Content Based Image Retrieval using Shape Features Extracted with Morphological Transformation and Block Truncation Coding

Pooja R Bhat
Dept of Computer Science and Engineering, B.V.B College Of Engineering and Technology,Hubballi,India
Email: poojabhatk1105@gmail.com

***Abstract*—In this paper, a Content Based Image Retrieval system is developed using a combination of texture, colour and shape features with the focus being the extraction of shape feature. Edge detection is carried out using morphological operations that include top hat and bottom hat transformation and block truncation coding. The obtained edge map is then enhanced using slope magnitude method in order to obtain connected boundaries. The shape feature is then combined with the texture and colour features. The features are extracted for both the database images and the query image. Similarity matching is then carried out between the database images and the query image and the most similar images are retrieved and displayed. The proposed Image retrieval system is tested across all the 11 categories of the Wang's dataset containing 1000 images. Performance evaluation of the system is conducted using Average precision and recall and a Precision-Recall Graph is generated for performance analysis.**

***Index Terms*—Image Retrieval, Edge detection, Shape feature, Block truncation coding, Morphological operations**

I. INTRODUCTION

Images play a very important part in human communication as they are illustrative, informative and easily understandable by everyone. With the advent of digital technology, the process of digitising images has become pervasive. Thus a large number of images are generated constantly everyday and add to the ever increasing database of images on the internet. This gives rise to the problem of searching for a given image among billions of images that exist on the digital platform. The image retrieval system has rapidly grown and evolved to meet this problem. Traditional methods for image retrieval included the method of adding keywords,captioning or some other metadata to the image and storing them in the databases. In these methods, a user who wishes to retrieve an image would enter the keyword of the desired image in terms of text and the retrieval would be performed using a text based search. The drawback of text based query was that annotating every image in a very large database with a different keyword was a very time consuming process and requiring a domain specialist with intricate knowledge of how the image would be stored. Also text descriptions are not entirely accurate in describing the complicated features of all images. Difference in languages of users also further complicate this image retrieval system. The content based image retrieval system came into focus to address the disadvantages of a text based image query system.

Content Based Image Retrieval (CBIR), also known as Query By Image Context (QBIC) was developed to focus on the inbuilt features of an image and to use these features to retrieve an image. The system worked on the concept that features of the images in the database from which the images were to be retrieved would be extracted and stored first. When a query image was presented, it's features would be extracted using the same methods. The features of the query image would then be matched with the already extracted features of the images in the database using a similarity measure and the most similar images would be then retrieved and displayed [1]. The process can be best illustrated with the following diagram

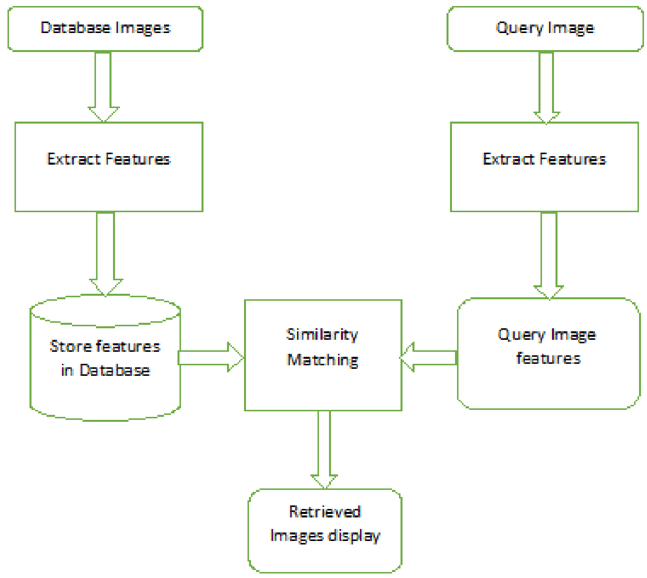


Fig. 1: CBIR overview

Colour, Texture and Shape are the basic descriptors or features of an image. Hence, a weighted combination of these three features is necessary for optimal image retrieval in any CBIR system.

A. Colour

Colour is an important feature for image retrieval as it is the most obvious differentiator in any image.In general, humans are not affected by small changes in colour in an image as much as they are affected by variations in gray levels. RGB, HSV and LAB are some of the most commonly used colour space representations. Irrespective of which colour space is being used, any colour information can be presented as either a single three-dimensional(3D) histogram or three separate one-dimensional(1D) histograms. It is essentially invariant to rotation and translation of the input image. In the proposed system, the normalized histogram method is used and the colour feature vector is extracted and stored [2]. The normalized histogram can be defined as

$$I(i) = \frac{H(i)}{\sum H(i)} \tag{1}$$

where *H(i)* is the histogram of a given image and index *i* stands for the histogram bin.

B. Texture

Texture in an image can be described as a property that any surface can possess containing descriptions of visual patterns, each having properties of homogeneity. It contains important information about the structural arrangement of the surface. It also describes the relationship of the surface to the surrounding environment. In short, it is a feature that describes the distinctive physical composition of a surface [3]. In the proposed system, the texture feature is extracted using wavelet features. In wavelet transform, a function known as wavelet is used which is essentially a waveform that is bounded both in frequency and duration. Examples of wavelet includes Coiflet, Mexican hat, Morlet and Haar. Discrete wavelet transform converts a given image into 4 parts : horizontal, vertical, diagonal and one approximate representative image. Then, mean and standard deviation of these components are taken and static textural moments are extracted and stores as texture feature vector [4].

C. Shape

Shape does not actually refer to the shape of the image but rather the shape of the region being sought out. Several applications in the areas of CAD/CAM and computer graphics require storage and access to large databases. A major data type stored and managed by these applications is representation of 2D objects.Hence it becomes imperative to develop an efficient way to detect images on the basis of shape feature. The most important aspect of shape feature extraction is the edge detection and/or edge segmentation. The process of Edge detection includes identifying and locating sharp discontinuities in the image. The discontinuity are sudden changes in pixel intensity which are indicators of boundaries of objects in the image [5]. Some of the traditional edge detectors include

- Gradient Based Edge detection : Roberts cross Edge Detector, Sobel Edge Detector, Prewitt Edge Detector are some of the most commonly used detectors.
- Laplace Of Gaussian : The Marr-Hildreth Edge Detector and Canny Edge Detector are the most common ones. Canny is the most widely used optimal edge detector in current times.

The disadvantages of the above listed edge detection methods are several including but not limited to sensitivity of noise which tends to degrade the magnitude of edges, diffraction due to existing edges in noisy images and complex gradient calculation leading to excessive time consumption. This calls for a simpler and efficient edge detection which is proposed in Section III

The rest of the paper is organized as follows. Section II discusses the literature that is relevant to the proposed work. Section III proposes a Content Based Image retrieval system designed with the primary focus being on the extraction of shape feature. Section IV introduces the experimental environment and also the evaluation of the selected parameters and concludes with the analysis of the results obtained with the proposed CBIR system. Section V concludes the report and also provides a future scope for the proposed system.

II. Literature Survey

A large number of Edge detection methods have been proposed in literature as it is the most fundamental operation performed during image analysis. An edge usually defines the boundaries between regions in an image which helps in recognizing the objects in an image as well as segmentation. The basic problem with general edge detectors is that their performance is usually not upto the mark. Adaptability of these detectors to different conditions is also poor. The performance of any edge detection method depends on the lighting situation of the image, the presence of objects with similar intensities, noise factor and the density of the edges in the image [6]. Each of the problem listed above can be solved by making some minor adjustments in the edge detection method or by reevaluating the threshold set for the edge, the fact still remains that this process is usually situation dependent and hence has to be performed manually. There is no absolute edge detection method that perform these tasks automatically.

One of the earliest popular edge detection method used was the Marr-Hildreth Edge detection technique. It is a gradient based operator which uses the Laplacian to take the second derivative of an image [6] [7]. This detector was then replaced by the Canny algorithm which is now considered to be the standard edge detector worldwide. The edge detection problem was visualized as a signal processing optimization problem and an objective function was developed to be optimized. The solution to this problem was a rather complex exponential function, but the algorithm figured out several ways to approximate and optimize the edge-searching problem [6] [8].Several other classic techniques like Robert's cross, Sobel and Prewitt edge detectors have also been utilised for the edge detection problem.

In recent years, research has forayed into new and non-conservative methods to identify and detect edges in an image. [9] discusses the application of fuzzy logic in edge detection. In this algorithm, using three $3x3$ spatial filters i.e., low-pass, high-pass and Sobel filters, through spatial convolution process, edginess of each pixel is determined. The edge values derived from these three masks then form the fuzzy input. Utilizing the membership functions and fuzzy rules, the pixel is then ruled to be either edge or non-edge. Similarly, a fuzzy logic based relative pixel algorithm is implemented by [10]. A set of 9 pixels, forming a $3x3$ window is subjected to fuzzy conditions that highlight the edges associated with the image. The image is scanned exhaustively using the relative pixel window and thus all the edges are obtained with the help of fuzzy rules.

Another technique involving fuzzy rules is described in [11] where in a fuzzy edge map is used. Fuzzy compactness vector is computed from fuzzy edge map thresholded at different levels of the unsegmented image, which also incorporates gray level contrast information embedded in the edges. [12] proposed a contour based approach for shape retrieval. Using a edge map provided by Canny operator,the contour of an image is traced and re-sampling is done to avoid any discontinuity in the contours. The result is then swept line by line and the neighbour of every pixel is explored to detect the number of surrounding points and to derive the shape feature vector. A novel approach was identified in [13] where in a shape matrix and snake model is used for representation and comparison of the shapes present in an image. The shape matrix is invariant and unique under rigid motions. The original template shape is deformed to adjust itself to shape images. The degree of deformation is used to find the similar between the shape images.

A novel method for edge extraction is discussed in [14] using block truncation coding and convolution measures. This method is primarily used for MRI images. A similar approach is utilised in [15]. Block truncation coding is combines with morphological operations to enhance the performance of the shape retrieval. Morphological operators as edge detectors is also discussed in [5] which elaborates on the advantages of using Morphological operations over the traditional edge detectors.

The following section III discusses the proposed system in detail.

III. Proposed work

The proposed CBIR system uses colour, texture and shape feature vectors for image retrieval. The colour and texture feature retrieval is already described in Section I. In this section, the shape feature retrieval is discussed with an emphasis on the edge feature extraction using Morphological operators and Block Truncation Coding. The morphological operators are used in Top hat and Bottom Hat Transformation.

A. Morphological Operations

Morphology refers to the group of image processing operations based on shape. Typically, a structuring element is applied to the input image to obtain the output image. The value of each pixel in the output image is dependent on not just the corresponding pixel in the input image but also the neighbouring pixels. So essentially, by choosing an optimal size and shape of the neighbourhood, a morphological image can be constructed. The basic mathematical morphological operators are Dilation and Erosion. Dilation is the process of adding pixels to an input image to create a swelling effect. Erosion is the process of shrinking an image by removing pixels from the object boundaries to create a sharper image [15].

1) Top Hat transformation: Top Hat transformation is a morphological process involves the opening operation. Opening operation is essentially the erosion process followed by the dilation process. The resultant image is called the opened image. The opened image is then subtracted from the original image to obtain the top hat image.

2) Bottom Hat transformation: Bottom Hat transformation also known as White Top Hat Transformation represents the complement of the top hat transformation. This uses the Closing operation where in the dilation process is first applied followed by the erosion process. The image thus obtained is called a Closed image. This image is then subtracted from the original image to produce a bottom hat image.

B. Block Truncation Coding

The images obtained from the morphological transformations mentioned above is subjected to block truncation coding (BTC).Block truncation coding is a simple and fast method for digital image compression. This method can also be extended to extract edges in an image. The image is first divided into blocks of $4x4$ which are non overlapping and each block is coded one at a time. Each block is evaluated and a mean of the whole block is found. If a pixel within a block is greater than or equal to the mean value, the corresponding pixel value in the output bit plane is assigned as 1. If a pixel is found to have a lesser value than the block mean, the corresponding pixel is assigned a value of 0 in the bit plane. This process is then repeated for all the blocks. The resultant bit planes are then combined to form a binary image which is further enhances using slope-magnitude convolution method.

C. Slope - Magnitude convolutions

The above mentioned process usually gives good results but there is a major drawback in that the edges are detected only in either horizontal or vertical directions and sometimes boundaries are disconnected. In order to obtain more connected boundaries in order to reflect the boundaries of the objects and shapes in the input image,slope magnitude method is used. This method can also be paired up with the traditional gradient operators like Sobel, Prewitt and Canny [15].The first step is to convolve the image with the G_x mask to get the X gradient of the image and the G_y mask to get the Y gradient. The G_x and G_y masks used are as follows

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Then the slope image is found by the following formula

$$Slope\_image = \sqrt{G_x^2 + G_y^2} \quad (2)$$

D. Similarity Matching

Similarity Matching is carried out using Canberra Distance which is a numerical measure of the distance between pairs of points in vector space. It is defined as

$$c\_dist = \sum \frac{|x - y|}{|x| + |y|} \quad (3)$$

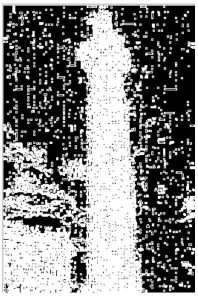
The Canberra distance is applied to all three feature vectors i.e., Texture, colour and shape. The three feature vectors are then combined by a weight which is assigned as $Texture = 0.2$,$Shape = 0.6$ and $Colour = 0.2$ and then the most similar images are retrieved and displayed.

IV. Result Analysis

The proposed system is implemented using MATLAB 7.12.0.365(R2011a) on $Intel$ Core i3 CPU M370(2.4 GHz) with 3 GB of RAM. Wang's database containing 1000 images spread over 11 categories is used as the test database. Figure 2 illustrates the results when the implemented system is tested against each category with the Edge Map and Slope Image displayed with the Original Image.

(a) Original Image

(b) Edge Map



(c) Slope Image

(d) Original Image

(e) Edge Map

(f) Slope Image

(g) Original Image

(h) Edge Map

(i) Slope Image

Fig. 2: Intermediate Results

Figure 3 depicts the CBIR system output. 20 images are fixed as the output size. An image from each category is executed on the implemented system and the results are obtained.

All the categories in the database of images are executed on the implemented system with varying input image and size of output image set. Top N images are charted against the number of iterations carried out. The system is run taking an image from each category as input one at a time and increasing the size of the output set in steps of 10. When the system is run taking a Bus image as input, in the first iteration of output size 10, 9 relevant images are returned. Whereas, in output size of 20, 14 relevant images are returned and so on. Each image category has 100 relevant images in the database. The following Figure 4 is the graph charted from the data obtained from the above mentioned operation. The X_Axis indicates the number of iterations carried out where the output set is varied by a factor of 10 and the Y_Axis indicates the Top N results obtained. Each of the different coloured lines indicate a different category in the Wang's dataset.

(a) Output 1

(b) Output 2

Fig. 3: System output

A. Precision and Recall Analysis

The Figure 4 graph can be used to determine the Precision and Recall of the system. Here Precision can be defined as

$$Precision = \frac{No.of Relevant Images Retrieved}{No.of Images Retrieved} \quad (4)$$

Recall can be similarly defined as

$$Recall = \frac{No.of Relevant Images}{No.of Relevant Images in the Database} \quad (5)$$

To give a graphical impression of the performance of a retrieval method, Precision Recall graphs can be created. To generate such a graph, a single query is repeatedly executed and the number of returned results is varied. For each of these results sets, the precision and recall are determined and both these values are plotted as a single coordinate in the graph. In an ideal Precision-Recall Graph (PR graph) , the linear relation

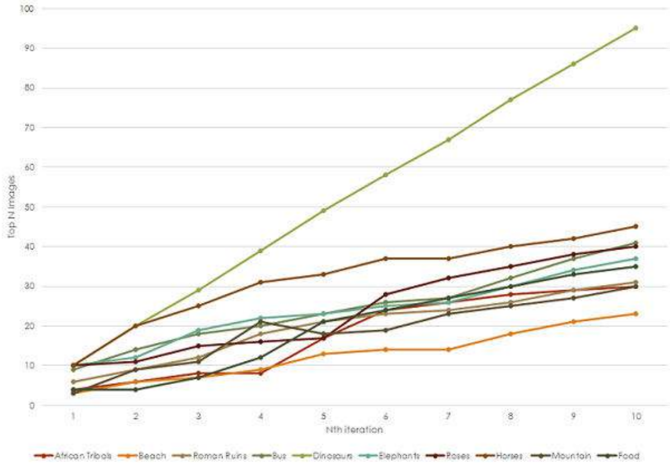


Fig. 4: Top N images over N iterations

between precision and recall indicates that with increasing recall, precision decreases until the point where all relevant images are retrieved.

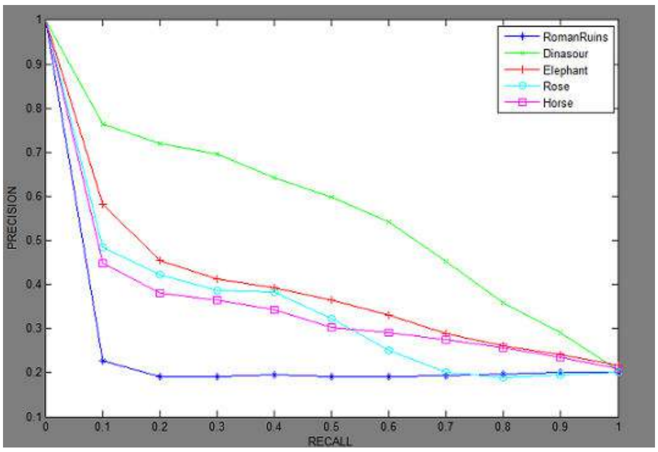


Fig. 5: Precision Vs Recall Graph

The Precision-Recall Graph depicted in the above figure 5 shows higher precision values for lower recall values. This indicates that for shorter result sets, the precision is very high but as the number of retrieved images increases, the precision decreases. This means that most of the relevant images are returned in the top of the result set whereas some of the relevant images are not detected by the CBIR system and are mixed with the rest of the results. The categories of Dinosaurs has the best performance in this retrieval method with the average precision = 0.9699. The categories of Horse,Bus,Elephants and Roses also perform well with the average precision being 0.68, 0.51, 0.51 and 0.49 respectively.The remaining categories have a relatively lower precision as compared to the above mentioned categories.

V. Conclusion

The Content based Image Retrieval System was successfully implemented with the extraction of Colour, texture and shape features. The novel approach of combining the Morphological operations with Block Truncation coding has produced good results in 7 categories of the standard dataset used. The performance of an image retrieval system is intrinsically linked with the speed of the retrieval. The implemented system reduces the retrieval time to below 15 seconds for every retrieval. Further improvements can be made by reducing the number of features extracted and thus reducing the complexity of the algorithm to a considerable extent. Also, some more improvements can be made to enhance the performance of the categories that did not perform well under the implemented algorithm.

References

[1] F. Joshi, "Analysis of existing cbir systems: improvements and validation using color features," *International journal of Emerging technology and advanced engineering*, vol. 3, no. 5, 2013.

[2] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern recognition*, vol. 29, no. 8, pp. 1233–1244, 1996.

[3] K. N. B. M. TECH, "Image retieval color, shape and texture features using content based," *Image*, vol. 2, no. 9, pp. 4278–4287, 2010.

[4] P. P. Buch, M. V. Vaghasia, and S. M. Machchhar, "Comparative analysis of content based image retrieval using both color and texture," in *Engineering (NUiCONE), 2011 Nirma University International Conference on*. IEEE, 2011, pp. 1–4.

[5] D. Rana and S. Dalai, "Review on traditional methods of edge detection to morphological based techniques," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 4, pp. 5915–5920, 2014.

[6] E. Nadernejad, S. Sharifzadeh, and H. Hassanpour, "Edge detection techniques: Evaluations and comparison," *Applied Mathematical Sciences*, vol. 2, no. 31, pp. 1507–1520, 2008.

[7] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.

[8] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[9] D. Aborisade, "Fuzzy logic based digital image edge detection," *Global Journal of Computer Science and Technology*, vol. 10, no. 14, 2010.

[10] S. Mathur and A. Ahlawat, "Application of fuzzy logic on image edge detection," 2008.

[11] M. Banerjee and M. K. Kundu, "Content based image retrieval with fuzzy geometrical features," in *Fuzzy Systems, 2003. FUZZ'03. The 12th IEEE International Conference on*, vol. 2. IEEE, 2003, pp. 932–937.

[12] S. Arivazhagan, L. Ganesan, and S. Selvanidhyananthan, "Image retrieval using shape feature," *International journal of imaging science and engineering (IJISE)*, vol. 1, no. 3, pp. 101–103, 2007.

[13] C. Sheng and Y. Xin, "Shape-based image retrieval using shape matrix," *International Journal of Signal Processing*, vol. 1, pp. 163–167, 2005.

[14] K. Somasundaram and V. Suresh, "A novel method for edge detection using block truncation coding and convolution technique for magnetic resonance images (mri) with performance measures," *IJITR*, pp. 42–45, 2015.

[15] H. Kekre, S. Thepade, P. Mukherjee, M. Kakaiya, S. Wadhwa, and S. Singh, "Image retrieval with shape features extracted using morphological operators with btc," *Image*, vol. 12, no. 3, 2010.