Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 8, August 2013, pg.169 – 173

RESEARCH ARTICLE

Image Color Extraction and Retrieval Using Classification Techniques

S.Kousalya[1], Dr. Antony Selvadoss Thanamani[2]

[1]Research scholar, Computer science (Aided) & NGM College, India

[2] Head of Department, Computer science (Aided) & NGM College, India

[1] *kousi.vs09@gmail.com*, [2] *selvdoss@gmail.com*

Abstract— Image mining is the thrust area in data mining which can be used to extract latent sequential data from the generic image series. Pixel measures or objects can be extracted effectively with high degree of precision if it is identical with the stored pattern of data sets in a right mode. The other area in the Image mining system is the Content-Based Image Retrieval (CBIR).Experiments with colour similarity mining technique of extracting colour from specific images and retrieving the similar pixels using Euclidean distance measure, pixels are grouped on the basis of nearest neighbour algorithm. This paper evaluating given the image and measure the pixels in image query processing.

Key Terms: - Content Based Image Retrieval; nearest neighbour; Euclidean distance; similarity measures

I. INTRODUCTION

Image mining presents special characteristics owing to the majesty of the data that an image can show. Every day, an immense amount of image data is being fabricated such as medical diagnosis, space research, remote sensing, agriculture, industries and even in the educational field [2]. But, regrettably there are certain technical hitches in gathering these data in an exact way. [1]Due to this gap, the data cannot be processed and to arrive at any finale. The web is measured the largest universal image repository. Image mining normally deals with the study and enlargement of new technologies that allow accomplishing this subject. Image mining is not only the simple veracity of recuperating relevant images; the aim is the innovations of image patterns that are noteworthy in extract the colours foreground and background from the given collection of images. Image retrieval is a fast rising and complicated research area with regard moving images. [4]Many Content Based Image Retrieval (CBIR) system prototypes have been prospect but few are used as commercial systems. CBIR aims at measuring colours for specific images that are similar to a given query colours. It also focuses at developing new techniques that support effective searching and extracting of image libraries based on automatically derived imagery colours and pixel. With the generated features, mining can be carried out using data mining techniques to discover significant patterns.

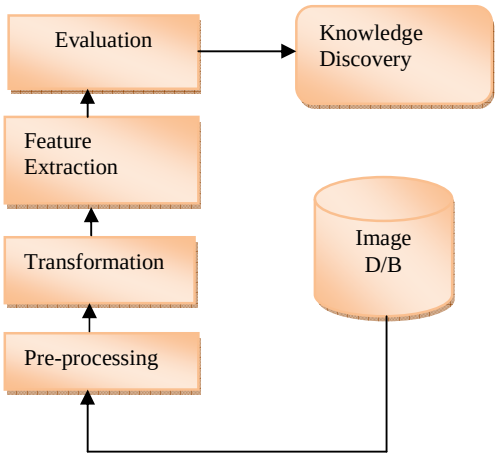

Fig 1: Image Process

The essential challenge in image mining is to expose out how pixel representation enclosed in a raw image or image sequence can be processed to recognize pixel level image objects and relationships. [6]The system then identifies those stored images whose feature values match those of the query most closely, and displays them. Colour features are usually represented as a histogram of intensity of the pixel colours.

II. RELATED WORK

Ji Zhang, Wynne Hsu and Mong Li Lee [1] proposed an efficient information-driven framework for image mining. In that they made out four levels of information: Pixel Level, Object Level, Semantic Concept Level, and Pattern and Knowledge Level. To achieve that High-dimensional indexing schemes and retrieval techniques are incorporated in the framework to maintain the flow of information among the levels. Wynne Hsu,Mong Li Lee and Ji Zhang [2] examines the research issues in image mining, developments in image mining, predominantly, image mining frameworks and suggests some future research guidelines for image mining. Dr.V.Mohan, A.Kannan[3] aims to the image mining is to remove the data loss and extracting the meaningful information to the human expected needs and colour based and texture based image retrieval yields high accuracy. That is, it retrieves the most matching images from the collection of the images, with respect to the query image. Peter Stanchev[4] proposed method is the possibility of retrieval using high level image semantic features.

III. IMAGE COLOR EXTRACTION SYSTEM

A. Colour Features

The colour feature extraction technique includes colour image segmentation. The standard RGB image is converted as L*u*v* image, where L* is luminance, u* is redness–greenness, and v* is approximately blueness–yellowness. Twelve hues are used as fundamental colours. There are yellow, red, blue, orange, green, purple, and six colours obtained as linear combinations of them. Five levels of luminance and three levels of saturation are identified. This results that every colour is transferred into one of 180 references colours. After that clustering in the 3-dimensional feature space is performed using the nearest neighbour algorithm. After this step the image is segmented as N regions, every of which is presented in extended chromaticity space.

B. Retrieval Systems

Image mining requires that images be retrieved according to some necessity terms. The requirement specifications can be classified into three levels of increasing complication: Comprises image retrieval by primitive features such as color, texture, shape or the spatial location of image elements. Comprises image retrieval by derived or logical features like objects of a given type or individual objects or persons. Comprises image retrieval by abstract attributes, involving a significant amount of high-level analysis about the meaning or purpose of the objects or scenes depicted.

C. Euclidean Distance

The Euclidean distance between points x and y is the length of the line segment connecting them ($\overline{xy}$).If x= (x_1,x_2,...x_n) and y= (y_1,y_2,..y_n) are two points in Euclidean n-space ,then the distance from x and y.

$dp(x, y) = ||x - y||_p = (\sum|xi - yi|p)1/p.$

This is the familiar straight line distance that most people are identified with if the two pixels that we are considering have coordinates (x_1,y_1) and (x_2,y_2) then the Euclidean distance given by

$D_{Euclid} = (x_2-x_1)^2 + (y_2-y_1)^2$

Most standard techniques treat color prediction as a problem of clustering points in three-dimensional space, where the points represent colors found in the original image and the three axes represent the three color channels. Almost any three-dimensional clustering algorithm can be applied to color quantization, and vice versa. After the clusters are located, typically the points in each cluster are averaged to obtain the representative color that all colors in that cluster are mapped to. The color channels are usually red, green, blue, yellow, magenta, purple and black but another popular choice is the Lab color space, in which Euclidean distance is more consistent with perceptual difference. If the palette is fixed, as is often the case in real-time color prediction systems such as those used in operating systems, color quantization is usually done using the "straight-line distance" or "nearest color" algorithm, which simply takes each color in the original image and finds the closest palette entry, where distance is determined by the distance between the two corresponding points in three-dimensional space. In other words, if the colors are (r_1, g_1, b_1) and (r_2, g_2, b_2),

We want to minimize the Euclidean distance:

$$\sqrt{\{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2\}}$$

Before this time, most color classifiers was done using the population algorithm or population method, which essentially constructs the equal-sized ranges and assigns colors to the ranges containing the most points.

D. Nearest neighbour algorithm

The nearest-neighbour method sufficiently large training set size **n**, the error rate of the NN classifier is less than twice the bayes error rate. Pixel quantify are usually in very low resolutions. They rely on careful placing of individual similar pixels often with a limited palette of colors.CNN (Condensed Nearest Neighbour) is an algorithm designed to reduce the data set for NN classification. It select's the set of prototypes U from the training data, such that NN with U can classified.

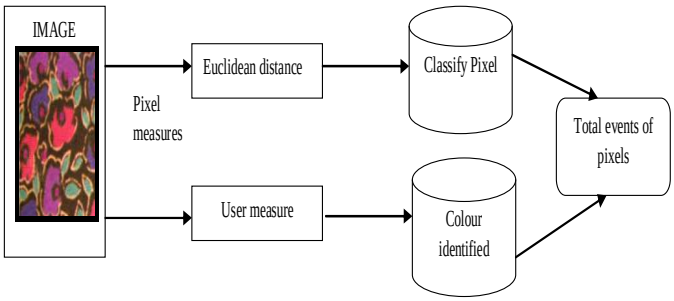

Fig. 2: Color retrieval taxonomy

E. Classify Each Pixel Using the Nearest Neighbor Rule

Each color marker now has an 'x' and 'y' value. It can classify each pixel in the specific image by calculating the Euclidean distance between that pixel and each color marker. The smallest distance will catch that the pixel most closely matches that color marker.

Color labels are 0 = background, 1 = red, 2 = green, 3 = purple, 4 = magenta, and 5 = yellow

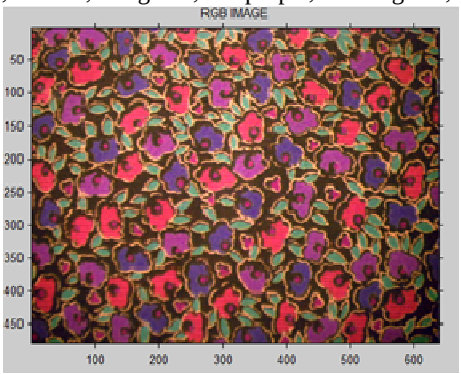

Fig 3: Raw image data

Present cases where significant colour shifts can be found. One of the numerical criteria for colour image prediction is to minimise the maximum variance between original pixel colour and the equivalent quantified colour, which provides better results than empirical algorithms. Another numerical criterion is to minimise the maximum difference between original and quantified pixel values.

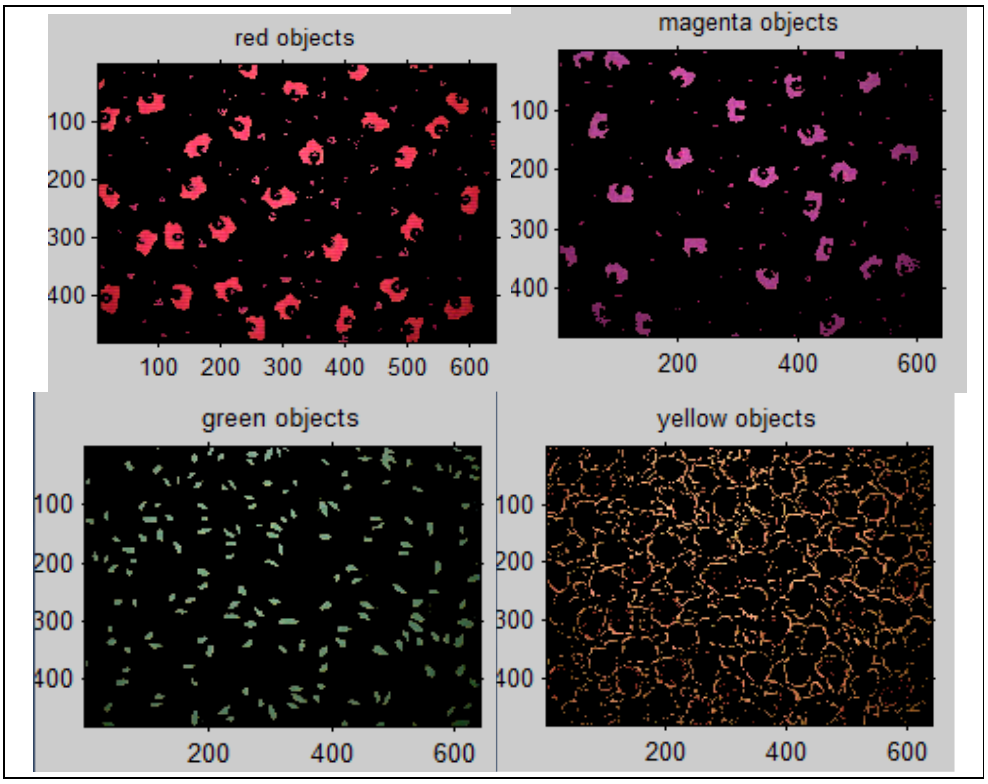

Fig 4: Nearest neighbor classification separated the different color pixels

On this set, subsets or groups of pixels with similar colour were identifying, like those presented on Fig3. For finding the similarity subsets, we asked four ranges to rank the image set with reference to their colour similarity. The interviewed people confirmed the groups of same colour images. A quantity of colour spaces have been used for colour illustration based on the perceptual concepts. There is no conformity on which is the best choice. Besides, its desirable characteristics are completeness, uniformity, compactness, and user oriented. Completeness means that it must include all perceptible different colours. Uniformity means that the measured proximity among the colours must be directly related to the rational similarity among them. Compactness means that each colour presents a perceptual difference from the other colours. Suitable spaces must be user oriented and based on an intuitive combination of the three basic attributes of the colour

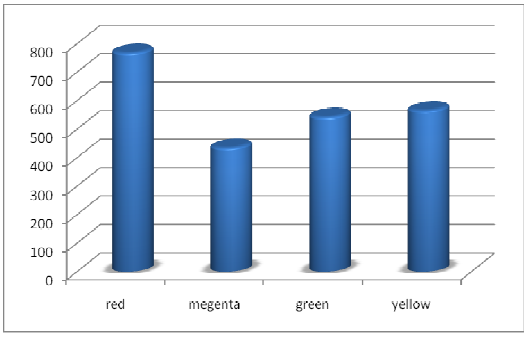

Fig 5: Segmented pixels x*y* space

Nearest neighbor classification separated the different color populations by 'x*' and 'y*' values pixel that were classified into separate colors.

IV. CONCLUSIONS

The main objective of the image mining extracting colour from specific images and retrieving the similar pixels using Euclidean distance measures. Combinations of colours resulting represent the image sequence. It is clear that all evidences show the proposed model is qualified for mining and segmenting image data.

REFERENCES

[1] Ji Zhang, Wynne Hsu and Mong Li Lee "An Information-Driven Framework for Image Mining" Database and Expert Systems Applications in Computer Science, 2001, Volume 2113/2001, 232-242, DOI: 10.1007/3-540-44759-8_24

[2] Wynne Hsu, Mong Li Lee and Ji Zhang "Image Mining: Trends and Developments" Journal of Intelligent Information Systems Volume 19, Number 1, 7-23, DOI:10.1023/A:1015508302797

[3] Dr.V.Mohan, A.Kannan " Color Image Classification and Retrieval using Image mining Techniques" International Journal of Engineering Science and Technology Vol. 2(5), 2010, 1014-1020

[4] Peter Stanchev " Using Image Mining For Image Retrieval" IASTED conf. "Computer Science and Technology", May 19-21, 2003 Cancum, Mexico, 214-218

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

[6] 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.

[7] W. Ma, Y. Deng and B. S. Manjunath. "Tools for texture/color based search of images."SPIE International conference - Human Vision and Electronic Imaging: 496-507, February 1997.

[8] Ji Zhang, Wynne Hsu, Mong Li Lee. Image Mining: Issues, Frameworks and Techniques, in Proceedings of the Second International Workshop on Multimedia Data Mining (MDM/KDD'2001), San Francisco, CA, USA, 2001.

[9] Aura Conci., Everest Mathias M. M. Castro "Image Mining By Color Content "

[10] V. Balamurugan and P. Anandhakumar "Neuro-Fuzzy Based Clustering Approach For Content Based Image Retrieval Using 2DWavelet Transform" International Journal of Recent Trends in Engineering, Vol. 1,No.1, May 2009

[11] Ji Zhang, Wynne Hsu and Mong Li Lee "An Information-Driven Framework for Image Mining" Database and Expert Systems Applications in Computer Science, 2001, Volume 2113/2001, 232-242, DOI: 10.1007/3-540-44759-8_24

[12] Kun-Che Lu And Don-Lin Yang "Image Processing and Image Mining using Decision Trees" Journal Of Information Science And Engineering 25, 989-1003 (2009)

[13] P. Rajendran, M. Madheswaran " Novel Fuzzy Association Rule Image Mining Algorithm for Medical Decision Support System" 2010 International Journal of Computer Applications (0975 - 8887) Volume 1 – No. 20

[14] P. Rajendran, M.Madheswaran "Hybrid Medical Image Classification Using Association Rule Mining with Decision Tree Algorithm" Journal Of Computing, Volume 2, Issue 1, January 2010, ISSN 2151-9617

[15] Rajshree S. Dubey "Image Mining using Content Based Image Retrieval System" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 07, 2010, 2353-2356

