A Review on Content Based Image Retrieval Using Feature Extraction

Devbrat Arya, Jaimala Jha (Asst. Prof)
Department.of CSE & MITS College, GwaliorDepartment. of CSE/IT & MITS College,
Gwalior, Madhra Pradesh, India

Abstract— Content based image retrieval (CBIR) is prominent technique for discovery various images from a large amount of image database. So image search and retrieval from large image data set is difficult task. In past few year CBIR is gaining more attention of researcher. CBIR uses the image visual cotents for example color, shape and texture to index and represent image. So the research focus has shifted from algorithms of low-level feature extraction to high level visual feature extraction mechanism. For better CBIR system, necessary to improve various retrieval process such as Feature extraction, similarity matching method and soon.In this paper, detailed on CBIR with feature extraction method and compared with performance parameter their advantages and disadvantages.

Keywords—CBIR, Color, Texture, Shape, Feature Extraction.

I. INTRODUCTION

Image processing (IM) is the technique for enhancing the pictorial data for human understanding and changing the way of a image into additional appropriate way for autonmus machine observation [1]. The benefit of IM machines over people is that protection just about the whole electromagnetic range, reaching from gamma to waves of radio where as person eye is restricted to electromagnetic spectrum visual band.They can work on images produced by sources such as electron microscopy, ultrasound, and PC created images. In this way IM has a gigantic scope of uses and practically every territory of science and innovation, for example, medicine, space program, farming, industry and law implementation make utilization of space program, One of the key subjects with any sort of picture handling is picture recovery which is the need to separate useful data from the raw data, for example, perceiving the vicinity of particular color or surfaces before any sort of thinking about contents image's is conceivable.Image recovery has been an exceptionally dynamic research regionsince 1970s, with push of two noteworthy exploration groups: management of database and PC vision[2].

In this way, image retrieval can be characterized as the assignment of looking for images in aimage database. CBIR systems look collection of images [3] in view of features that can be separated from files of image themselves without distinct. The target of CBIR systems is to backing image retrieval in light of substance e.g., figure, color, texture. In past decades numerous CBIR systems have been produced, the shared belief for them is to extricate a coveted image. Contrasting two images and choosing in the event that they are comparative or not is a moderately simple thing to accomplish for a human. Getting a PC to do likewise successfully is however an alternate matter. The algorithms utilized as a part of these methods are regularly isolated into three errands: extraction, classification and selection. The extraction mission is to changes over well off substance of images into different substance features. Extraction of feature is the methodology of creating elements to be utilized as a part of the selection and classification tasks. Feature selection declines the quantity of features gave to organization task.Those types which are prone to help in inclination are chosen and utilized as a part of the classification task. Features which are not chose are discarded. Of these three actions, extraction of feature is most basic on the critical that the specific features made accessible for separation specifically impact the adequacy of classification task.
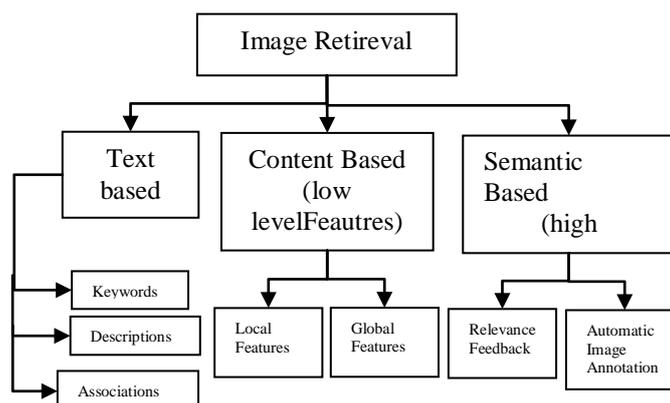

Figure 1.Image retrieval categories.

The final consequence of the extraction assignment is an arrangement of components, regularly called feature vector, which constitutes image representation. In the most recent couple of decades, various above uncovered systems of picture substance highlight extraction affirmed sufficiently solid for practical uses in industrial automation, social security, biomedicine, crime prevention and biometric authentication.

II. FEATURE EXTRACTION

In extensive sense, features may visual based and text based. Textual features are annotations, tags, keywordsetc.features of the visual image aretexture, shape, coloretc. The visual features are further categorized as common features and domain particular features. Common features are shape, texture, colorand domain particular features are application dependent for e.g. finger prints and human faces. Domain particular features are related to recognition of pattern.A characteristic is defined as an fascinating a part of an image and elements are used as a beginning point for a lot of computer imaginative and prescient algorithms. Considering the fact that elements are used as the beginning factor and common primitives for subsequent algorithms, the total algorithm will typically handiest be as good as its characteristic detector.

A. Color:

Color defines one of the features of significant visual in CBIR. There are various examples, where features of color in image retrieving are used for exampleblock-based,histograms,moments. Color histogram is used formeasures computingdistance based on the similarity of color for all image. A color histogram is used to define the distribution of global color in an image and is extra frequently used technique because of its benefits like high efficiency. Other representation of feature for example color sets and color moments are also used than histogram of color. The color is aextensively used significant feature forrepresentation of image. This is mostsignificant as it is invariant with respect to image scaling, translation and rotation [4]. Color space, color quantification and similarity measurement are the color feature extractionkey components. Color feature is not dependent upon image size. The models of color canbe categorized as models of User & Hardware based; for example HSV and RGB. Numerousspaces of color are there which offers various applications.

B. Texture:

Texture includesignificantknowledge about structural surfaces arrangement and their relationship to the surrounding atmosphere. It is an inherent property of effectivelyeachsurfacesconcludinghair, trees, bricks, clouds, and fabric. Texture givesvaluable surfaces data about their relationship and structures with surrounding. Analysis and texture can be studied at three different levels i.e. on statistical level, statistics extracted set from the image is known as texture. On the level of structural, the image primitives and theirrules of placement are well-known as its texture. On the level of spectral, texture is describes as coefficients set in transform domain. With various level use the textures can be recognized but textures may not agree with human evaluating the texturesway.[5] These reasons are human perception and semantic gap subjectivity. Texture feature defines spectral features which are taken applyingstatistical features, wavelet transform, tamura texture features etc. Tamura explored the representation of texture from a various viewpoint. [6] colorand texture queries can be formulated in similar way, through selecting preferred textures or throughproviding an query image.

C. Shape:

fShape does not refer to the image shape but to the specific region shape that is being sought out. In the image retrieval, and on applications depending, fewneed the representation of shape to be invariant to translation, scaling and rotation, while others do not. Objects or regions shape features have been used in numeroussystems of CBIR. Compared with the featuresoftexture and color, shape features are generallydefined after images have been segmented into objects or regions. Features of shape are separated into two different classes region based and boundary based. Features of boundary based shape uses only shape boundary whereas shape features of region-based uses complete shape region[5]. The term shape refers to data that can be assumed directly from image. Shape is represented through means of perceptually grouped geometric cubes,vertices to edges, joints, contours, and polygonal areas removed from an image. One of these grouping can function a spatial design or as a rough sketch throughappliedadditional post processing featuresof shape are called as geometric features. shape feature are usually used – global features for example aspect ratio, circularity and moment invariants and local features for example consecutive boundary sets.[7]

Table I Comparison of the various techniques [9]

Feature	Techniques	Accuracy	Dimension	Advantage	Disadvantage
Color Feature	Color Moment	Low	Low	Lower computational complexity	Precision is low
	HSV histogram	High	Medium	Simple ,Fast computation	No spatial information
	Color Correlogram	High	High	Contains spatial correlation of colors ,Simple to compute	Very slow computation
Texture Feature	Gabor Filter	High	High	Achieves highest retrieval results	Computationally intensive
	Gabor Moment	Low	Low	Lower Dimensionality	Low retrieval result compare to Gabor filter

	Gray level co-occurrence matrix	High	High	Contain pixels positions having similar values of gray level	High Dimensionality
Shape Feature	Moment Invariant	High	Low	Invariable to translation, rotation and scale	Limited recognition power
	Zernike moments	High	Low	Invariable to translation, rotation and scale	Computational Complexity is High

III. ARCHITECTURE OF CBIR SYSTEM

Content-based retrieval utilizes the substance of images to speak to and access the images from the large database.
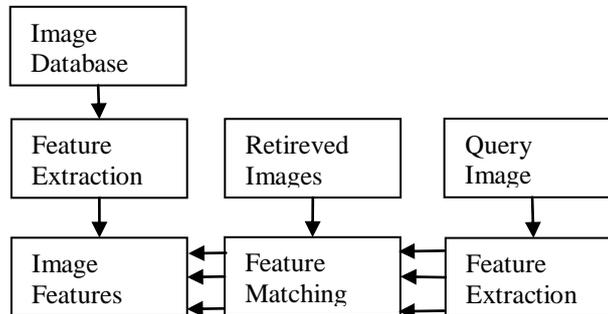

Figure 1.1. General Architecture of CBIR

The general architecture of CBIR system is for the given image database, elements are extricated first from individual images. The features can be visual elements like color, texture, shape, locale or spatial elements or some compacted domain features. The separated elements are depicted by highlight vectors. These element vectors are then put away to shape image feature database. For a given query image, we comparably remove its elements and structure a feature vector. This feature vector is coordinated with the already stored vectors in image feature database.

IV. APPLICATIONS OF RETRIEVAL SYSTEM

The benefits of such structures range from simple users searching a specific image on web.
Numerous professionals kinds like police force for recognition of image in crime prevention.

- Medicine diagnosis
- Engineering and architectural design
- Publishing and Fashion
- Remote sensing systems and Geographical information
- Home entertainment etc.

V. METHODS COMPARISON

Paper name	Feature extraction method	Performance evaluation parameter	Advantages	Disadvantages
Color & Texture Feature of CBIR [17]	Color Histogm Standrd Wavelet	Retrieval Accuracy	Improve the retrieval accuracy	Insufficient Feature Set
CBIR using colour, textur& Shape feature [18]	Colour moment, Gabor filter, GVF	Retrieval Accuracy	Create robust feature set	High Semantic gap
CBIR using Feature Combination & RF[19]	Colour moment, Gabor wavelet co-occurence matrix filter,	Precision	Minimize the semantic gap applying RF with SVM	It is the time consuming to label negative examples

VI. LITERATURE REVIEW

Choras et al.,[12] proposed an integrated colour, shape and texture characteristic removal system where Gabor filtration is used for opting for the areas of interest quantity (ROIs). They calculated elements of colour and texture from the ROIs based on threshold Gabor facets and histograms, color moments in YUV house, and form facets headquartered on the Zernike moments. The features provided proved to be efficient in making a choice on the similarity between image.

Jain and Singh [13] provided content material functionality overview centered image retrieval programs through combining benefits of HC and divide and overcome k-Menas approach. He proposed HDK procedure to use both HC and Divide benefits and overcome k-means with the aid of introducing equivalency and suitable relation standards.

Singha and Hemachandran [14] exhibited the CBIRutilizing features like surface and color, called Wavelet Based Color Histogram Image Retrieval (WBCHIR). The surface and color elements are extricated through wavelet change and color histogram and blend of these features is strong to scaling and objects interpretation in animage. He likewise showed a promising and quicker recovery technique on a WANG picture database having 1000 common-purpose color images.

Pinjarkar et al.,[15] talked about different systems utilized as a part of the exploration range of CBIR methods utilizing Relevance Feedback. To enhance the retrieval execution

of the CBIR the Relevance feedback manner will also be included in CBIR approach to acquire the greater values of the general estimation parameters used for CBIR evaluation approach which can lead to higher outcome of retrieval efficiency. He additionally mentioned various relevance suggestions procedures for content established image Retrieval programs, the more than a few parameters used for experimental systems estimate and the these methods analysis on the foundation of their results.

NitishBarya et.al [3] In this review paper, we focused on a survey of different methods of CBIR to improve the performance of image retrieval system. Also, this paper gives a review of the works carried out, in eliminating the semantic gap between the low level features and the concepts of high level semantic. Fuzzy color histogram is one of the appropriate method of image retrieval over conventional color histogram method and further SVM- RBF (Radial basis function) can be used to reduce the semantic gap and improve the retrieval performance. But due to creation of vast amount of digital images, users are not satisfied with the current image retrieval techniques. Hence further improvement is required to increase the image retrieval quality. A key aspect of this survey is to suggest a system that, in addition to use of existing low level features combine with new feature to match images that may give better result.

Rajinder Kauret.al [16] In early era, to search image text based image retrieval techniques were used but those were not applicable for large scale databases because of text description problems. In present methodology acquisition, transmission, storing, and manipulation are permitted on bigset of images. With the expansion in prominence of the network and advancement of multimedia technologies, users are not fulfilled by the customary data retrieval procedures. So these days, the substance based image retrieval is turning into a wellspring of definite and fast retrieval. CBIR is a system which utilizes visual elements of image, for example, color, shape, texture and so on to hunt user required images from expansive image database as indicated by users solicitations as an query image. The feature extraction technique has been applied for image segmentation using HSV and RGB. In this paper first we review the building of CBIR system based on color features in RGB and HSV color space and compare both method.

This paper elucidates extraordinary proposed ways and strategies utilized through researchers. It propagates the importance of content situated image retrieval methods. The perfect goal of CBIR methods is to extract the points like form, colourand image texture from database and examine it with the question image with a view to retrieve the wanted image. Additional approaches and tactics are in progress to make content centered image retrieval process extra powerful and efficient. This overview specializes in content material founded implementations of image retrieval, usability and challenges. It additionally supplies conceptual overview of methodology.

VII. PERFORMANCE MEASUREMENT

A. Euclidean Distance (ED)
It is defined as the displacement of a pixel from the nearest background point. The equation of ED is[14].

$$D = \sum (A_i - B_i)^2$$

B. Chi Square Distance
The ED between the components of profiles, on which weighting is defined (weight means the inverse of its frequency), is called the chi-square distance. The equation is as: [14].

$$\chi^2 th = \sqrt{\sum_{j=1}^{p} \frac{1}{a+j} \left[\frac{a_{hj}}{ah+} - \frac{a_{ij}}{ai+} \right]^2}$$

C. Weighted Euclidean Distance (W.E.D):
Multiply squared differences by corresponding weights are known as Weighted Euclidean Distance. The formula of W.E.D is[14]

$$dx, y = \sqrt{\sum_{j=1}^{i} \frac{1}{S_j^2} \left(X_j - y_j \right)^2}$$

D. Performance Parameter
Evaluation of retrieval process is a crucial problem in CBIR. Different methods are used for measuring performance of retrieval system. The most common performance parameters are Precision and Recall .

E. Precision
Precision rate is defined as a ratio of various retrieve relevant images similar to the query to the total number of retrieved images in response to query[1][10[11].

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}$$

F. Recall

Recall rate is defined as a ratio of number of retrieve relevant images similar to the query to the total number of relevant images presented in the database[1][10][11].

$$\text{Recall} = \frac{\text{The number of relevant images retrieved}}{\text{Total number of images in database}}$$

VIII. CONCLUSIONS

In this paper, we have compared variousfeature extraction methods of texture, colorand shape that are most commonly used in image understanding studies. Our comparison present that there is considerable performance variability between numerous feature extraction approaches. One of the features of this study is the use of a publicly presented benchmark that further studies can use.

REFERENCES

[1]   Kenneth R. Castleman, (1996)*"Digital Image Processing"* . Prentice Hall.

[2]   A. Blaser, (1979) *"Database Techniques for Pictorial Applications"*, Lecture Notes in Computer Science,Vol.81, Springer Verlag GmbH.

[3]   NitishBarya, HimanshuJaiswal*"Survey on Content based Image Retrieval to Deal with Rapid Growth of Digital Images"*International Journal of Computer Applications (0975 – 8887) Volume 124 – No.12, August 2015.

[4]   Henning Müller, Nicolas Michoux, David Bandon, Antoine Geissbuhler (2004), ―*A review of content- based image retrieval systems in medical applications*――clinical benefits and future directions‖, International Journal of Medical Informatics 73, 1―23.

[5]   Chiu, C.Y., Yang, S. and Lin, H.C.(2002), ―*Learning human perceptual concepts in a fuzzy content based image retrieval system*‖, 5th International Conference on Computational intelligence and Multimedia Applications, pp. 330-335.

[6]   Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying Ma.(2007),‖ *A survey of content-based image retrieval with high-level semantics*‖ The Journal of the pattern recognition society. Pattern Recognition 40 (2007) 262 – 282

[7]   Jayant Mishra, Anubhav Sharma and KapilChaturvedi(2011),‖ *An Unsupervised Cluster-based Image Retrieval Algorithm using Relevance Feedback*‖ International Journal of Managing Information Technology (IJMIT) Vol.3, No.2, May 2011

[8]   Dr. Anna SaroVijendran and S. Vinod Kumar *"A New Content Based Image Retrieval System by HOG of Wavelet Sub Bands"*International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 8, No. 4 (2015), pp. 297-306.

[9]   Keyuri M. Zinzuvadia, Prof. Bhavesh A. Tanawala, Prof. KeyurN.Brahmbhatt*, "A Survey on Feature Based Image Retrieval Using Classification and Relevance Feedback Techniques",* International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 1, January 2015, pp: 508-513.

[10]   Mohd. Danish,RitikaRawat ,Ratika Sharma *"A Survey: Content Based Image Retrieval Based On Color,Texture,Shape&Neuro Fuzzy"* Journal Of Engineering Research And Applications ISSN : 2248-9622, Vol. 3, Issue 5,Sep-Oct2013,Pp. 839-844.

[11]   Ms. K. Arthi and Mr. J. Vijayaraghavan, *"Content Based Image Retrieval Algorithm Using Colour Models"*, International Journal of Advanced Research in Computer and Communication Engineering ((IJARCCE), Vol. 2, Issue 3, March 2013.

[12]   R.S. Choras, T. Andrysiak, M. Chorus, *"Integrated color, texture and shape information for content-based image retrieval"*, Pattern Anal Applic. 10: 333-343, 2007.

[13]   Jain, M. and S.K. Singh, *"A survey on: Content based image retrieval systems using clustering techniques for large data sets",* Int. J. Manage.Inform. Technol., 3: 23-39, 2011.

[14]   Singha, M. and K. Hemachandran., *"Content based image retrieval using color and texture".* Signal Image Process. Int. J., 3: 39-57, 2012.

[15]   Pinjarkar, L., M. Sharma and K. Mehta*., "Comparison and analysis of content based image retrieval systems based on relevance feedback",*J.Emerg. Trends Comp. Inform. Sci, 2012.

[16]   Rajinder Kaur1, PardeepKadyan *"Survey On Content Based Image Retrieval Using HSV and RGB"* International Journal of Computer Engineering And Electronics Technology Vol. 1, Issue 1, 2015.

[17]   LeninaBirgale, ManeshKokare and DharmpalDoye, *"Colour and Texture Features for Content BasedImage Retrieval"*, Proceedings of the International Conference on Computer Graphics, Imaging andVisualisation, 2006.

[18]   P. S. Hiremath , JagadeeshPujari, *"Content Based Image Retrieval using Colour, Texture and Shapefeatures"*, 15th International Conference on Advanced Computing and Communications,2007.

[19]   Lijun Zhao, Jiakui Tang, *"Content-Based Image Retrieval Using Optimal Feature Combination andRelevance Feedback"*, International   Conference on Computer Application and SystemModeling,volume-4, v4-436 – v4-442, 2010.