IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-6, Special Issue-2, March 2018 E-ISSN: 2347-2693

Feature Extraction Techniques in Image Mining System –A Survey

D. Angayarkanni[1*], L. Jayasimman[2]

[1] Dept. of Computer Science, Srimad Andavan Arts & Science College (Auto), Tiruchirappalli,TamilNadu.
[2] Dept. of Computer Science, Srimad Andavan Arts & Science College (Auto), Tiruchirappalli,TamilNadu.

Available online at: www.ijcseonline.org

Abstract— Nowadays huge amount of images are produced with the rapid development of digital imaging technology in various fields such as medical, astronomy, weather forecasting, photography, satellite imaging etc. So maintaining images in the large databases, extracting useful information from the images and retrieval similar images are the emerging research area in current scenario. A point of interest in an image is called feature that transform pictorial information into alpha numeric data. That feature can be used for solving many problems such as reducing the dimension of the image, classifying the images, indexing the images in the image database, automatic data analysis and retrieval of images from the database, etc. One of the main tasks in imaging technology is to extract useful and important features from the images. This paper presents a study on various low level feature extraction techniques used in image mining system that is used for various applications of imaging technology.

Key words— Image Mining, Feature extraction, color, Texture, Shape, Image retrieval

I. INTRODUCTION

Image mining is the process of extracting implicit knowledge, image data relationship, or other patterns that are not explicitly stored in the images, between images and other alphanumeric data [1].In image mining, various techniques such as Object Recognition, Image retrieval, Image indexing, Image Clustering and Classification are used to mine knowledge from the image database[2]. In image mining process, initially images in the database are preprocessed to improve the quality of images and then the images are go through various transformation and feature extraction technique to retrieve important features from the images. Using these features, various data mining techniques such as clustering, classification are carried out to discover knowledge from the images[3].Features are the information that can be extracted from images in terms of numerical values that are difficult to understand and correlate by human. Feature extraction is an essential step for all images mining techniques and noticing as much information as possible for large image database. Most common image visual features are color, texture, shape etc. that can be used for many image retrieval and image processing tasks such as clustering, classification, indexing etc. Generally, features in the images are classified as low level features and high level features. Low-level features are directly operated on the image pixels. It is used to capture the visual property an image, either globally locally. The most commonly used features are color, texture, shape, and salient/interest points in an image. In global extraction, features are computed to capture the overall characteristics of an image. In local feature extraction, features are extracted from a small group of pixels [4]. High level feature extractions are based low level features and uses Artificial Neural Network (ANN) to extract features from multiple layers. The extracted features are stored in an *n*-dimensional vector that contains a set of values where each value represents a certain feature. This vector can be used to classify an object, or provide us with condensed higher-level information regarding the image.Rest of the paper is organized as follows, Section I contains the introduction of Image Mining and features present in the image , Section II contain the information about the techniques used in the low level feature extraction, Section III contains the study of various literatures related to the feature extraction technique used in image mining system, Section IV concludes the research work

II. LOW LEVEL VISUAL FEATURES

A. Color

Color is one of the most commonly used low level visual features in image mining process. It is closely related to human visual perception. It is used to encode the spatial distribution of features in images and compact to provide efficient storage and retrieval. Basically, Color is defined in three dimensional spaces. Colors are represented by variety of color spaces which denotes the range of colors as tuples of numbers, typically as 3 or 4 values or color components (e.g. RGB).There are five major color models i.e,CIE, RGB, YUV, HSL/HSV, and CMYK[4]. To extract the color feature, Color histogram, colorcorrelogram, color coherence vector and colormoments are used with different color

© 2018, IJCSE All Rights Reserved

spaces. Histogram based Descriptors such as Scalable Color Descriptor, ColorStructure Descriptor and Dominant ColorDescriptor are also used to represent the color information [5] [6].

B. Texture

Image texture gives information about the spatial arrangement of color or intensities in an image or in a selected region of an image. Textures is a repeated pattern of information or arrangement of the structure with fixed intervals and have quantifying properties such as smoothness, coarseness and regularity. It's often used as a region descriptor in image analysis and computer vision [4]. Texture features of an image is extracted using statistical, structural, Model and transform methods.Statistical texture techniques describe texture of regions in an image through statistical properties of gray level of the points in the surface region and computed using gray level histogram or gray level co-occurrence matrix (GLCM). The GLCM approach use second-order statistics of the greyscale image histograms. Structural texture techniques describe a texture as the arrangement of well-defined texture elements in regularly spaced parallel lines with some predefined rules. Transform based texture techniques use spatial frequency properties of the pixel intensity variations of an image to convert the image into a new form. Model based texture analysis techniques generate a realistic model of each pixel in the image based on a weighted average of the pixel intensities in its neighbourhood. The texture descriptors used to observe the homogeneity of the region and the histograms of the corresponding region borders. The set of descriptors used to find texture patterns are Homogeneous Texture Descriptor, Texture Browsing Descriptor and Edge Histogram Descriptor [5] [6].

C. Shape

Shape feature of an image contains semantic information that can be easily acquired by many segmentation techniques. Shape of an image is extracted by external method (boundary/Contour based) and internal method (region based) that describes regions, contours and shapes for 2Dimages. Moment descriptors like geometric moments, Legendre moments, Zernike moments and pseudo Zernike moments are used as Region based Descriptor and they are used in more general situations and involves more computation. Contour-based methods need extraction of boundary information. Contour shape representations comprises global shape descriptors such as eccentricity and circularity, shape signatures such as chain code, centroid distance and cumulative angles, spectral descriptors such as Fourier descriptors and wavelet descriptors, and curvature scale space descriptors [5][6].

III. LITERATURE SURVEY

Various methods are proposed for the extraction of low level features from the image in the literature. The aim of feature extraction is to reduce the complex and high dimensional information present in the images. In image mining, features can be used for finding similarity between the images, classifying and clustering the images. From that, images are retrieved and detect knowledge from the images very easily and efficiently.

Sreelekshmi et al. [7] proposed a texture feature extraction technique which uses four statistic features contrast, homogeneity, energy and correlation of GLCM in all four directions of an image. This system reduces the searching time of query image and leads to an increase the speed of image retrieval.

Hire Dnyanda et al. [8] suggested the combination of HSV color space for extracting color features and sobel edge detection method for shape feature extraction to retrieve images from the database.

S.Asha et al. [9] presented an analysis on various edge detection techniques to extract edge features in content based image retrieval. This system proposed canny edge detector which gives enhanced result based on the evaluation of various edge detection techniques and in future, to detect edges in color image without gray image conversion and for automatic extraction of moving object a canny algorithm can be improved.

AnkitaTripathi et al. [10] proposed an image mining technique for classification of textual images using low-level image features including GLCM features mean, skewness, energy, contrast, homogeneity. The differences between images are measured and used to classify the textual images by performing classification and clustering techniques on datasets using the above features.

FatemehAlamdar et al. [11] proposed a feature color extraction using quadHistogram in HSV color space and encode the histogram using Haar Transform.

JeyanthiPrabhu et al. [12] presented a color histogram technique for image retrieval. To extract texture feature, this system presents gray level co-occurrence and color co-occurrence matrix and combine the color histogram method with gray level co-occurrence technique and color co-occurrence matrix respectively. This sytem also proposed that
the integrated method of color and texture is giving better results than the single color image retrieval and also the wavelet based image retrieval using gray level co-occurrence matrix produces better result than the color co-occurrence matrix.

Alphonsa Thomas et al.[13] presented a survey on the essential concepts of content based image retrieval systems and suggested that Combining of low level features such as color, texture and shape in retrieval has lots of advantages and the retrieval based on the combination gives

© 2018, IJCSE All Rights Reserved

are more exact and efficient results than other methods based on single feature.

SumanKhokhar et al.[14] used multiple features of an image such as color ,texture and shape to classify the images for content based image retrieval This system uses LAB Color space and Zernike Moments for shape ,GLCM for texture and RGB color histogram for color feature extraction and all the extracted features are applied to the feature selection algorithm ReliefF which selects top 20 features for BFNN classification and results show that the performance of the proposed technique overtakes the other conventional image retrieval methods by attaining higher precision values, recall values, and accuracy of 94.28%.

BhagyashriAmrutkar et al.[15] used Dominant color, Texture ,edge features to extract the image features and then clustering the images in the feature database to present an efficient content based image retrieval system.Dominantcolors extracted using color-quantization technique the Local Binary Pattern is applied for texture feature extraction shape feature extraction with improves sobel color edge detection technique is used. Then all features of image are combined to form a single feature vector and KMeans clustering is applied over combined feature vector of database images.

Kannan et al. [16] presented a new clustering technique that extract color, texture property from the query image and the images in the image database based on classification of the image like High-texture detailed Image, Average-texture detailed Image and Low-texture detailed Image. This technique is also compared with existing methods and performance is measured using precision and recall.

ParagDhonde et al. [17] proposed a hierarchical; KMeans based image retrieval method that gives the better accuracy performance with the combined approach of the feature extraction such as color, texture, and shape as compared to individual features.

Rahul Mehta et al. [18] proposed a combination of color, texture based image retrieval system that uses conventional color histogram (CCH) for color feature extraction and to measure the similarity between the query image and retrieved images from the image database Quadratic Distance Metric (QDM) is used. And the texture feature is retrieved using pyramid structure wavelet transform model and Euclidean distance used to measure the similarity between images .This combination gives better result than the traditional one.

ManimalaSingham et al.[19] proposed a wavelet based color histogram(WBCH) for image retrieval in which texture and color features are extracted through HAAR wavelet transformation and color histogram and the combination of these features is robust to scaling and translation of objects in an image. Similarities between images are measured with Histogram intersection distance method. This system is implemented with matlab7.3 and uses WANG database. This method gives better performance based on precision, recall as compared to existing methods.

J.C.Kavitha et al. [20] proposed a GLCM texture feature extraction combined with RGB Color Histogram that gives better classification accuracy for dermoscopy images into melanoma and non-melanoma as compared with the combination of GLCM with color histogram with other color spaces such as HSV, OPP. Support vector machine (SVM) is used for the process of image classification.

Shijin Kumar et al. [21] proposed a GLCM for texture feature extraction and Connected Regions for shape feature extraction, the extracted features can be used for the classification of images for malignant, benign and normal MRI and this method gives better accuracy rate.

Jigisha M. Patel et al. [22] presented an overview of color and texture feature extraction techniques like color histogram, colorcorrelogram, color co-occurrence matrix and tamura texture feature, steerable pyramid, wavelet transform, Gabor wavelet transform respectively and the comparative analysis also have been presented in the paper. This system suggested that Gabor wavelet transform is used when there is requirement for efficient discrimination of texture feature.

AmanChandha et al. [23] suggested the combination of feature extraction techniques such as Average RGB, Color Moments, Co-occurrence, Local Color Histogram, Global Color Histogram and Geometric Moment provides better results as compared with individual technique based on the comparative analysis of CBIR. Performance of the retrieval is improved by cropping of query image to identify the user's point of view in a specific image and resulting more precise and personalized search results.

G.Nagarajan et al. [24] suggested a hybrid genetic algorithm for feature extraction and selection in medical images. In the first phase of this system, Texton based contour gradient extraction algorithm, intrinsic pattern extraction algorithm and modified shift invariant feature transformation algorithms are used to extract the features of the image. In the second phase relevant feature vector is selected using a combined approach of "Branch and Bound Algorithm" and "Artificial Bee Colony Algorithm" using several types of medical images such as breast cancer, Brain tumour and thyroid images. Chi-Square distance function is used to measure the similarity between the query and database image. To increase the performance of retrieval system, a diverse density based relevance feedback method is used in the third phase. This system gives better performance than the previous algorithms.

Shijin Kumar P.S et al. [25] suggested a technique to retrieve information form MR images that extract texture using GLCM and shape features from the segmented image using Connected Regions. These feature vectors are used for classification of MR image. This system gives increased accuracy of large dataset. In future, artificial Neural Network is used to develop classification algorithm and categorize MR Images.

© 2018, IJCSE All Rights Reserved

SushantShrikantHiwale et al. [26] proposed Color Histogram, Color Auto-Correlogram, Color Moment, Gabor Wavelets and Discrete Wavelets transform to extract image features. Using Support Vector Machine (SVM) classifier which effectively distinguishes between relevant and irrelevant images based on the extracted image features. The results of this system depicts that proposed method has better precision and recall rate compared to other methods.

BalvatTarulatha et al. [27] presented indexing techniques to mine images in the database using the color features based on VIBGYOR colors by the highest color percentage. In this system, images are automatically classified by its color feature.Muhammad Imran et al. [28] proposed an image retrieval technique based on color histogram using HSV color space to extract color feature and First Order Statistics to extract texture features of the image. A performance result shows that the proposed technique achieved 15% higher accuracy as compared to Variance Segment and Histogram-based techniques. This technique can help for identification of suspects in forensic department.

Avinash N Bhute et al. [29] presented a survey of various methods and algorithms for the retrieval of images and proposed an indexing scheme that uses thecolor, texture, and shape features of the image. Based on the features, Antipole tree method is used to create indexing structure of the images in the database so that similar images are retrieved very efficiently from the database.

Naushad Varish et al. [30] considered color feature in RGB colorspace. The three probability histograms of each color component are divided into several numbers of significant bins respectively and from each bin, this system computed several statistical values standard deviation, skewness and kurtosis. The computed statistical values are used as extracted features of the image data. The processing costs of the presented CBIR techniques is significantly low and has been tested on standard image databases and have been achieved good result.

Suresh et al. [31] proposed a texture Structure Histogram that integrates color and texture information .In which, hsvcolor space for color information and edge orientation for texture pattern are used with non-equal interval quantization scheme and this system gives better performance.

JuliRejito et al. [32] proposed a color feature extraction using color histogram in RGB color space and similarity of images in the database is measured using Histogram intersection method. And extracted features are clustered by K-means algorithm and stored as an index for images in the database to produce featured database. This system is used to retrieve images with high rate of accuracy.

IV.CONCLUSION

This paper present a survey on various low level feature extraction method can be used in image mining system. Features from the images are extracted and converted to feature vectors. Those vectors are used as an index for images in the large database or used as a key for image classification and clustering. Feature extraction is also used for dimensionality reduction of images in the database. Selection of relevant features vector and the combination of feature extraction methods gives more accurate results for retrieval of images in the database and mining knowledge form the images.

REFERENCES

[1] AboliW.Hole, Prabhakar L.Ramteke, *"Design and Implementation of Content Based Image Retrieval Using Data Mining and Image Processing Techniques",* International Journal of Advance Research in Computer Science and Management Studies, March 2015.

[2] AnkitaTripathi, Hitesh Jangir *"A Study on Image Mining Methods and Techniques ",* International Journal of Innovative Research in Computer and Communication Engineering April 2016.

[3] K.Saraswathi, Dr.V.GaneshBabu,"*A Review on Image Mining Techniques",*International Journal of Scientific & Engineering Research, May-2017.

[4] https://en.wikipedia.org.

[5] 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, January 2015.

[6] B. S. Manjunathm,Jens-Rainer Ohm, Vinod V. Vasudevan,Akio Yamada ," *Color and Texture Descriptors"* IEEE Transactions on Circuits and Systems for Video Technology, June 2001

[7] Sreelekshmi. Anil A.R *"A Survey on Feature Extraction Techniques for Image Retrieval using Data Mining & Image Processing Techniques " ,* International Journal Of Engineering And Computer Science Nov. 2016.

[8] Hire Dnyanda N & S. S. Katariya " *Color And Shape Based Image Retrieval* ",International Journal Of Computer Science Engineering And Information Technology Research, Dec 2012

[9] S.Asha, R.Rajeshkanna *"A Survey on Content Based Image Retrieval Based on Edge Detection",* (IJCSIT) International Journal of Computer Science and Information Technologies, vol. 5 (6), 2014.

[10] AnkitaTripathi,Shivam Pandey, Hitesh Jangir, *"An improved and efficient image mining technique for classification of textual images using low-level image features",* IEEE International Conference on Inventive Computation Technologies, 2016.

[11] FatemehAlamdar, Mohammad Reza Keyvanpour "*A New Color Feature Extraction Method Based on Quad Histogram" ,* 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011)

[12] JeyanthiPrabhu and JawaharSenthil Ku ,"*Wavelet Based Content Based Image Retrieval Using Color And Texture Feature Extraction By Gray Level Concurrence Matrix And Color Concurrence Matrix* ", Journal of Computer Science Volume10(1),2014 Science Publications

[13] Alphonsa Thomas, SreekumarK,*"A Survey on Image Feature Descriptors-Color, Shape and Texture, International Journal of Computer Science and Information Technologies, 2014.*

[14] *SumanKhokhar , SatyaVerma , "Content Based Image Retrieval with Multi-Feature Classification by Back-propagation Neural*

© 2018, IJCSE All Rights Reserved

Network", International Journal of Computer Applications Technology and Research, 2017.

[15] *BhagyashriAmrutkar, Lokesh Singh "Efficient Content Based Image Retrieval Using Combination Of Dominant-Color, Shape And Texture Features And K-Means Clustering" International Journal Of Engineering And Computer Science ,January 2016*

[16] *S.KannanDr.V.MohanDr.N.Anbazhagan "Image Clustering and Retrieval using Image Mining Techniques" IEEE International Conference on Computational Intelligence and Computing Research 2010*

[17] *ParagDhonde, Prof. C. M. Raut "Precise & Proficient Image Mining Using Hierarchical KMeans Algorithm", International Journal of and Research Publications, January 2015.*

[18] *Rahul Mehta, Nishchol Mishra, Sanjeev Sharma, "Color - Texture based Image Retrieval System", International Journal of Computer Applications, June 2011.*

[19] *ManimalaSingham and K.Hemachandran "Content Based Image Retrieval using Color and Texture", An International Journal of Signal & Image Processing, February 2012.*

[20] *J. C. Kavitha; A Suruliandi "Texture and color feature extraction for classification of melanoma using SVM",* International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE), 2016

[21] Shijin Kumar P.S, Dharun V.S *"Extraction of Texture Features using GLCM and Shape Features using Connected Regions",* International Journal of Engineering and Technology (IJET) 2016

[22] Jigisha M. Patel, Nikunj C. Gamit *"A Review on Feature Extraction Techniques in Content Based Image Retrieval"*, IEEE WiSPNET 2016 conference

[23] Aman Chadha ,SushmitMallik ,RavdeepJohar *"Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval",*International Journal of Computer Applications, August 2012.

[24] G.Nagarajan, R.I.Minu, B Muthukumar , V.Vedanarayanan &S.D.Sundarsingh *"Hybrid Genetic Algorithm for Medical Image Feature Extraction and selection",* International Conference on Computational Modelling and Security (CMS 2016) Published by Elsevier

[25] Shijin Kumar P.S, Dharun V.S *" Extraction of Texture Features using GLCM and Shape Features using Connected Regions"* International Journal of Engineering and Technology (IJET),2017

[26] Sushant Shrikant Hiwale, *"Quick Interactive Image Search in Huge Databases Using content-Based Image Retrieval"*, IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems(ICIIECS), 2015.

[27] BalvantTarulatha,Ms. NamrataShroff,Prof. M B Chaudhary *"VIBGYOR Indexing Technique For Image Mining"*, IEEE International Conference on , Data Mining and Advanced Computing (SAPIENCE),2016.

[28] Muhammad Imran,Rathiah Hashim,Noor Eliza Abd Khalid ,"*Color Histogram and First Order Statistics for Content Based Image Retrieval",* International conference on Recent Advances on Soft Computing and Data Mining 2014.

[29] Avinash N Bhute, B. B. Meshram *"Content Based Image Indexing and Retrieval"* International Journal of Graphics & Image Processing ,Nov. 2013

[30] NaushadVarish , Arup Kumar Pal. *"Content based image retrieval using statistical features of color histogram",* IEEE, 3rd International Conference ON Signal Processing, Communication and Networking (ICSCN), 2015.

[31] Prof. Suresh M B, Dr. B MohankumarNaik *"Content Based Image Retrieval Using Texture Structure Histogram and Texture Features",* International Journal of Computational Intelligence Research 2017.

[32] JuliRejito,AtjeSetiawan Abdullah, Akmal, DeniSetiana, and Budi NuraniRuchjana *"Image Indexing using Color Histogram and k-means Clustering for Optimization CBIR in Image Database",* The Asian Mathematical Conference 2016, IOP Conf. Series: Journal of Physics: Conf. Series **893** (2017)

Authors Profile

Ms.D.Angayarkanni is working as a Assistant Professor in Computer Science,Bharathidasan University Constituent College,Lalgudi. She is pursuing Ph.D in Srimad Andavan Arts and Science College,Trichy,India.

Dr. L. Jayasimman is working as a Driector,MCA Department, Srimad Andavan Arts & Science College, Trichy, India. He received his M.Tech Degree in Bharathidasan University, Trichy, India in 2008 and has completed his PhD (Computer Science) in Bharathidasan University in 2014.He has more than 15 years of teaching experience and published more than 50 research papers in national and international journals. He has received an award Doctor of Letters from International Economics University,SAARC countries. He is also a member in International Association of Engineers.

© 2018, IJCSE All Rights Reserved