

Content Based Image Retrieval Based on Shape, Color and Structure of the Image

T. Venkat Narayana Rao¹, Somasani Jyothi²
Professor¹, Student², Department of C.S.E
Sreenidhi Institute of Science of Technology
Yamnampet, Hyderabad, TS, India.

Abstract: In the recent era, as technology is growing rapidly the usage of social media is also increasing as a result large databases are required for storing the images. With the advancements in the technology, the storage of these images in computers has become possible. But retrieving the images is becoming a big task. We need to store them in a sequential manner and retrieve them when required. This paper details retrieval of images by considering the features related to content like shape, color, texture is called CBIR (content based image retrieval). As it is very difficult to extract the pictures in such huge data bases so we chose this technique which aim at high efficiency.

Keywords: Image retrieval, color, texture, shape and content based.

I. Introduction

With the advancements in the technology, the storage of the images in computers had become possible. But how to retrieve the required image with precision is still a vital problem. In the early years we retrieved the images using text. It is one of the methods for pictures extraction. But CBIR(content based image retrieval) is the best method and one of the well planned and structured technique for extracting or reviving the images from such huge databases. It considers the features of an image such as shape, color and structure of image instead of searching by text. So it has very good efficiency and it is an easy way to retrieve images. In this process the alike images are revived by simply giving an input as a query image and based on that images can be retrieved from data base by matching with shape, texture and color etc.

The word "content-based image retrieval" was introduced in 1992 by T. Kato. It was used to define experimentations into instinctive revival of images from a data stores or databases i.e. on the basis of colour and shapes present. From then, this word is used for describing the process of extracting similar pictures or images from a huge data bases by considering certain features like shape, colour and texture etc. For this process tools, algorithms, and techniques are based on the fields such as pattern matching and identification, processing of signal and statistics etc[4][5].

This paper describes how the user would give a query image and by taking that image how the system detects different properties of the input image and provide the similar images as output. During this process it extracts the properties from a given input image and matches with the pictures in the databanks, and then the same pictures are projected to user as output. Figure 1 represents the architecture of CBIR.

II. Architecture

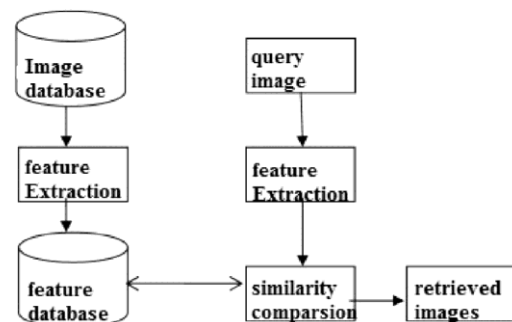


Figure 1: Architecture of CBIR

III. CBIR Methods

A. Colour

The similar images are extracted by considering some features like shape, color, size. By considering the color feature, images can be extracted very easily because we have color histograms and RGB values to identify the colors. It will divide the image into segments and by considering the regions it will identify the similar images. Generally rather than size and shape we can identify the similar images, objects, etc. very easily by color. Color is defined as the property controlled by an object of making different perceptions on the eye resulting in the way it mirrors or produces light. Colors can be represented in three dimensional way like RGB(red,blue,green) or HSV(hue, saturation,value) or HSB(hue,saturation,brightness). HSB and HSV are based on person's approach of hue, saturation and brightness. As we have images in the format of JPEG, JPG,GIF,we can use RGB colour cache the information of image. Simply the definition of RGB is a unit in the form of a three dimensional cube which has red, blue, green axes. Thus,

this vector describes a 3D colour space which has three colours as mentioned in the figure 2. If we have all the 3 axes coordinates as zeros then the colour will be black, if all the coordinates are 1 then the colour will be black [6] as shown in figure 3.

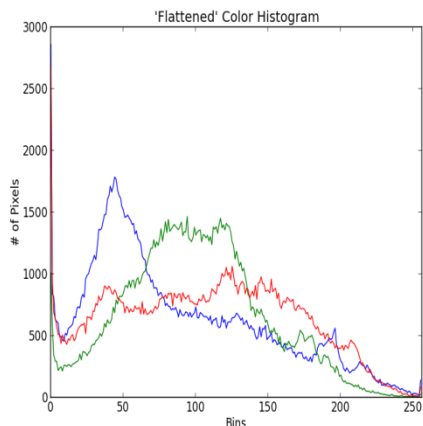


Figure 2: Color histograms representation

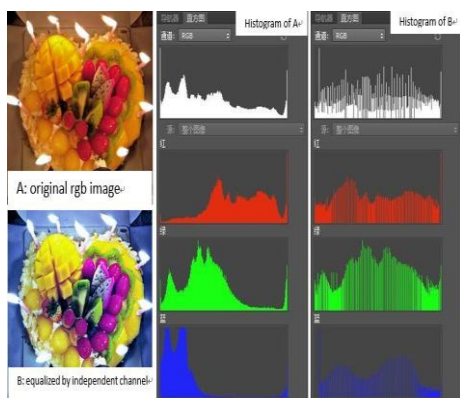


Figure 3: Image comparison using Color histograms

An image may contain perceptible data or syntactical data. The perceptible data of image can be shown as colour, shape of picture texture and pertaining to relations. After extracting the perceptible features, they are considered as feature vectors and they are to be stored in the data base. Then the images retrieval will happen. By considering several techniques we match the query image with the pictures in the huge data bases[8][9]. Then the same features images are extracted and they are displayed to the user. The result of this system produces one or more images shown in figure 4 and 5.

Metrics of color

i. Minkowski-Form Distance

It is considered for measuring the space between any two pictures when we find all the features of those images are having same importance. The Minkowski-form distance is very useful for the totaling or calculating of space between two pictures when each of the picture feature vectors are

equally treated and they both does not depend on each other. That is they are independent.

ii. Histogram Intersection

Histogram Intersection is mainly used for measuring of the similarity among color images. It is one of the form of Euclidian distance.

iii. Quadratic Form Distance

The major drawback of Minkowski distance is it considers each and every bin in an independent way. Sometimes it will not consider the likeness between certain pairs of bins which has exactly the similar features.

iv. Mahalanobis Distance

It is used to measure the distance between a distribution D and point P in the 3D space. It is a multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. If the point is at mean D then distance is considered to be zero, and increases as P moves away from the mean along every component axis space. It is used to measures the number of standard deviations from P to the mean of D. If each of these axes is re-scaled to have unit variance, then the Mahalanobis distance corresponds to standard Euclidean distance in the transformed space. Mahalanobis distance (or "generalized squared interpoint distance" for its squared value) can also be defined as a dissimilarity measure between random vectors and of the same distribution with the covariance matrix S.

If the covariance matrix is a diagonal , then that results in distance measure which is known as *standardized Euclidean distance*.

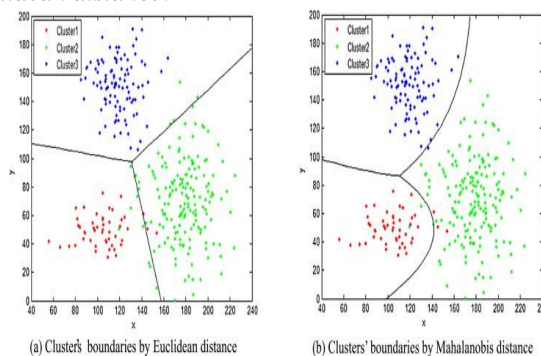


Figure 4: Clusters boundaries by Euclidean distance and Mahalanobis distance

B. Comparing features using colour based CBIR

Retrieval of pictures on the basis of color is the important method for CBIR. They are insensitive to translation, scale changes and rotation etc. The calculations for color retrieval is very easy. The color histogram firstly divides the color space into several regions to compute various metrics. Each region or interval is considered as a bin.

Thus, it counts the number of pixels that each color falls into each interval. Therefore, the color becomes quantized, by counting pixels in color histogram i.e. calculated where each color falls into its respective intervals. Color features are divided into both global color histogram and block color histogram[11][12][13].

C. Global Color Histogram

In general the RGB color space converts to other color spaces. As we know that RGB color space was not able to meet the human visual supplies of people for picture retrieval, this conversion takes place. The global color histogram can be computed as:

Step 1. Firstly convert the pictures in the database into HSV space from RGB space.

Step 2. Quantization of the images takes place using some formulas of HSV.

Step 3. Count the value of each feature.

Step 4. After calculating each value similarity is evaluated by Euclidean distance using formula.

By using global color histogram we can extract features of image easily. It involves highly efficient calculation and matching can be done based on that calculation. This is the main advantage. The major drawback is that it will only calculate the frequencies of colors.

D. Block color histogram based CBIR

In block color histogram model it is divided into $m \times m$ blocks. The computation process if retrieval is easy if the block is large. In this a 2-D space is segmented into 3×3 and it is used most efficiently[6][7].

For each block in color space the calculation takes place. With this normalized color features for each block can be calculated. Here each block will be given a weight coefficient. Middle block weight is always high.

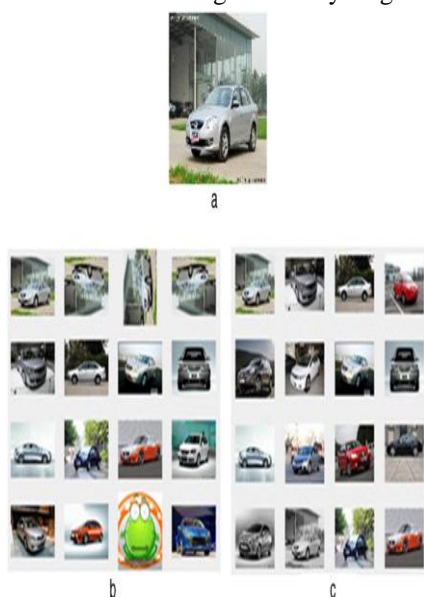


Figure 5: The color based CBIR.

Given a query image as car and based on the color feature similar images can be retrieved.

3.2 Texture

Texture is also considered to be one of the important characteristic of an image. Even by using this feature retrieval of images is possible. To attain this many algorithms like gray level concurrence matrices, the Marcov random field model, and Tamura texture feature are available[2].

In texture analysis many terms are considered like:

- Coarseness
- Line likeness
- Roughness
- Directionality
- Contrast etc.

We can approach the texture analysis by three ways:

- Statistical
- Structural
- Spectral

Statistical characteristics of the grey levels of the points/pixels which have a surface picture of the texture are represented by the Statistical techniques. Generally these types of properties are calculated using Grey level co-occurrence matrix of the surface.

In Structural techniques we consider textures as simple and small structures called “Texel’s” (or texture elements). These are arranged in a regular manner on some surface according to rules of surface arrangement.

Spectral techniques mainly depend on the Fourier spectrum properties and it gives information about global periodicity of the grey levels of a surface by detecting high-energy peaks in the Fourier spectrum.

Most used statistical texture representations are:

- Co-occurrence Matrix [figure 6]
- Tamura Texture
- Wavelet Transform

Co-occurrence Matrix:

This was developed by R.M. Haralick. This is the illustration of texture features that describes the grey level three-dimensional dependency of texture[1][3]. It is simple in mathematical form as shown below:

- Let $q(i,j)$ be a position operator
- let k be an $m \times m$ matrix
 $k[i][j]$ is defined as number of times that points with grey level (intensity) $j[i]$ occur, in the position specified by q , relative to points with grey level $j[j]$.
- Let d is $m \times m$ matrix is produced by dividing k with the total number of point pairs that satisfy q , $d[i][j]$ is a measure of the joint probability that a pair of points satisfying q will have values $j[i], j[j]$.

- C is a co-occurrence matrix and it is represented by q .
 Examples for the operator q are: “ i above j ”, or “ i one position to the right and two below j ”, etc

1	2	1	3	4
2	3	1	2	4
3	3	2	1	1

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	0	0	0	0	0
2	0	1	0	2	0	0	0	0
3	0	0	1	1	0	0	0	0
4	0	1	0	0	1	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Figure 6. Image and Co-Occurrence Matrix

Wavelet Transform

Textures are modelled as some quasi-periodic patterns which has either spatial or frequency representation. The wavelet transform is used to transform the picture into a simple multi-scale illustration which has both spatial and frequency features.

In Fourier transform we use wavelets to describe and represent signals. But in general to represent wavelet transform sine functions are used. The basic idea of this wavelet transforms is that there should be a transformation change in time extension, but not shape.

Wavelets are limited in period, yet the mean value of a wavelet is zero. Simply a wavelet can be defined as a waveform that is confined in both frequency and period. A Fourier transform is used to convert a signal into a continuous series of sine waves, each has certain amount of constant frequency and amplitude and of boundless period, mainly the objects of real world.

Examples of different kind of wavelets are Haar, Daubechies, Coiflet, Morlet. Among all these Haar, is one of the most simplest wavelet and it is used popularly. Daubechies are very important for current wavelet applications [2]. They are explained below:

Haar Wavelet

Figure 7 represents Haar Wavelet.

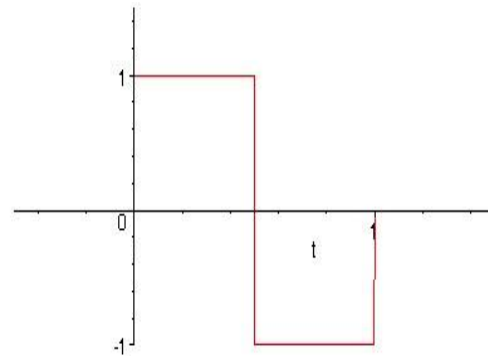


Figure 7: The Haar Wavelet

Daubechies Wavelet

Figure 8 represents Daubechies wavelet family.

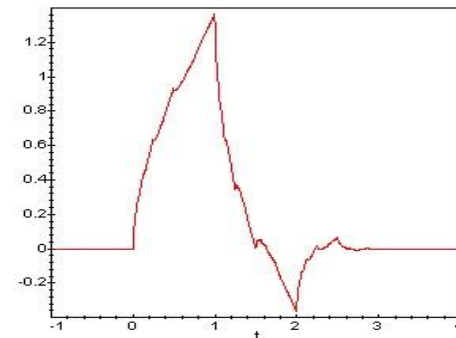


Figure 8: The Daubechies Wavelet

3.3Shape

In addition to texture, colour, shape is also considered to retrieve the images. The characteristic surface configuration of an object is probably defined as shape. But it is not an efficient way to retrieve the images because it is very hard to describe the methods in it. One of the major problem is segmentation of intersected target [7][14].

Understanding user's visual point of view, Tamura analysed the texture portrayal by performing computations to coarseness, contrast, and directionality which are the important features of texture. All these features of textures are calculated using various algorithms.

Coarseness is simple stated as quality of image being rough, or calculated as mean size of areas in the space that has similar concentration or intensity. Measure of quality of being very clear and being bright.

Directionality is the calculation or estimation of directions of the grey values within the image.

In general by seeing the shape of an object we can guess what the object is. So by using this shape feature we can identify the image easily. Generally this shape representation can be carried out in two ways:

1. Boundary-based.

In Boundary-based shape depiction we consider the outer boundary of the shape. By considering the outer characteristics of the image i.e. number of pixels which are along the boundary of object, the classification can be done figure 9.

Example:

- Models of polygon
- partitioning boundaries
- Fourier Descriptors

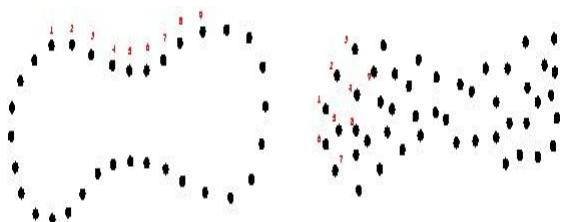


Figure.9: Boundary-based & Region-based

2. Region-based.

In Region-based shape representation we deal with complete shape region considering only its internal features but not the external characteristics and metrics. The number of pixels of image that is considered is in that particular region only.

Example :

- Super quadrics
- Fourier Descriptors
- Implicit Polynomials
- Blum's skeletons

The important representations for shape categories are: Moment Invariants and Fourier Descriptor.

The transformed Fourier boundary as the shape feature is considered in the model of Fourier Descriptor whereas Moment invariants is to use moments of region. Representation of shape happens when a 3-D real world object onto a 2-D image plane, the object data is lost on a single dimension. This may result in inaccurate results, even this is not the one problem but also shape is corrupted with noise, distortion, defects etc. All these may distract the shape of the image. Shape-based image retrieval mainly focus on calculating the likeness between shapes by considering the features. Shapes can be detected by mainly considering the geometric figures.

Some of the parameters of shape are variance, centre of gravity, convexity, mean, dispersion, Euler number, hole area ratio, profiles and mass etc.

Some of these are described as follows[10].

Mass represents the number of pixels contained in one class. It is calculated as:

$$mass = \sum_{xy} h(x, y)$$

where

$$h = \begin{cases} 1 & \text{if } s(x, y) \in c \\ 0 & \text{if } s(x, y) \notin c \end{cases}$$

Centre of mass is defined as Centroid.

Where h is a small mask of cluster cover required image S(x, y). The co-ordinates (xc, yc) of the centroid are defined by the below formulae:

$$x_c = \frac{\sum_{xy} x * h(x, y)}{mass}$$

$$y_c = \frac{\sum_{xy} y * h(x, y)}{mass}$$

Below formulae calculates the mean and variance.

$$\mu_c = \frac{\sum_{x,y} I_{x,y} * h_c(x, y)}{mass}$$

$$\sigma_c^2 = \frac{\sum_{x,y} (I_{x,y} - \mu_c)^2 * h_c(x, y)}{mass}$$

From centroid each region of the class is defined and sum of both will give dispersion. The distance is computed by the formula of Euclidean distance.

The dispersion is computed as:

$$Disp = \sum_i dist(O_c, O_{i,c})$$

Here, Euclidean distance is dist (Oc,Oi,c)
 centroid of class c= Oc
 centroid of region I of class c= Oi,c

IV. Future Enhancements

1. Extending the query processing to accept all file types as query image.
2. Testing the algorithms of CBIR for precision.
3. Extending features for face recognition and improvement in extraction using this face recognition.
4. Improving the performance of prominent boundaries detection method with advanced Wavelets.
5. Extraction of texture feature with time and cost effective.
6. Extraction based on the colour codes of the image, Shape correlation and the fore ground colour codes.

V. Limitations

The limitations of the developed CBIR system are:

1. The developed algorithms and application are resource hungry. Sometimes processing of images with large size may require high calculations which is possible with only some machines.
2. This system does not yield the accurate and expected results with low resolution images.
3. No proper indexing technique and there is a confusion while considering the input image.

VI. Conclusion

The extraction of the pictures earlier depends on text but in present scenario this can be done using content. The technique CBIR deals with pictures mainly on the basis of features such as colour, texture and shape, which is efficient than retrieving the image by text. We need not give any indexing or numbering to the images. In this technique it considers the images which are kept in database and compares with the image user gave as input (ie, query image) then by considering certain features like texture, colour and shape, similar images are retrieved using colour histograms and various algorithms. CBIR compares the images using these techniques and different distances are calculated like *Quadratic Distance Equation* etc. and for texture it considers wavelet equation *Euclidean Distance Equation* etc. Always CBIR system has a increasing pace of development in the research field. Studies are being developed to retrieve the images in faster way and in an efficient manner. This paper improves the accurateness of a CBIR system. It allows the system to extract more similar pictures by comparing to the query picture. The proposed methodology had increased the average accuracy when compared to the previous studies.

References:

- [1]. Barbeau Jerome, Vignes-Lebbe Regine, and Stamon Georges, "A Signature based on Delaunay Graph and Co-

- occurrence Matrix," Laboratoire Informatique et Systematique, Universiyt of Paris, July 2002.
- [2]. Sharmin Siddique, "A Wavelet Based Technique for Analysis and Classification of Texture Images," Carleton University, Ottawa, Canada, Proj. Rep. 70.593, April 2002.
- [3]. Thomas Seidl and Hans-Peter Kriegel, "Efficient User-Adaptable Similarity Search in Large Multimedia Databases," in Proceedings of the 23rd International Conference on Very Large Data Bases VLDB'97, Athens, Greece, August 1997.
- [4]. FOLDOC, Free On-Line Dictionary Of Computing, "co-occurrence matrix," May 1995, Available at: <http://foldoc.doc.ic.ac.uk/foldoc/foldoc.cgi?cooccurrence+matrix>
- [5]. Colin C. Venteres and Dr. Matthew Cooper, "A Review of Content-Based Image Retrieval Systems", Available at: <http://www.jtap.ac.uk/reports/htm/jtap-054.html>
- [6]. Linda G. Shapiro, and George C. Stockman, Computer Vision, Prentice Hall, 2001. Shengjiu Wang, "A Robust CBIR Approach Using Local Color Histograms," Department of Computer Science, University of Alberta, Edmonton, Alberta, Canada, Tech. Rep. TR 01-13, October 2001.
- [7]. R. Jain, R. Kasturi, and B. G. Schunck, Machine Vision, McGraw Hill International Editions, 1995.
- [8]. R. Datta, D. Joshi, J. Li, J.Z. Wang. Image retrieval: ideas, influences, and trends of the new age. ACM Computing Surveys, 40 (2) (2008), pp. 1–60. CrossRef | View Record in Scopus | Citing articles (2)
- [9]. P. Muneesawang, L. Guan. An interactive approach for CBIR using a network of radial basis functions. IEEE Transactions on Multimedia, 6 (2004), pp. 703–716. CrossRef | View Record in Scopus | Citing articles (60)
- [10]. Mathematical and computer modelling. Volume 54, Issues 3–4, August 2011, Pages 1121– 1127. Edited By Daoliang Li, Yande Liu and YingyiChe 11. Transitions of Mesogens. ISSN: 2319-8753 Co-Occurrence Matrix and Its Statistical Features as an Approach for Identification of Phase International Journal of Innovative Research in Science, Engineering and Technology (An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 9, September 2013.
- [11]. Škrjanec Marko. Bsc Thesis, (Mentor: MatejKristan), Fakultetazaračunalništvo in informatiko, Univerza v Ljubljani, 2013. Fruit Image Data set .
- [12]. Y. Liu, D. Zhang, G. Lu, and W. Ma , "A survey of content based image retrieval with high-level semantics," Journal of Pattern Recognition, vol. 40, Nov. 2007.
- [13]. Benjamin B. Kimia, "Symmetry-Based Shape Representations," Laboratory for Engineering Man/Machine Systems (LEMS), IBM, Watson Research Center, October 1999.
- [14]. Marinette Bouet, Ali Khenchaf, and Henri Briand, "Shape Representation for Image Retrieval", 1999.