin, and counting of the number of image pixels of each cell and storing this count in the corresponding histogram bin. This descriptor is invariant to translation and rotation. The similarity between two color histograms can be performed by computing the L1, L2, or weighted Euclidean distances.

Color Moments

Color moments have been successfully used in many retrieval systems such as QBIC. The first order (mean), the second order (variance) and the third order (skewness) color moments have been proved efficient in representing color distributions of images. They are defined as:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij} \qquad (1)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^2\right)^{1/2} \qquad (2)$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^3\right)^{\frac{1}{3}} \qquad (3)$$

Where f_{ij} is the value of the i-th color component of the image pixel j, and N is the number of pixels in the image.

Usually, color moments can be used as the first filter to cut down the search space before other sophisticated color features are used for retrieval. Examples of color descriptors that incorporate color spatial distribution include *Color Coherence Vector (CCV), Border/Interior Pixel Classification (BIC), and Color Correlogram* . The color correlogram was proposed to characterize not only the color distributions of pixels, but also the spatial correlation of pairs of colors. The first and the second dimension of the three-dimensional histogram are the colors of any pixel pair and the third dimension is their spatial distance. Compared to the color histogram and CCV, the color auto correlogram provides better retrieval results, but is also the most computationally expensive due to its high dimensionality. In the BIC approach, each image pixel is classified as a border or interior pixel, if it is at the border of the image itself or if at least one of its 4 neighbors has a different color [10].

The MPEG-7 initiative, formally known as Multimedia Content Description Interface, focuses on the description of multimedia content, including content of various modalities such as image, video, speech and graphic. One of the most important components of the MPEG-7 framework is the proposing of image descriptors. For the color property, MPEG-7 has defined a number of histogram descriptors, a dominant color descriptor, and a color layout descriptor [10].

Texture

Texture is the property of an image, characterized by the existence of basic primitives whose spatial distribution creates some visual patterns defined in terms of granularity, directionality, and repetitiveness. Generally, texture representation methods are classified into two main categories: structural and statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primaries and their rules. Statistical methods, including Tamura feature, shift-invariant principal component analysis (SPCA), Fourier power spectra, Wold decomposition, co-occurrence matrices, Markov random field, fractal model and multi-resolution filtering techniques such as Gabor and wavelet transform, define texture by the statistical distribution of the image intensity [9].

Tamura Features

Today's CBIR systems mostly use a set of six visual features based on psychological experiment, namely, coarseness, contrast, directionality, linelikeness, regularity and roughness.

Coarseness

Coarseness relates to distances of notable spatial variations of grey levels, that is, implicitly, to the size of the primitive elements (texels) forming the texture. Coarseness is a measure of the granularity of the texture. To calculate the coarseness, moving averages $A_k(x,y)$ are computed, At each pixel (x,y), compute six averages for the windows of size $2^k \times 2^k$, (k=0,1,...,5) around the pixel.

$$A_k(x,y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i,j)/2^{2k} \tag{4}$$

Where g (i,j) is the pixel intensity at (i,j).

At each pixel, compute absolute differences $E_k(x,y)$ between the pairs of no overlapping averages in the horizontal and vertical directions.

$$E_{k,h}(x,y) = |A_k(x+2^{k-1},y) - A_k(x-2^{k-1},y)| \tag{5}$$
$$E_{k,v}(x,y) = |A_k(x,y+2^{k-1}) - A_k(x,y-2^{k-1})| \tag{6}$$

At each pixel, find the value of k that maximizes the difference E_k (x,y) in either direction and set the best size $S_{best}(x,y)=2^k$.

$$S_{best}(x,y) = 2^k \tag{7}$$

Compute the coarseness feature F_{crs} by averaging S_{best} over the entire image

$$F_{crs} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} S_{best}(i,j) \tag{8}$$

Contrast

Contrast measures how grey levels vary in the image and to what extent their distribution is biased towards black or white. The Contrast is defined as:

$$F_{con} = \frac{\sigma}{\alpha_4^{1/4}} \tag{9}$$

Where the kurtosis $\alpha_4 = \mu_4/\sigma^4$, μ_4 is the fourth moment about the mean, and σ^2 is the variance.

Directionality

Degree of directionality is measured using the frequency distribution of oriented local edges against their directional angles. To compute the directionality, the image is convoluted with two 3x3 arrays and gradient vector at each pixel is computed.

$$\begin{matrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{matrix} \quad and \quad \begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{matrix}$$

The magnitude and angle of this vector are defined as

$$|\Delta G| = (|\Delta_H| + |\Delta_V|)/2 \tag{10}$$
$$\theta = tan^{-1}(\Delta_V/\Delta_H) + \pi/2 \tag{11}$$

Where Δ_H and Δ_V are the horizontal and vertical differences.

Then, by quantizing θ and counting the pixels with the corresponding magnitude $|\Delta G|$ larger than a threshold, a histogram of, denoted as H_D, can be constructed. The histogram is relatively uniform for images without strong orientation and exhibits peaks for highly directional images. The degree of directionality relates to the sharpness of the peaks:

$$F_{dir} = \sum_{p}^{n_p} \sum_{\emptyset \in W_p} (\emptyset - \emptyset_p)^2 H_D(\emptyset) \tag{12}$$

In this sum p ranges over n_p peaks; and for each peak p, w_p is the set of bins distributed over it; while $\emptyset_p$ is the bin that takes the peak value.

Regularity

The regularity feature is defined as $F_{reg}=1-r(s_{crs}+s_{con}+s_{dir}+s_{lin})$ where r is a normalizing factor and each *s*... means the standard deviation of the corresponding feature *F*... in each sub image the texture is partitioned into.

Roughness
The roughness feature is the summation of coarseness and contrast measures: $F_{rgh}=F_{crs}+F_{con}$. The first three features are mostly used in many CBIR Systems.

Shape
Many content-based image retrieval systems use shape features of object or region. Shape features are usually described after images have been segmented into regions or objects as compared with color and texture features. The most frequently used methods for shape description can be boundary-based (rectilinear shapes, polygonal approximation, finite element models and Fourier-based shape descriptors) or region-based (statistical moments). A good shape representation feature for an object should be invariant to translation, rotation and scaling.

Moment Invariants
Moment invariants are used for classical shape representation. Central moments of order p+q for the shape of object R are defined as:

$$\mu_{p,q} = \sum_{(x,y)\in R}(X - X_C)^p \ (Y - Y_C)^q \tag{13}$$

where R is the object represented as binary image and (X_c, Y_c) is the center of the object. This central moment can be normalized to be scale invariant :

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^{\gamma}} \ , \gamma = \frac{p+q+2}{2} \tag{14}$$

Depending on these moments translation, rotation and scale moments can be derived.

4.2 Similarity and Indexing Schemes
Content-based image retrieval system usually calculates visual similarities between a query image and images in a database. In recent years many similarity measures have been developed for image retrieval based on empirical estimates of the distribution of features. Different similarity/distance measures will affect retrieval performances of an image retrieval system significantly. Here we will discuss some of the commonly used methods.

Similarity Measures

Mahalanobis Distance
Mahalanobis distance is a distance measure introduced by P. C. Mahalanobis in 1936 [17]. It is based on correlations between variables by which different patterns can be identified and analyzed. It gauges *similarity* of an unknown sample set to a known one.

$$D(I,J) = \sqrt{(F_I - F_J)^T C^{-1} (F_I - F_J)} \tag{15}$$

We denote *D (I, J)* as the distance measure between the query image I and the image J in the database; and $f_i(I)$ as the number of pixels in bin i of I .

If feature dimensions are independent then Mahalanobis distance can be simplified as, in that case only a variance of each feature component $c_{i,}$ is needed:

$$D(I,J) = \sum_{i=1}^{N} (F_I - F_J)^2 / c_i \tag{16}$$

Minkowski-Form Distance
Minkowski-form distance is the most widely used method for image retrieval. If each dimension of image feature vector is independent of any other and is of equal importance, the Minkowski-form distance L_p is appropriate for calculating the distance between two images [17]. This distance is defined as:

$$D(I,J) = \left(\sum_i |f_i(I) - f_i(J)|^p\right)^{1/p} \tag{17}$$

When p=1, 2, and ∞, *D* (*I*, *J*) is the *L1*, *L2* (also called Euclidean distance), and L_∞ distance respectively. The intersection of the two histograms of I and J is defined as:

$$S(I,J) = \frac{\sum_{i=1}^{N} min(f_i(I), f_i(J))}{\sum_{i=1}^{N} f_i(J)} \tag{18}$$

Quadratic Form (QF) distance:
Quadratic form distance can lead to perceptually more desirable results than Euclidean distance and histogram intersection method as it considers the cross similarity between colors. Quadratic form distance has been used in many retrieval systems for color histogram-based image retrieval [17]. It is given as:

$$D(I,J) = \sqrt{(F_I - F_J)^T A (F_I - F_J)} \tag{19}$$

Where *A*=[*aij*] is a similarity matrix, and *aij* denotes the similarity between bin *i* and *j*. F_I and F_J are vectors that list all the entries in *fi*(*I*) and *fi*(*J*).

Indexing
Efficient indexing is critical to the building and functioning of very large text-based databases and search engines. Effective indexing is an important issue in content-based image retrieval. Because the feature vectors of images tend to have high dimensionality and are not well suited to traditional indexing structures, dimension reduction is usually used before setting up an efficient indexing scheme. One of the techniques commonly used

for dimension reduction is principal component analysis (PCA). It is an optimal technique that linearly maps input data to a coordinate space such that the axes are aligned to reflect the maximum variations in the data. In addition to PCA, many researchers have used Karhunen-Loeve (KL) transform to reduce the dimensions of the feature space. Apart from PCA and KL transformation, neural network has also been demonstrated to be a useful tool for dimension reduction of features. After dimension reduction, the multi-dimensional data are indexed. A number of approaches have been proposed for this purpose, including R-tree (particularly, R*-tree), linear quad-trees, K-d-B tree and grid files [9].

User Interaction

For content-based image retrieval, user interaction with the retrieval system is very important because modification of queries can only be obtained by involving the user in the retrieval process. User interfaces in image retrieval systems typically consist of a query formulation part and a result presentation part. Specifying what kind of images a user wishes to retrieve from the database can be done in many ways. Commonly used query formations are: category browsing, query by concept, query by sketch, and query by example.

Relevance Feedback

Human perception of image similarity is subjective, semantic, and task-dependent. Although content-based methods hold promises for image retrieval generally, retrieval results based on similarities of pure visual features may not always be meaningful in a perceptual or semantic way. Also, each type of visual feature tends to capture only one aspect of image property and it is usually hard for a user to specify clearly how different aspects are combined. To address this problem, interactive relevance feedback, a technique in traditional text-based information retrieval systems, was introduced. With relevance feedback, it is possible to establish a link between high-level concepts and low-level features [9]. Relevance feedback is a supervised active learning technique used to improve the effectiveness of information systems. The main idea is to use positive and negative examples from the user to improve system performance. For a given query, the system first retrieves a list of ranked images according to a predefined similarity metrics. The user marks the retrieved images as relevant (positive examples) to the query or not relevant (negative examples). The system will refine the retrieval results based on the feedback and present a new list of images to the user. Thus the key issue in relevance feedback is how to incorporate positive and negative examples to refine the query and/or to adjust the similarity measure [9].

Evaluation:

To evaluate the performance of the retrieval system, two measurements, namely, recall and precision, are used. For a query q, the data set of images in the database that are relevant to the query q is denoted as R(q), and the retrieval result of the query q is denoted as Q(q). The precision of the retrieval is defined as the fraction of the retrieved images that are indeed relevant for the query:

$$precision = \frac{|Q(q) \cap R(q)|}{|Q(q)|} \quad (20)$$

The recall is the fraction of relevant images that is returned by the query:

$$recall = \frac{|Q(q) \cap R(q)|}{|R(q)|} \quad (21)$$

Discussion

To automate the process of Image annotation, gap between low level features of an image and high level semantics must be reduced, anything ranging from image similarity function to a robust image annotation engine falls under the purview of Content based image retrieval. CBIR, Image Mining, Image Annotation etc. are interlinked areas of research, and is promising area. This paper tries to put a perspective view in the area of Image Mining and Content based Image Retrieval.

5. Conclusion

In recent years there has been tremendous growth in the quality (resolution and color depth), nature (dimensionality) and throughput (rate of generation) of the images acquired and this trend of upward growth is likely to continue. This poses ever changing challenges to image retrieval research. With the advent of very large-scale images (e.g., Google and Yahoo!, Facebook, Linkedin, Flickr), and high resolution, high dimension biomedical and astronomical imagery, often captured at high throughput, image retrieval research faces an immediate challenge of incorporating the ability to make high-resolution, high-dimension, and high-throughput images searchable by content. Meanwhile, we do hope that the quest for a robust and reliable image understanding technology will continue. Together these will constitute the agenda for future research. The future of CBIR will depend on the progress made in each aspect of image retrieval, and the extent to which the individual user benefits by it.

References

[1] Yang Changchun, Yi Li, "A Data Mining Model and Methods Based on Multimedia Database," *Internet*

Technology and Applications, 2010 IEEE International Conference on, vol., no., pp.1-4, 20-22 Aug. 2010.

[2] C. Amitkumar Bhatt, Mohan S. Kankanhalli , “Multimedia data mining : state of the art and challenges,” *Multimedia Tools and Applications*, Springer, vol. no. 51, issue 1,2011, pp. 35-76.

[3] OR Zaïane, J Han, ZN Li, Je Hou, “Mining Multimedia Data”, Proceedings CASCON '98 Proceedings of the 1998 conference of the Centre for Advance Studies on Collaborative research, ACM IBM Press ©1998

[4] Hsu Wynne, Lee Mong Li, Zhang Ji, “Image Mining: Trends and Developments” *Journal of Intelligent Information Systems*,Springer Netherlands, vol. 19, issue 1, 2002, pp. 7-23.

[5] T. N. Manjunath, R. S. Hegadi, and G. K. Ravikumar, “A Survey on Multimedia Data Mining and Its Relevance Today” , *Journal of Computer Science*, vol. 10, no. 11, 2010, pp. 165-170.

[6] Z. Liu, “A Survey on Social Image Mining,” Communications in Computer and Information Science, Springer, 2011, Volume 134, pp. 662-667.

[7] S. Silakari, M. Motwani, M. Maheshwari, “Color Image Clustering using Block Truncation Algorithm”, *IJCSI International Journal of Computer Science Issues*, Vol. 4, No. 2, 2009, pp. 31-35.

[8] S. Nandagopalan, B. S. Adiga, and N. Deepak, “A Universal Model for Content-Based Image Retrieval,” *WASET*, 2008, pp. 644-647.

[9] Fuhui Long, H. Zhang, D. Dagan Feng: "Fundamentals of Content-Based Image Retrieval", Book Chapter, *Multimedia Information Retrieval and Management*, Springer, 2003.

[10] S. Torres and A. X. Falcão, “Content-Based Image Retrieval : Theory and Applications”, RITA, Vol. 13, 2006, pp. 165-189.

[11] A. Yoshitaka, T. Ichikawa, “A Survey on Content-Based Retrieval for Multimedia Databases”, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11, No.1, 1999, pp. 81-93.

[12] R. Datta, D. Joshi, J. Ali, J. Z. Wang, “Image Retrieval: Ideas, Influences, and Trends of the New Age”, *ACM Computing Surveys*, Vol. 40, No. 2, Article 5, 2008, pp. 5:1-5:59.

[13] Michael S. Lew, N. Sebe, C. Djeraba, R. Jain, “Content-Based Multimedia Information Retrieval: State of the Art and Challenges”, *ACM Transactions on Multimedia Computing, Communications and Applications*, Vol.2, No. 1, 2006, pp. 1-19.

[14] Arnold W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, “Content-Based Image Retrieval at the End of the Early Years”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, 2000, pp.1349-1380.

[15] M. Bouet, M.-Aude Aufaure, “New Image Retrieval Principle : Image Mining and Visual Ontology,” *Book Chapter*, Multimedia Data Mining and Knowledge Discovery, Springer London, 2007, pp. 168-184.

[16] Krishna K. Pandey, N. Mishra, H. Sharma, “Enhanced of colour matching algorithm for image retrieval”, *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 3, No. 2, 2011, pp. 529-532.

[17] Http://en.wikipedia.org/wiki/Content-based_image_retrieval.