A STUDY OF COLOR, TEXTURE FEATURE OF IMAGE RETRIEVAL IN IMAGE MINING APPLICATIONS

Hlaing Htake Khaung Tin
Faculty of Information Science
University of Computer Studies, Yangon, Myanmar
Email - hlainghtakekhaungtin@gmail.com

Abstract: *Image mining is the process of searching and discovering valuable information and knowledge in large volumes of data. Some of the methods used to gather knowledge are, Image Retrieval, Data Mining, Image Processing and Artificial Intelligence. Image retrieval system is used in a wide range of applications such as geography, medical, architecture, advertising, design, military and albums. Image mining is an extended field of image and data mining that is concerned with the process of knowledge discovery concerning digital images. The main aim of this paper is to present a study of color, texture feature of image retrieval in image mining applications. The further study of image mining finds much application such as like satellite image mining, weather prediction etc., can be used in various image retrieval based application.*

Key Words: *Image Mining, image retrieval, text based, content based, color feature, text feature.*

1. INTRODUCTION:

Multimedia, digital media and image mining are perhaps among the top ten most overused terms in the last decade. The field of multimedia and digital media is at the intersection of several major fields, including computing, telecommunications, desktop publishing, digital arts, the television/movie/game/broadcasting industry, audio and video electronics.

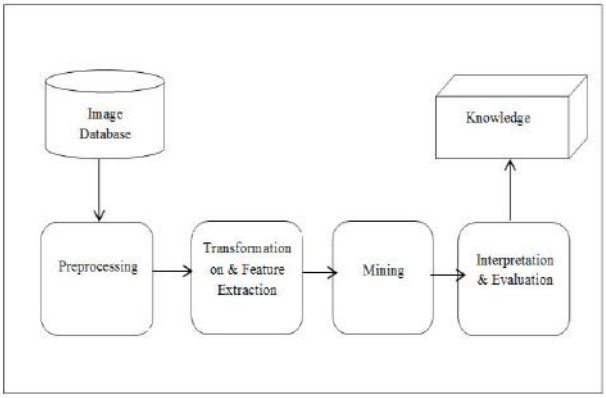

Figure 1. Image Mining Process

(1) Image Preprocessing: According to the users and application requirements, the preprocessing of image data can be done at region level and edge level. Proximate features retrieval and extrication should be performed either automatically or manually so that the image can be best captured. Preprocessing of images is performed to improve quality.

(2) Image Transformation and Feature Extraction: Preprocessed images transformed in varying ways and features are extricated to produce the useful and desired features from image content. Shapes, color, texture the traits that are generally used for mining to extract features.

Image mining process comprised of following steps as shown in figure 1: pre-processing, transformations, feature extraction, mining significant patterns and features, evaluation, interpretation and finally extracting knowledge.

(3) Image Mining: The extricated features than used for mining using various mining techniques to discover significant patterns. Image mining is not just to apply currently existing data mining techniques or algorithms; it is much more than that. There difference between image and relation database and they are-
1) In image database, the data values are not relatively significant while in relational database, the data values are significant.
2) In image database, it is critical to interpret implicit spatial information as the position is dependent and interrelated while in relational database there is no difficulty in such interpretation because position is independent.

(4) Interpretation, Evaluation and Knowledge discovery from images: After mining, patterns are obtained and these patterns finally evaluate and interpret the knowledge that is required. The knowledge retrieved can be used by individual or organization for various purposes to make predictions and profitable output. [1]

2. LITERATURE REVIEW:

In this section, we discuss recent literature on some key aspects of image mining and image retrieval. Image mining deals with the extraction of image patterns from a large collection of images. Clearly,

image mining is different from low-level computer vision and image processing techniques because the focus of image mining is in extraction of patterns from *large* collection of images, whereas the focus of computer vision and image processing techniques is in understanding and/or extracting specific features from a *single* image. While there seems to be some overlaps between image mining and content-based retrieval (both are dealing with large collection of images), image mining goes beyond the problem of retrieving relevant images. In image mining, the goal is the discovery of image patterns that are significant in a given collection of images [4].

Image mining is an interdisciplinary endeavour that draws upon expertise in various fields like computer vision, image retrieval, matching and pattern recognition. Some methods allow image mining to have two different approaches. First method extracts images from image databases or collection of images. Second method mines a combination of associated alphanumeric data and collection of images. Research in Image mining can be broadly classified in two main directions (1) Domain specific applications (2) General applications. Both are used to extract most relevant image feature and later to generate image patterns. A vast amount of image data is generated in daily life and in various fields like medical, astronomy, sports and all kinds of photographic images. It is still at the experimental stage and growing field of research. Lack of understanding in the research issues of image mining is the obstacle to rapid progress [5]

Stanchev [6], using image mining in image retrieval, described a new method for image retrieval using high level semantic features. It is based on extraction of low level color, shape and texture characteristics and their conversion into high level semantic features using fuzzy production rules, derived with the help of an image mining technique. Dempster-Shafer theory of evidence is applied to obtain a list of structures containing information for the image high level semantic features. Johannes Itten theory is adopted for acquiring high level color features. The main advantage of this method is the possibility of retrieval using high level image semantic features. After the full system realization it will be able to obtain statistic characteristics about the usefulness of the suggested method.

S.Balan and T.Devi[6] explains that the retrieval process represents a visual query to the system and extracts the images based on the user request such mechanism referred to as query-by-example and used to compare some similarity metrics to compare query and target images. The greater demand for retrieval and management tools for visual data and visual information is a more capable medium of conveying ideas and is more closely related to human percept ion of the real world. In Text based Image Retrieval images are indexed and retrieved based on the descriptions such as their size, type, date and time of capture, identify of owner, keywords or some text description of the image. This is often called description based or text based image retrieval process .In Query Based Image Retrieval query image can be extracting the visual features and can be compared to find matches with the indices of the images stored in the database. These features are used to retrieve the similar images from the image database.

3. IMAGE RETRIVAL

In general, image retrieval is classified into two methods such as: Text Based Image Retrieval (TBIR) and Content Based Image Retrieval (CBIR).

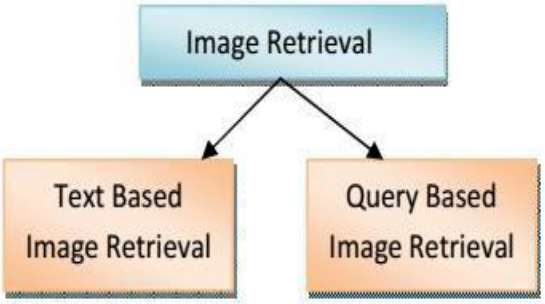

Figure 2. General form of Image Retrieval

The advantages of TBIR are easy to implement and fast retrieval. The disadvantages of TBIR are manual annotation is impossible for a large database, time consuming process, not accurate, loss of information and more than one object can be refer by the same word, inefficiency.

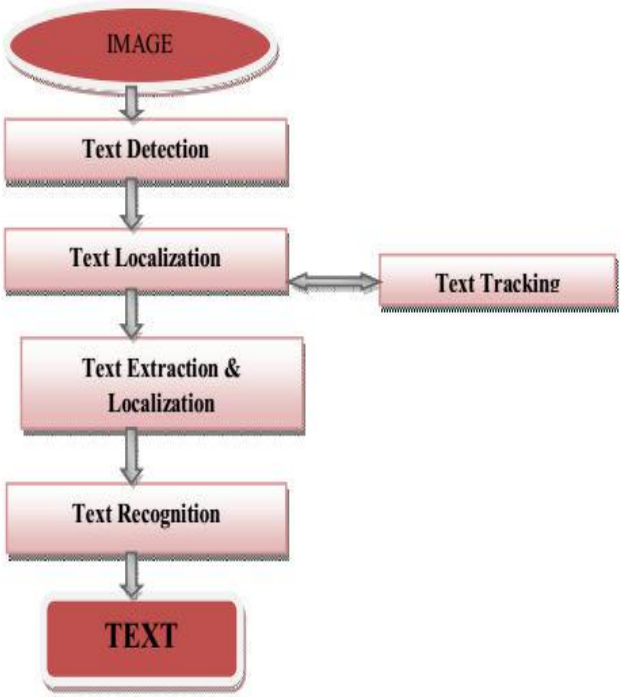

Figure 3.Text based Image Retrieval

In CBIR systems, a feature is a characteristic that can capture a certain visual property of an image .In

feature extraction, features such as color, texture or shape from image are extracted and creates a feature vector for each image. Content based image retrieval (CBIR) technique indexes images based on their content like color, shape, texture etc., image retrieval in CBIR covers image processing and information retrieval mechanism. The CBIR system performs 2 main tasks:

a) *Feature Extraction*: where feature vector (set of features) are generated to accurately represent the content of each image in the database.

b) *Similarity Measurement*: where distance between the query image and other images in the stored database based on feature vectors is used to retrieve the "closest" images.[1]

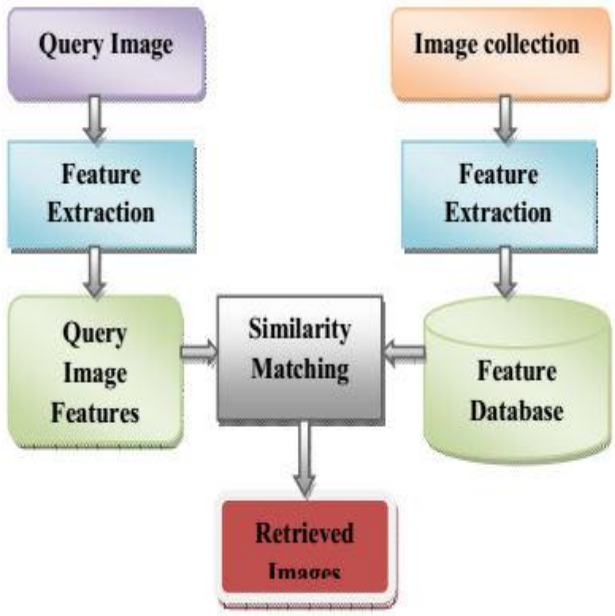

Figure 4.Content based Image Retrieval

4. DIFFERENT METHODS OF COLOR AND TEXTURE FEATURES

In this section is shown in different methods of color and texture features of image retrieval in image mining applications.

A. Color Feature

The color distribution of the pixels in an image contains huge amount of information about the image content. The image attribute can be obtained from the image color distribution by employing the color-occurrence matrix color space, histogram, moment, and color coherence vector and color structure descriptor. Color space consists of three dimensional spaces and color is used as a vector in it[2]. Color spaces are required for description of color based retrieval of image. Mostly RGB, HSV, HSI, YCrCb, LAB. Most color images are recorded in the well-known RGB color space.[2].

Table 1. Color Feature

Color Feature	Pros	Cons
Conventional color histogram	Simple, fast computation, invariant to any transformation	No spatial information
Color moment	Low complexity	Precision low
Fuzzy color histogram	Fast computation. Robust to quantization noise.	More computation
Color Correlogram	Encode spatial info	Very slow computation
Color Autocorrelogram	Encode spatial info, low complexity	Precision low
Color Structure Descriptor	Both color distribution and local structure of the color, Hmmd color space	No spatial information
Scalable Color Descriptor	Simple, fast, Hsv color space	No spatial information
Color Cooccurrence Feature	Include spatial info, low complexity, high speed, reduce the feature dimensionality	Codebook needed
Color Coherence Vector	Include spatial information	Complexity high

B. Texture Feature

Texture refers to the visual patterns that have properties of uniformity that do not result from the presence of only one color or intensity. It is Fourier transformed boundary as the shape feature. The Moment invariants are to use region-based moments, which are invariant to transformations as the shape feature [3].

Textual image indexing is a very efficient and easy technique for image retrieval. Basically, the approach emphasizes on keywords given for a specific image. Text data present in images contain useful information for, keyword additions, Standard titles of content, caption indexing, etc. There are differences in text style, text size, text orientation and alignment and all these can be used for image indexing.

Table 2: Texture Feature

Texture Feature	Pros	Cons
Tamura feature	Effective retrieval	Highly complex
Wavelet filter	Detect different frequency and orientation	Precision low
Gray level co-occurrence matrix	Including position of pixels having similar gray level values.	High dimensionality
Gabor filter	High retrieval	Computation expensive
Steerable pyramid	Support any no of orientation	Storage space is high
Edge histogram descriptor	Computation easy	Retrieval result poor
Gabor moment	Low dimensionality	Computation expensive
Bit Pattern feature	Low complexity, high precision, capture edge visual pattern, texture information	Codebook needed
LBP, LTP, LTRP	Best feature for texture retrieval	Only suitable for gray scale image

5. CONCLUSIONS:

This study of image mining implies on challenges and accountability of various prospects. The implementation of CBIR system raises several research challenges, such us how to retrieve images efficiently, size of the image database and time complexity for feature extraction. The successfulness of the image retrieval system is measured with the precision, recall, and average retrieval rate. These values indicate the percentage of relevant image returned by a CBIR system with a specific number of retrieved images. Many researchers have done lot of research in described applications but there is lack of efficiency that is required at current stage. The further study of image retrieval finds much application such as like medical diagnosis, crime prevention, fingerprint, face recognition, ear recognition, objection recognition and iris recognition.

ACKNOWLEDGMENT:

I would like to thank all of my current and previous colleagues at Computer University for their support and encouragement during my research work. A big thanks you to my husband Oo Tun Shwe and daughter SanThitSarNyi for their help and support.

REFERENCES:

1. Ankita Tripathi, Hitesh Jangir, "A Study on Image Mining Methods and Techniques", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 4, April 2016.
2. Dr.Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng. "Fundamentals of content-based image retrieval", In Multimedia Information Retrieval and Management, D. Feng Eds, Springer, 2003.
3. Felci Rajam, S Valli, A Survey on content based image retrieval,Life Science Journal, 2013.
4. *Dr.*Hilal M. Yousif, *Dr.*Abdul- Rahman Al-Hussaini and Mohammad A. Al-Hamami "Using Image Mining to Discover Association Rules between Image Objects".
5. T.Y.Gajjar,N.C.Chauhan ,"A review on image mining frameworks and techniques" International journal of computer science and information technologies, vol3(3),2012.
6. Peter-Stanchev,"Using image mining in image retrieval",IASTEDconference,"Computer science and technology"(2003).
7. S.Balan and T.Devi,"Design and Development of an Algorithm for Image Clustering In Textile Image Retrieval Using Color Descriptors", International Journal of Computer Science, Engineeringand Applications (IJCSEA) Vol.2, No.3, June 2012.

Author's Biography:

I received the B.C.Sc and B.C.Sc (Hons:) degree from Government Computer College, Lashio, Northern Shan State, Myanmar in 2003, the Master of Computer Science (M.C.Sc) degree from University of Computer Studies, Mandalay, Myanmar in 2006 and the Ph.D(IT) degree from University of Computer Studies, Yangon, Myanmar in 2014 January. My research interests mainly include face aging modelling, face recognition, perception of human faces and information technology (IT).