

Content Based Image Retrieval Method using Fuzzy Heuristics

P.E. Rubini*

Department of Computer Science, SRM University, Chennai, India

*Corresponding author email: iniya29@gmail.com, Tel.: +91 9894188837

ABSTRACT

Content based image retrieval (CBIR) refers to image content that is retrieved directly, by which the images with features or containing certain contents will be searched in an image database. The main idea of CBIR is to analyze image information by low level features of an image, which includes color, texture, shape and space relationship of objects etc., and to set up feature vectors of an image as its index. A new CBIR search engine is proposed using three features and similarity is measured and controlled by fuzzy heuristics. CBIR Search Engine relies on the characterization of primitive features such as color, shape, and texture that are automatically extracted from the images. There are several techniques to deal with CBIR problems for retrieving the relevant images. CBIR proposed by using three methods. Colour feature is extracted by using histogram based method, texture feature is extracted by using Gabor filter and shape feature is by moment invariant algorithm. For searching the similar images with the database similarity measure is calculated and is controlled by using fuzzy. Fuzzy similarity measure is implemented by using mamdani fuzzy inference method. The use of these three algorithms ensures that the image retrieval approach produces images which are relevant to the content of an image query.

Keywords - Content-based, Fuzzy heuristics, Image retrieval, Search engine.

1. INTRODUCTION

Nowadays, with large number of digital images available on Internet, efficient indexing and searching becomes important for large image storage. In traditional approach labeling of images with keywords, provides the diversity and ambiguity of image contents. So, content-based image retrieval (CBIR) approach indexes images by low-level visual features such as color, texture and shape.

A typical CBIR system consists of two main parts: (1) feature extraction and (2) similarity measurement as shown in Fig.1. First, features such as shape, texture and colour, which constitute the image signature, are generated to represent the content of a given image. The similarity of a query image to the images in database is then measured using an appropriate distance metric. In typical content based image retrieval approach, a user submits an image based query which is then used by the system to extract visual features from images. The visual feature is based on the type of image retrieval. These features are examined in order to search and retrieve similar images from image database. The similarity of visual features between query image and image in a database is calculated by applying fuzzy rules.

2. BACKGROUNDS

In content-based image retrieval systems, a desirable image is retrieved, from the large collection of images stored in the image database, based on their visual content. The visual content of an image is represented by common attributes which are called features. They include 'shape of the image', 'colour histogram of the image' and 'texture of the image'.

2.1 Colour features

Colour feature is the most commonly used visual feature for image retrieval. Many colour models are available that can be used to represent images such as HSI, HSV, LAB, LUV and YCrCb. Colours play a major role in human perception. The most commonly used colour model is RGB, where each component represents colour, red, green and blue.

2.2 Texture features

Texture is another important feature of an image that can be extracted for the purpose of image retrieval. Image texture refers to surface patterns which show granular details of an image. It also gives information about the arrangement of different colours. There exist two main approaches for texture analysis. They include structural and statistical approaches. In structural texture approach, the surface pattern is repeated. In statistical texture; the surface pattern is not regularly repeated in the same pattern such as different flower

objects in a picture. Co-occurrence matrix is a popular representation of texture feature of an image. It is constructed based on the orientation and distance between image pixels. The wavelet transform is used for image classification based on multi-resolution decomposition of images. Among the different wavelet transform filters, Gabor filters were found to be very effective in texture analysis.

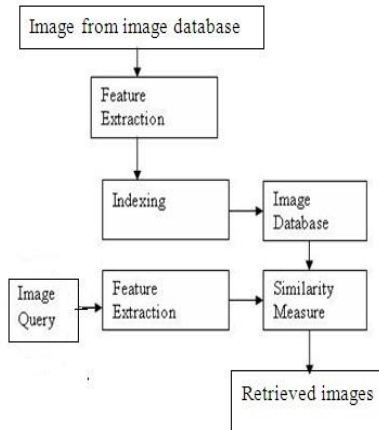


Fig.1 A typical content based image retrieval [5]

2.3 Shape features

Shape feature plays a vital role in object detection and recognition. Object shape features provide robust and

efficient information of objects in order to identify and recognize image. Shape features are important in describing and differentiating the objects in an image. Shape features can be extracted from an image by using two kinds of methods: contour and regions. Contour based methods are normally used to extract the boundary features of an object shape. Region-based methods that rely on shape descriptors are normally able to extract both kinds of features: boundary and region. Region-based methods normally use a moment based theory such as Hu moments, Legendre moments and Zernike moments.

3. SYSTEM ARCHITECTURE

A content-based image retrieval approach is based on colour, texture and shape features and controlled by fuzzy heuristics and the architecture is shown in Fig.2. The objective of using three algorithms is to develop an integrated image retrieval approach capable of producing efficient results. For better results, the approach ensures that the retrieved images are highly relevant to the query image. When a user inputs an image query, the image retrieval approach extracts features based on colour, shape and texture by applying relevant algorithms. The extracted features are stored in a feature vector. Then a similarity measure based on Euclidean distance and a set of fuzzy rules are applied to produce results relevant to the image query.

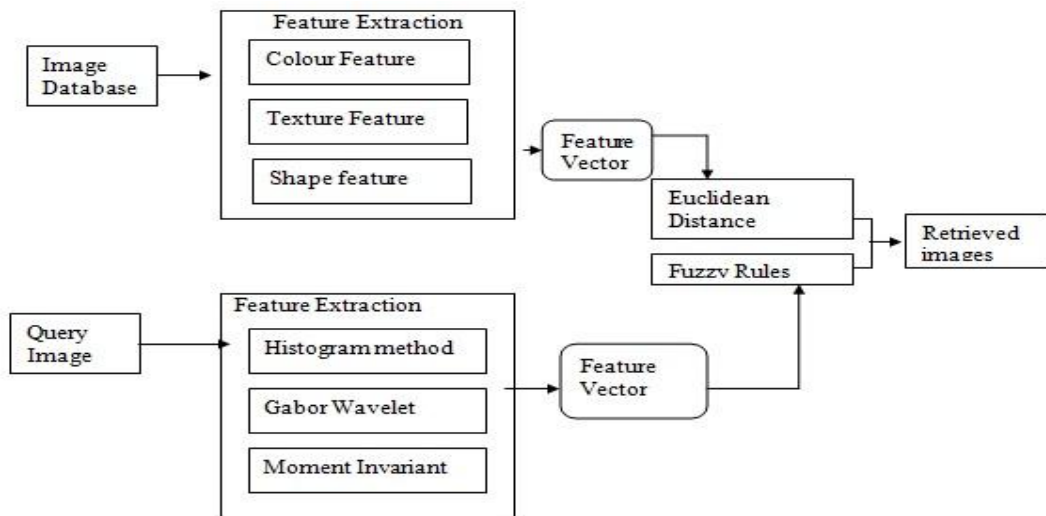


Fig.2 Content Based image retrieval architecture [5]

3.1 Histogram based method

For the colour feature, we integrate two types of histogram based methods using a colour image histogram and an intensity image histogram. For the colour image the RGB colour model that is based on the Red, Green and Blue components. An image histogram can be generated as follows:

$$h_b = \sum_{i=1}^M \sum_{j=1}^N \delta_b(i, j) \quad \forall_b = 0,1,2, \dots \quad (1)$$

where $\delta_b(i, j) = 1$ if the v at pixel location $[i, j]$ falls in b , and $\delta_b(i, j) = 0$ otherwise. Similarities between different histograms can be calculated using different methods such as Euclidean distance and histogram intersection as a similarity measure. Every pixel in an image is basically represented as a point in the colour

model such as RGB. This colour point is represented by three values that hold the information of colour. The image is represented by its histogram. The colour histogram helps to find the images which contain similar colour distribution. It is achieved by measuring the similarities through computing distance between two histograms.

3.2 Gabor Wavelet method

The second element is the texture feature. For this purpose, the Gabor wavelet algorithm is used.

The wavelet transformation provides a multi-scale decomposition of an image data. The Gabor filter is normally used to capture energy at a certain scale and at a certain orientation. Scale and orientation are two most important and useful features that are used for texture analysis. The Gabor filter is also known as scale and rotation invariant. A 2D Gabor function consists of a sinusoidal plane wave of some orientation and frequency, modulated by a 2D Gaussian. The Gabor filter in spatial domain is denoted as:

$$g, \lambda, \theta, \psi, \sigma, \gamma(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (2)$$

Where,

$$x' = x \cos(\theta) + y \sin(\theta)$$

$$y' = y \cos(\theta) - x \sin(\theta)$$

wavelength of cosine factor is represented by λ ; θ represents the orientation of the normal to parallel stripes of a Gabor function in the degree; the phase offset in degree is represented by Ψ ; the spatial aspect ratio which specifies the elliptically of the support of the Gabor function is represented by γ ; and σ is the standard deviation of the Gaussian that determines the linear size of the receptive field. When an image is processed by Gabor filter; the output is the convolution of the image $I(x, y)$ with the Gabor function $g(x, y)$ which is

$$r(x, y) = I(x, y) * g(x, y) \quad (3)$$

where $*$ represents the 2D convolution. The process can be performed at various orientation and scale; and prepared filter bank. To generate the filter bank, different scale and orientation parameters help to cover the entire spatial frequency space to capture mostly texture information with filter design. After applying Gabor filters on the image by orientation and scale, to obtain an array of magnitudes

$$E(m, n) = \sum_x \sum_y |G_{mn}(x, y)|$$

$m=0,1,\dots,M-1; n=0,1,\dots,N-1$

The magnitudes represent the energy content at different orientation and scale of image. The main purpose of texture based retrieval is to find images or regions with similar texture. The following mean μ_{mn} and standard deviation σ_{mn} of the magnitude of the transformed coefficients are used to represent the texture feature of the region:

$$\mu_{mn} = E(m, n) / P * Q$$

$$\sigma_{mn} = \sqrt{\sum_x \sum_y (|G_{mn}(x, y)| - \mu_{MN})^2 / P * Q} \quad (4)$$

where M represents the scale and N represents the orientation. The feature vector that represents the texture features is created using mean μ_{mn} and standard deviation σ_{mn} as feature components and these components are saved into two feature vectors and then these two vectors are combined in order to make the single feature vector that will be treated as an image texture descriptor.

3.3 Moment Invariant Method

The third main element is shape feature. In this approach Hu moment invariant algorithm is used. The Hu moment invariants algorithm is known as one of the most successful techniques for extracting image features for object recognition application. The 2D moment of order $(p + q)$ of a digital image $f(x, y)$ is defined as :

$$m_{p,q} = \sum_x \sum_y x^p y^q f(x, y) \quad (5)$$

for $p, q = 0, 1, 2$ where the summations are over the values of the spatial coordinates x and y spanning the image. The corresponding central moment is defined as:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (6)$$

$$\bar{x} = m_{10} / m_{00}$$

$$\bar{y} = m_{01} / m_{00}$$

and x, y are called the centre of the region. Hence the central moments of order up to 3 can be computed as:

$$\mu_{0,0} = m_{0,0}$$

$$\mu_{1,0} = 0$$

$$\mu_{1,1} = 0$$

$$\mu_{0,1} = 0$$

$$\mu_{1,1} = m_{1,1} - \bar{y}m_{1,0}$$

$$\mu_{2,0} = m_{2,0} - \bar{x}^2m_{1,0}$$

$$\mu_{0,2} = m_{0,2} - \bar{y}^2m_{0,1}$$

$$\mu_{3,0} = m_{3,0} - 3\bar{x}m_{2,0} + 2m_{1,0}\bar{x}^2$$

$$\mu_{2,1} = m_{2,1} - 2\bar{x}m_{1,1} - \bar{y}m_{2,0} + 2\bar{x}^2m_{0,1}$$

$$\mu_{1,2} = m_{1,2} - 2\bar{y}m_{1,1} - \bar{x}m_{0,2} + 2\bar{y}^2m_{0,1}$$

$$\mu_{0,3} = m_{0,3} - 3\bar{y}m_{0,2} + 2\bar{y}^2m_{0,1}$$

The normalized central moment of order (p +q) is defined as

$$\eta_{p,q} = \mu_{p,q} / \mu_{0,0}^Y$$

Φ_1 – Φ_6 moments are scaling, rotation and translation invariants and the Φ_7 moment is skew invariant which enables it to differentiate the mirror images. Φ_1 – Φ_7 moments are used to calculate the feature vectors:

$$\phi_1 = \mu_{20} + \mu_{02}$$

$$\phi_2 = (\mu_{20} + \mu_{02})^2 + (4\mu_{11})^2$$

$$\phi_3 = (\mu_{30} + 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} - \mu_{03})^2$$

$$\phi_5 = (\mu_{30} + 3\mu_{12}) + (\mu_{30} + \mu_{12}) [(\mu_{30} + 3\mu_{12})^2 + 3(\mu_{21} - \mu_{03})^2]$$

$$\phi_6 = (\mu_{20} - \mu_{02}) [(\mu_{30} + 3\mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})$$

$$\Phi_7 = \mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) + (\mu_{30} + \mu_{12}) [(\mu_{30} + 3\mu_{12})^2 + 3(\mu_{21} - \mu_{03})^2]$$

This set of normalized central moment is invariant to translation, rotation and scale changes in an image.

3.4. Euclidean Distance Method

Similarity between two images is measured numerically that reflects the strength of connections between them. Euclidean distance is used to calculate the similarity between two feature vectors as:

$$ED(M^k, M^t) = \sqrt{\sum_{i=1}^n (M_i^k - M_i^t)^2} \tag{7}$$

Where M^k and M^t are image query and image database respectively, i is a feature range. Closer distance represents the higher similarity between images.

3.5 Fuzzy similarity measure

Fuzzy heuristics is used to measure similarity between the query image and the database images in order to retrieve and display relevant or similar results to the user query. Three types of preferences are taken; the first priority is given to the shape features, as it is not affected by external factors, and invariant to the rotation, translation and orientation. The second priority is given to the colour features, as it is invariant to the rotation and translation. The third priority is given to the texture features. The Mamdani fuzzy inference method is used to perform fuzzy rules. After obtaining the relevant images to the query image, the common images between X, Y and Z set of images can be retrieved. The common set of images is considered the most relevant images. Commonality is measured using the below criteria.

X = Shape features are used to calculate the distance between query and database image

Y = Colour features are used to calculate the distance between query and database image.

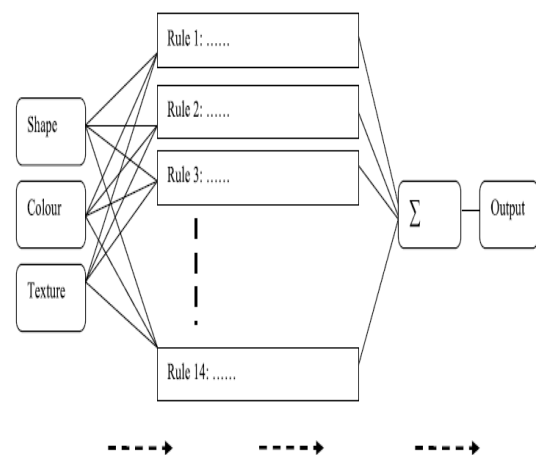


Fig.3 Fuzzy System Architecture [5]

Z = Texture features are used to calculate the distance between query and database image

S = Image similarity

By adopting the steps, a set of fuzzy rules to process the results achieved by applying the three distinct algorithms.

- Step 1: Define a number of inputs. In this case three inputs are used such as shape distance, colour distance and texture distance between query image and database images.
- Step 2: The membership functions for three types of input have been defined. There are three different types of fuzzy set that identified each input as low, medium and high.
- Step 3: Three types of output fuzzy sets have been declared such as high similar, medium similar and low similar.
- Step 4: A fuzzy rule can be defined as a conditional statement such as if then. Fuzzy rules applied using logical operator.
- Step 5: To process the Mamdani fuzzy inference method the crisp inputs are taken and fuzzify them to determine the degree to which these inputs belong to each of the appropriate fuzzy set.
- Step 6: Apply the AND fuzzy operator to get one number that represents the result of antecedent of rules. The output is a single truth value.
- Step 7: The process of unification of the output of all the rules that have been used until last step. The output of this step is one fuzzy set for each output variable. The process is called an aggregation.
- Step 8: Lastly the aggregate output fuzzy set should transform to a single crisp number. Then process of defuzzification is to done it.

4. EXPERIMENTAL RESULTS

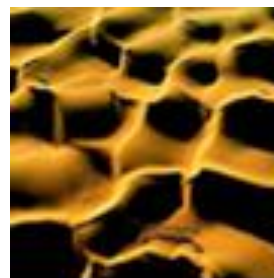
The colour feature is extracted and the feature vector for the image is:



The RGB value for the image is:

Red : 115.29
Green: 86.36
Blue : 63.50

The texture feature is extracted and the feature vector for the image is:



Mean : 20.0635
Standard deviation: 21.98975

The shape feature is extracted for the image is:



Invariant Moment 1	0.184008285283884
Invariant Moment 2	0.0122793899446014
Invariant Moment 3	0.00598437108339944
Invariant Moment 4	0.00013626555808644
Invariant Moment 5	7.83015755727509
Invariant Moment 6	1.13814777461541
Invariant Moment 7	2.68245961059209

The similarity between two images is calculated by using Euclidean distance:



Distance = 0.01876.

The precision and recall is calculated for the evaluation of the result based on the following:

$$\text{Precision} = I / V$$

I is number of images retrieved that are similar to the query image

V is total number of images retrieved

$$\text{Recall} = I / R$$

I is number of images retrieved that are similar to the query image

R is the total number of images in the database that are similar to the query image

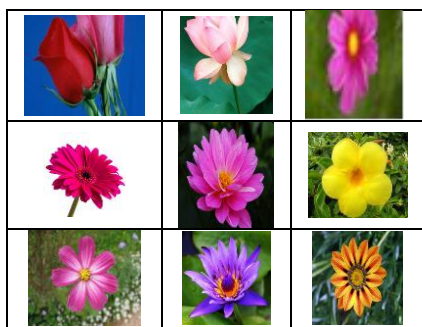


Fig.4 Image Retrieved results with precision and recall

The precision and recall value for the query image is

$$\text{Precision} = 16 / 28 = 0.5$$

$$\text{Recall} = 16 / 50 = 0.3$$

Precision and recall values for various images

Images	Precision	Recall
Image 1	0.5	0.3
Image 2	0.7	0.4
Image 3	0.6	0.33
Image 4	0.53	0.3
Image 5	0.60	0.34
Image 6	0.85	0.48
Image 7	0.67	0.38
Image 8	0.9	0.52
Image 9	0.35	0.2

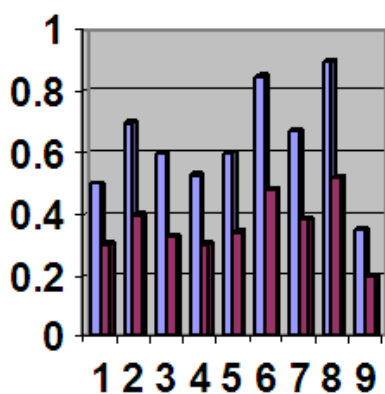


Fig.5 Precision and recall on the obtained result

5. CONCLUSION

Image retrieval approach which is based on colour, texture and shape features, controlled by fuzzy heuristics is used. The approach is based on the three well known algorithms: colour histogram, texture and moment invariants. The use of these three algorithms ensures that the image retrieval approach produce

results which are highly relevant to the content of query image, by taking into account the three distinct features of the image and similarity metrics based on Euclidean measure. The colour histogram is used to extract the colour features of an image using four components such as Red, Green, Blue and Intensity. The Gabor filter is used to extract the texture features and the Hu moment invariant is used to extract the shape features of an image. The evaluation is carried out using the standard Precision and Recall measures, and the results obtained are compared with the image query. The work based on space relationship will be further analyzed in the future enhancement.

REFERENCE

- [1] A.K. Jain, A. Ross and S. Prabhakar, An introduction to biometric recognition, *IEEE Trans. Circuits Syst. Video Technol.*, 14 (1), 2004, 4–20.
- [2] K. Iqbal, M.O. Odetayo and A. James, Integrated image enhancement method for biometric security, *Proc. The 2011 IEEE International Conference on Systems, Man, and Cybernetics*, Anchorage, Alaska, 2011.
- [3] K. Ericsson and A. Lehmann, Expert and exceptional performance: Evidence of maximal adaptation to task constraints, *Annu. Rev. Psychol.*, 47 (1), 1996, 273–305.
- [4] M. Smids, *Background subtraction for urban traffic monitoring using Webcams*, Master thesis, Universiteit Van Amsterdam, FNWI, 2006.
- [5] K. Iqbal, M.O. Odetayo and A. James, Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics, *Journal of Computer and System Sciences*, 78, 2012, 1258–1277
- [6] S. Jeong, C.S. Won and R.M. Gray, Image retrieval using color histograms generated by Gauss mixture vector quantization, *Comput. Vision Image Underst.* 94 (1–3), 2004, 44–66.
- [7] R. Mudigoudar, S. Bagal, Z. Yue, P. Lakshmi and P. Topiwala, Video super-resolution: From Qvga to HD in real-time, *Proceedings of the SPIE*, vol. 7443, 2009, 74430W1–74430W12.
- [8] T. Chaira and A.K. Ray, Fuzzy measures for color image retrieval, *Fuzzy Sets and Systems*, 150, 2005, 545–560.
- [9] T. Acharya and A.K. Ray, *Image Processing Principles and Application*, John Wiley & Sons, Inc., 2005.