# Cluster Oriented Image Retrieval System

Mahip M.Bartere
M.E(Scholar)
Sipna College of Engineering & Technology,Amravati(M.S)

Prashant R.Deshmukh
Asst. Professor
Sipna College of Engineering & Technology,Amravati(M.S)

## ABSTRACT

Image mining presents special characteristics due to the richness of data that an image can show. Effective evaluation of results of image mining by content requires that the user point of view is used on the performance parameters. Comparison among different mining by similarity systems is particularly challenging owing to the great variety of methods implemented to represent likeness and the dependence that the result present of the used image set. Other obstacle is lag of parameter for comparing experimental performance. In this paper we propose an evaluation framework for comparing the influence of THE distance function by image mining by color and also a way to mine an image from its name. Experiments with color similarity mining by quantization on color space and measure of likeness between a sample and the image results have been carried out to illustrate the proposed scheme. Important aspects of this type of mining are also described.

## Keywords

Color Based Image segmentation, Deviation Factor, Image comparison, Clustering, Text Based Mining.

## 1. INTRODUCTION

Images are generated at increasing rate by sources such as military reconnaissance flights ; defense and civilian satellites; fingerprinting devices and criminal investigation; scientific and biomedical imaging; geographic and whether information system; stock photo database for electronic publishing and News Agency; Fabric and fashion design; art galleries and museum management; architectural and engineering design; and WWW search engine. Most of the existing image management system is based on the verbal description to enable there mining. A keyword description of the image content, created by some user on input, in addition to a pointer to the image data is the base of the system. Image mining is based on standard mining. However verbal description is always inadequate, error prone and time consuming. The majority of the pictorial information in the real world images (as that in figure cannot be fully captured by text and numbers due to limitation power of languages. A more efficient approach is gather when image example is given by the user on input to the mining process. Automatically generate matching is required then for an efficient image mining. The basic idea is to extract the characteristics feature similar to that of an object recognition schemes. After matching the images are ordered with respect to the query image according to their similarity measured and displayed for viewing. In this work, we present a framework for considering the influence of this distance function on color mining. The framework accesses a system's quality from the point view of users; it provides a basic set of attributed to characterize the ultimate utility of systems. Then we analyze example of mining by color and present some conclusions.

## 2. LITERATURE REVIEW

Early Techniques were not generally based on visual features but on the textual annotation of images. In other words, images were first annotated with text and then search using text based approach from traditional database management system. Text based image retrieval system uses traditional database techniques to manage images. Through text description, images can be organized by topical or semantic hierarchies to facilate easy navigation and browsing base on standard Boolean queries. However since automatically generating descriptive text for wide spectrum of images is not fisible, most text based image retrieval system requires manual annotation of an images .Obviously, annotating images manually is a cumbersome as an expensive task for large image databases, and is often subjective, context sensitive and incomplete. as a result it is difficult for traditional text based method to support variety of task dependant queries.

## 3. PROPOSED METHOD

Cluster based image mining techniques based on

- Mining by color content
- Mining by Text

### 3.1 MINING BY COLOR CONTENT

In a proposed system, we are having two approaches for images retrieval that is mining by color contents and mining by text. Images are mostly RGB, Gray Scale and Binary types which are easily possible to segments as per their pixels values.RGB images are segmented as per the color content, Gray Scale images are segmented as per their average RGB pixels value and Binary Images as per their ON/OFF pixels status. Image mining by color contents include search for images based on their segmented color contents. Let see Dataflow Diagram for proposed image mining technique based on their color contents is shown in figure 1.

Segmentation of images based on their color contents is a major area of research. In a proposed method we are using C-means clustering technique to clusterize an image based on their RGB colors as per equation

$$Cn = \int_1^n \sum_{i=1}^{M} |Pi - Pi + 1| \quad \text{.....eq}^n \text{ 3.1.1}$$

Where

$C_n$= C means cluster.

n= No of Clusters.

$P_i$= Backward pixel

$P_{i+1}$= Forward Pixel

M= No of Pixels in an image.

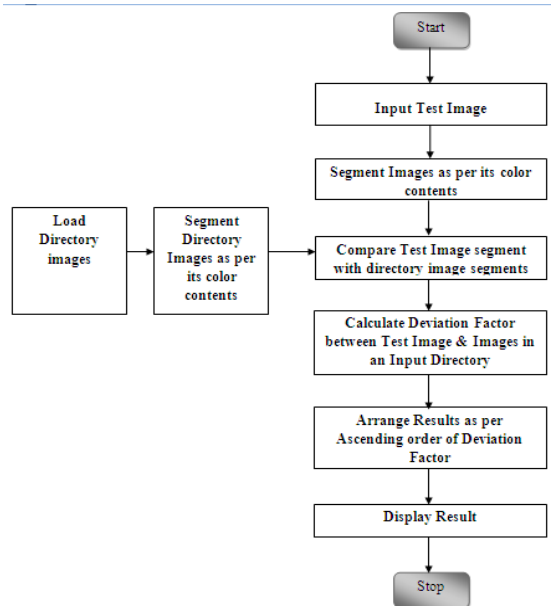


**Fig.1. DFD for Mining by Color**

In c-means clustering technique the pixels are scanned as per Raster Scan method from left to right and top to bottom way. Whenever difference occurs with backward and forward pixel we set it as OFF else we proceeds our scanning as per the equation

If $|P_i - P_{i+1}| = 0$ $\quad P_{i+1} = P_i$

Else

$P_i - P_{i+1} \neq 0$ $\quad P_i = 0$ … eq$^n$ 3.1.2

Deviation Factor $D_f$ is a difference between number of segments of directory image and input test image. Deviation factor is a measure of similarity and difference between two images which can be represented with an equation

$$Df = |Ctr - Cts| \text{ ….. eq}^n \text{ 3.1.3}$$

Where

$C_{tr}$= Directory Image Segments.

$C_{ts}$= Input Test Image Segments.

An overall searching time for color based image mining is depended on number of color segments in an image.

$$O(tc) \alpha\ n \text{ …….. Eqn 3.1.4}$$

Where O (tc) = searching time for color based image mining

We use best fit searching technique to find out an image segment express with

$$O(C) \alpha\ O(n) \text{ …….. Eqn 3.1.5}$$

Where O(C) = no of segments comparisons.

## 3.2 MINING BY TEXT

From the literature survey, it is observed that, image mining by text is a complicated process and fails in a domain where single image contains multiple images content (like single image with nature, man, aero plane etc.).To have accuracy in image mining process through text is quite impossible. We proposed a novel method that uses any one of the input name image as a reference image for search and rest of the searching mechanism as per section 3.1.Let see a data flow diagram for image mining based on the text as follows

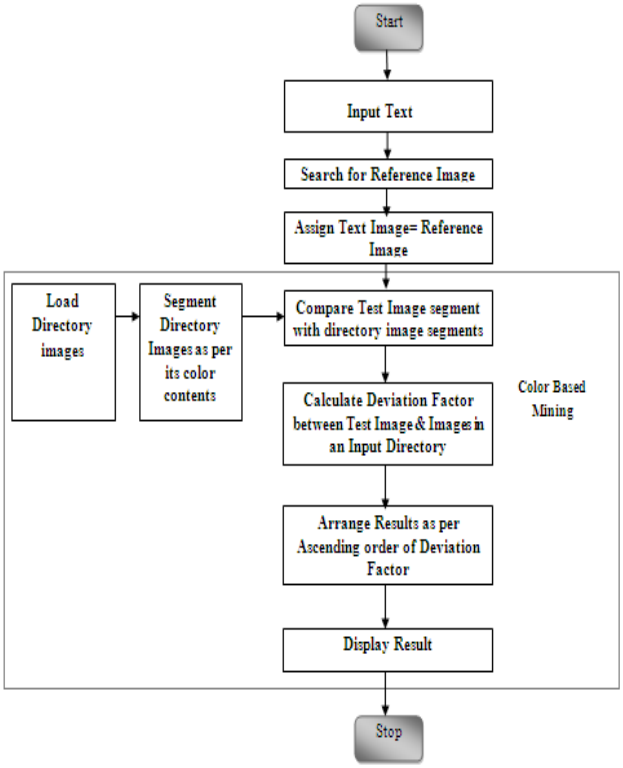


**Fig.2.DFD for Mining by Text**

Overall searching time for text based image mining is express as

$$O(Tt) = O(ft) + O(tc) \text{… Eqn 3.2.1}$$

Where O (Tt) = Searching time for text based image mining

O (ft) =First fit search time.

Rest of the mechanism for text based image mining is as per section 3.1.

## 4. EXPERIMENTAL RESULTS


Figure 4.1 Search by color content

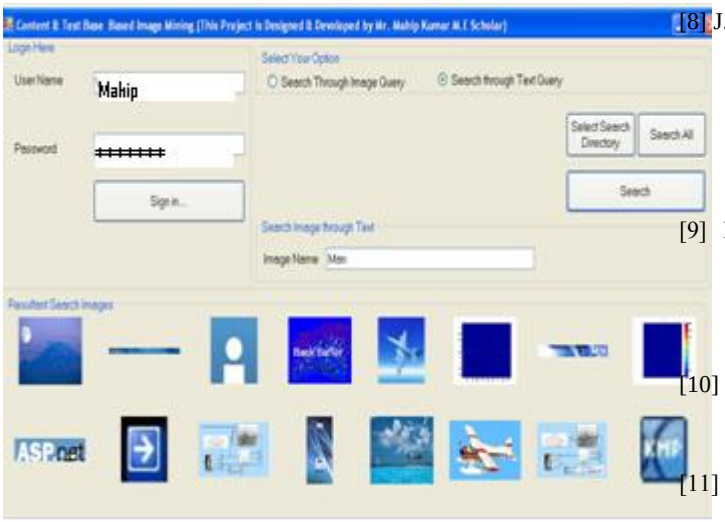


Figure 4.2 Search by text

Original Image Color Base Clustered Image

Figure 4.3 Color Based Segmentation

## 5. REFERENCES

[1] V. N. Gudivada, V. V. Raghavan, "Content-Based Image Retrieval Systems", IEEE Computer,September, 18-22, 1995.

[2] S. A. Stricker, "Bounds for discrimination power of color indexing techniques", Proc. SPIE, pp. 15-24, 1994.

[3] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner and W. Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions", IEEE Trans. Pattern Analysis Machine Intell., Vol. 17, No.7, pp. 729-736, 1995.

[4] A. Pentland, R. W. Picard, S. Sclaroff, "Photobook: content-based manipulation of databases", Int. J. Computer Vision, Vol. 18 , No. 3, 233-254, 1996. http://www-white.media.mit.edu/~tpmink/photobook

[5] M. J. Swain, D. H. Ballard, "Color Indexing", Int. J.Comp.Vision, Vol. 7 , No. 1, 11-32, 1991.

[6] M. Flickner, H. Sawhner, W. Niblack, J. Ashley, Q.Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petrovic, D. Steele, P. Yanker, "Query by image video content: the QBIC system", IEEE Computer, September, 23-32, 1995.http://wwwqbic.almaden.ibm.com/~qbic/

[7] P. K. Kaiser, R. M. Boyton, Human Color Vision, Second Ed., Washington, D.C.: Optical Society of America, 1996.

[8] J. R. Bach, C. Fuller, A Gupta, A Hampapur, B. Horowits, R. Humphrey, R. C. Jain, C. Shu, "Virage image search engine: an open framework for image management", Symposium on Electronic Imaging: science and technology-storage & retrieval for image and video databases IV, IS&T/SPIE, 76-87, 1996. - http://www.virage.com

[9] B. Hill, Th. Roger, F. W. Vorhagen, "Comparative analysis of the quantization of color spaces on the basis of the CIELAB color-difference formula",ACM Transaction on Graphics, Vol. 16, No. 2, April, 109-154, 1997.

[10] H. Zhang, Y. Gong, C. Y. Low, S. W. Smoliar, "Image retrieval based on color features: an evaluation study: Proc. of SPIE, 2606, pp. 176-187, 1997.

[11] C. Z. Ren, R. W. Means, "Context Vector Approach

To Image Retrieval", Proc. IEEE ICIP, Vol. 1, No 407, 1997.