IMAGE RETRIEVAL BY SEMANTIC INDEXING

[1] CH.GANAPATHI REDDY, [2]Dr. G. R. BABU, [3]Dr. P. V. D.SOMASEKHAR

[1]HOD, Department of Electronics and communication Engineering, GNITS, Hyderabad, India-500008
[2]Professor & HOD, ECE KMIT, Hyderabad, India
[3]Professor, Department of Electronics and communication Engineering, JNTU, Hyderabad, India
E-mail: ganapathi7898@gmail.com

ABSTRACT

Image retrieval is analogous to conventional text-based querying of databases except that both the query and the retrieved data are in the form of images. For example, a query image of a mountain must retrieve images of mountains from a database of images. Such matches are done using features easily extracted from images such as colour, texture, shape and geometry. The major challenge facing the researchers is that humans measure similarities between images based on the objects found therein whereas a computer based system uses "low-level" features listed above. Humans, looking at an image of an elephant, would retrieve images of elephants from a database while a computer-based system would "look for" images containing gray regions of certain size and shape. The result is a serious mismatch between human expectation and the system performance and is called the "*semantic gap*" in literature.

Key words: *Image Retrieval, K-means algorithm, Connect component algorithm*

1.INTRODUCTION

Image database systems are being generated at an ever-increasing rate. The conventional image databases use an alphanumeric index for images retrieval. However, human beings are not used to retrieving images based on their alphanumeric indices. Recently, an integrated feature extraction and object recognition technique is being investigated by researchers to allow queries on large databases using example images, user-constructed sketches and drawings, colors and texture patterns, and other graphical information. The algorithms and techniques being developed in such a context have been
given the name of **content-based image retrieval (CBIR)**. In CBIR systems, the content of an image such as color, shape and texture is used to retrieve images that are similar to a query image. The user makes a request like "find pictures of dogs" or even "find pictures of Abraham Lincoln". This type of open-ended task is very difficult for computers to perform - pictures of Chihuahuas and Great Danes look very different, and Lincoln may not always be facing the camera or in the same pose. A human can retrieve these types of images looking at the images easily but it becomes difficult for a computer-based system.

2. MOTIVATION

Content-based image retrieval (CBIR) systems mainly perform extraction of visual features typically color, shape and texture as a set of uncorrelated characteristics. Such features provide a global description of images but fail to consider the semantics of an image. At a more abstract level of knowledge about the content of images, extraction of object descriptions and their relative positions provides a spatial configuration and a logical representation of images.

Because of the lack of low-level features extraction such methods generally fail to consider the physical extension of objects and their primitive features. An image retrieval system should perform similarity matching based on the representation of visual features conveying the content of segmented regions; besides, it should capture the spatial layout of the depicted scenes in order to face the user expectations.

3.SCOPE

In this work, we develop algorithms to assign high-level or semantic labels to different regions of an image, and to retrieve them based on the assigned labels. The labels are assigned in a semi-automatic or human-assisted way. Such a method enables our tool to retrieve images even if they differ in low-level features such as colour and size.

It is also invariant to rotation and translation of the image. All the images that we took are color images. It works well for gray scale images also. We extract images that do not have fine level features i.e. we do not use hierarchical level of classification images .For example we extract images of a flower but not to a very fine level of extracting rose or jasmine flowers.

In our project we have a database of about 1000 images. It consists of 5 categories. The categories are independent of each other. In future if required we can increase the images in the database, that can be done easily through the tool thus and it is directly added to the database and is included during retrieving next time.

The system is specific to these five categories (cars, currency, flowers, sports, waterfalls) presently. The querying image must be placed among the existing categories. If there are any ambiguous images they are placed to the closest category of images. The proposed system yet does not the ability to reject the images belonging to any of the categories. For adding more categories the system must be remodeled. The input (query) to a system is an image or text and output is the confidence with which the system assigns the input to the given categories.

4. BLOCK DIAGRAM

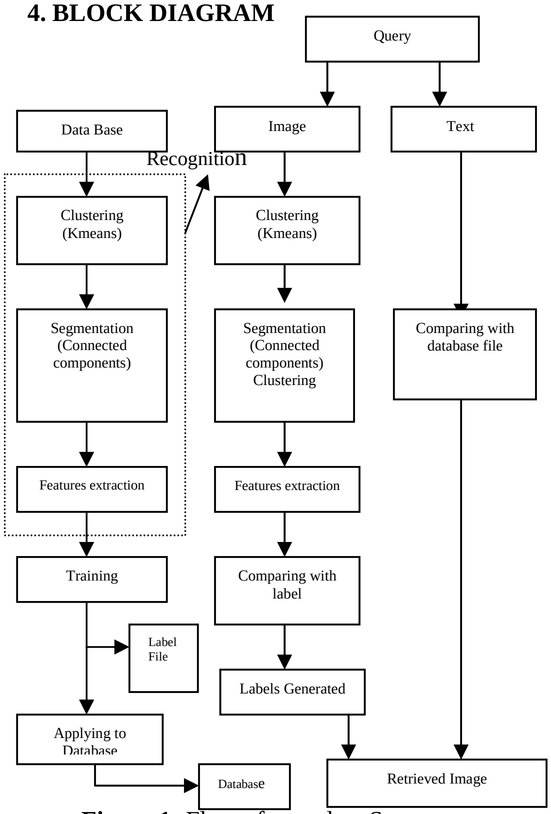


Figure 1: Flow of complete System.

5. CLUSTERING

Clustering is the classification of objects into different groups, or more precisely, the partitioning of the dataset into subsets (clusters) so the data in each subset (ideally) share some common trait – often proximity according to some defined distance measure. Clustering is used in many fields including machine learning, data mining, pattern recognition, image analysis and bioinformatics.

Kmeans is one of the simplest unsupervised learning algorithms that will solve the clustering problem

5.1 KMEANS

The Kmeans algorithm is an iterative technique that splits a large set of co-ordinates into groups called clusters. Its main objective is to minimize the average square distance between points in the same cluster. Although it offers no accuracy guarantee, its simplicity and speed are very appealing in practice. The Kmeans is a simple and fast algorithm that attempts to locally improve an arbitrary Kmeans clustering.
The algorithm goes in this way.

- Pick *K* cluster centers randomly
- Assign each pixel in the image to the cluster that minimizes the variance between the pixel and the cluster center
- Re-compute the cluster centers by averaging all of the pixels in the cluster
- Repeat steps 2 and 3 until convergence is attained
- The objective it tries to achieve is to minimize total intra-cluster variance or the squared error function.

After clustering the output will be.

Figure 2: Source Image for clustering.

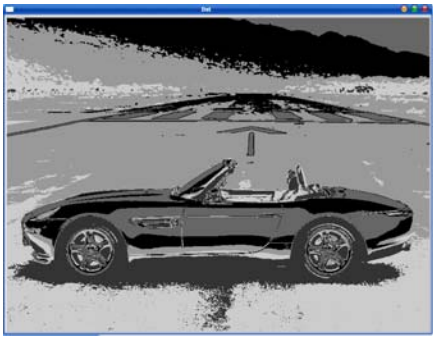


Figure 3: Output of clustering.

6. SEGMENTATION

In computer vision, segmentation refers to the process of partitioning a digital image into multiple regions (sets of pixels). Another way of extracting and representing information from an image is to group pixels together into regions of similarity. This process is commonly called segmentation. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. For this segmentation we have been using two pass connectivity algorithm .

6.1 CONNECTED COMPONENT ALGORITHM

Connected component labelling works by scanning an image, pixel by pixel (from top to bottom and left to right) in order to identify connected pixel regions i.e., regions of adjacent pixels that share some common features. Once all the groups have been determined, each pixel is labelled with a grey level. According to the component it was assigned.

Connected component labelling works on binary or grey level images and different measures of connectivity are possible. We used 8-connectivity. The connected components labelling operator scans the image by moving along a row until it comes to a point 'P' (where 'P' denotes the pixel to be labelled at any stage in the scanning process.) It examines the 4 neighbours of 'P' which has alredy been encountered in the scan (i.e., the neighbours (1) to the left of 'P' (2) above it and (3 and 4) the two upper diagonal terms). Based on this information point 'P' is labelled.

After completing the scan the equivalent labelling pairs are sorted into equivalence segments and unique label is assigned to each segment. In the first pass we get ambiguity at the corner pixel of '**V**' shaped segments.At this stage we come to know both the segments are same.So in the second pass we assign the lowest segment number to it.

The algorithm is as follows:

P1	P2	P3
P4	P	P5
P6	P7	P8

First pass - initial labelling

- If none of P1…..P4 are labeled, assign a new label.
- If any one of P1……P4 are labeled, assign same label.
- If more than one of P1……P4 are labelled and their labels are identical, assign same label.
- If more than one of P1…..P4 are labelled and their labels are different.

Second pass – renumbering and merging.

- Assign any of the labels.
- Mark different labels of P1……P4 are equivalent.
- Renumber equivalent labels on second pass.

Output of segmentation:

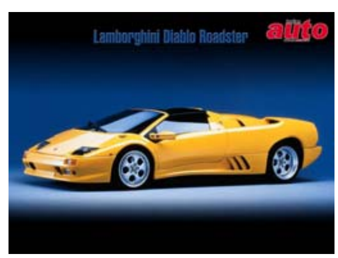


Figure 4: Source Image for segmentation.

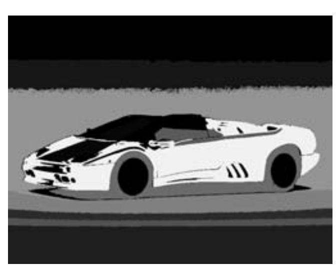


Figure 5: Output after segmentation.

7. FEATURES EXTRACTION

Features extraction plays an important role in image retrieval. It is a challenge to select a good feature set for classification. A selected feature is used as a discrimination power if its intra-class and the inter-class distances are large. In the above segmented image we extract topmost ten segments. We extract 8 features for each of the segments.

The features are

- **MBR (Minimum Bounding Rectangle):** The MBR is calculated using boundaries of the segment i.e. x_{min}, x_{max}, y_{min}, y_{max} for each segment in the database
- **Occupancy ratio**: It is defined as size to the area of the MBR.
  Occupancy ratio = size / (area of the MBR)
- **Average Colour values:** Average values of R, G and B components.
- **Variance values**: Variance of R, G and B values.

We normalize all the features in an image i.e. we bring all the features to a range between the interval [0, 1]. A normalized method is preferred over others because it gives equal weight age to others. Based on all these features training is given to the tool to recognize an image.

8. TRAINING PHASE

We take few images say 5 images from each category i.e. a total of 25 images and label the segments manually. We maintain files for each label. When number of labels for a particular label are very less (i.e. one or two) we mark them as miscellaneous and they are not taken in to account .We store all the above feature extracted values in to the corresponding label files. Based on the labels of all the segments in an image, label is given to the entire image. Some of the problems encountered by us while retrieving the images are as follows.

In the features the positions (corner values) of the Minimum Bounding Rectangle are included. Due to the retrieved images are not up to the mark, we got only (10-20) % images are relevant. The reason for this is the segment can be anywhere in the image. So this approach is not up to the mark.

In the next approach we omitted the positions. We have done the extraction of images based on the new results but the images retrieved are only (20-30) % relevant. This is because some individual values are low and some values are high so the irrelevancy came into picture. For example the values for ravg are 0.223, 0.888, and 0.454. The average value of above three is 0.521 .So the values may match another label. The average values of the color values are different because we work for different colour images. So this approach does not benefit much for different colour images.

We go for the next approach that is nearest neighboring approach. In this we do not average the values but we use all the labels and compare them with the label we want.

This approach is better than the approach described above. The efficiency of this approach is 80%.This is the best approach among the above three and we followed this approach

9. APPLYING TO DATABASE

Apply recognition to all the database images and extracted features for the images are compared with label files. We use nearest neighbour approach for measuring the similarity between the labels. By this we assign labels to the image using this label file and stored in a database file. The database file is stored as image name and segment numbers in a database file. Here the labels are given in a semi automatic manner. The database file is maintained as shown below:

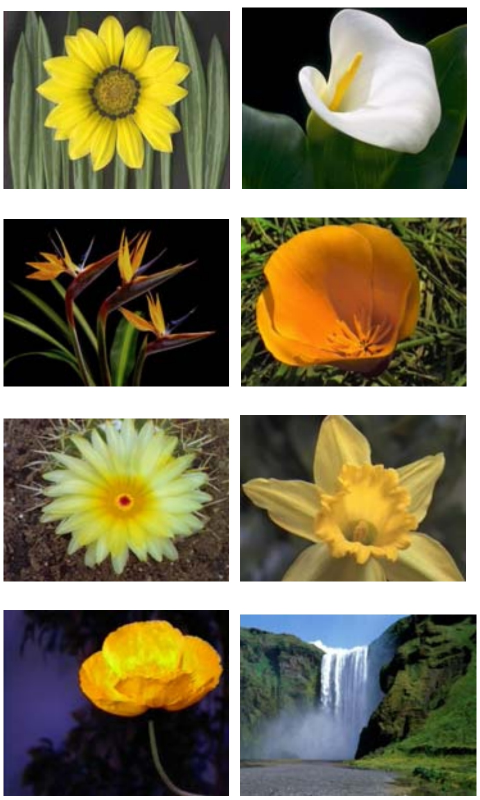
<Image 1 name><Number of labelled segments><Seg0 label No><Seg1 label No>...…………………<Seg9 label No>
Thus all the label numbers are maintained in the database.

10. QUERYING

By giving an image as query this tool will search with the existing set of labelled images (trained images).The querying image must be placed among the existing categories .If there are any ambiguous images they are placed to the closest category of images. The query may be in the form an image, text-based or combination of two or more categories. This is an automated stage.

If the query is in the form of an image we apply the recognition stage to our query image and then label the file based on the label file that we obtained from the manual labeling of the image in the training phase. The labels are compared with the labels in the database file and finally retrieve most similar top10 images.

11. RETRIEVED IMAGES



12. QUERY IMAGE

If the query is in the form of a text then we check out the labels in the database file and retrieve the top 10 images which mostly contain the label as desired. For example the query is a car then we check for the images which mainly

13. CONCLUSION

In this project we proposed an approach using high level features called as semantic indexing. From our results we can infer that this approach can retrieve images of different shape, colour and size. Secondly we can say that the images that occupied maximum area that images are retrieved accurate. Nothing in the world is perfect so our system has also some limitations for this approach. The approach does not work well for the images in which the segment size is small. Lastly the limitation of labeling is we cannot label the segments correctly.

REFERENCES:

[1]Sameer Antani, Rangachar Kasturi, and Ramesh Jain. A Survey on the Use of Pattern Recognition Methods for Abstraction, Indexing and Retrieval of Images and Video. *Pattern Recognition*,**35**:945–965, 2002.

[2] I. Daubechies. *Ten Lectures on Wavelets*. Capital City Press, 1992.

[3] J. M. Fuertes, M. Lucena, N. Perez de la Blanca, and J. Chamorro-Martinez. A Scheme of Colour Image Retrieval from Databases. *Pattern Recognition Letters*, **22**:323–337, 2001.

[4] Vito Di Gesu and Valery Starovoitov. Distance-Based Functions for Image Comparison. *Patter Recognition Letters*, **20**:207–214, 1999.

[5] Anil K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, New York, 1988.

[6] Mohan S. Kankanhalli, Babu M. Mehtre, and Hock Yiung Huang. Colour and Spatial Feature for Content-Based Image Retrieval. *Pattern Recognition Letters*, **20**:109–118, 1999.

[7] Jia Li, J. Z. Wang, and Gio Wiederhold. IRM: Integrated Region Matching for Image Retrieval. In *Proc. of 8th ACM Int. Conf. on Multimedia*, pages 147–156, Los Angeles, California, 2000.

[8] Mingjing Li, Zheng Chen, and Hong-Jiang Zhang. Statistical Correlation Analysis in

Image Retrieval. *Pattern Recognition*, **35**:2687–2693, 2002.

[9] G. Pass, R. Zabih, and J. Miller. Comparing Images Using Color Coherence Vectors. In *Proc. Of the ACM Multimedia'96 Conference*, pages 65–73, 1996.

[10] B. G. Prasad, S. K. Gupta, and K. K. Biswas. Region-Based Image Retrieval Using Integrated Color Shape and Location Index. *Computer Vision and Image Processing*, **94**:193–233, 2004.

[11] E. Di Sciascio, F. M. Donini, and M. Mongiello. Spatial Layout Representation for Query-by-Sketch Content-Based Image Retrieval. *Pattern Recognition Letters*, **23**:1599–1612, 2002.

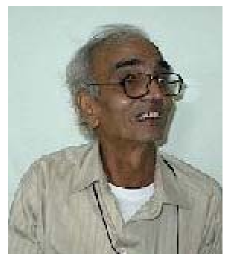
[12] Hamdi A. Taha. *Operations Research*. Prentice Hall, 1982.

[13] Kian-Lee Tan, Beng Chin Ooi, and Lay Foo Thiang. Indexing Shapes in Image Databases Using the Centroid-Radii Model. *Data and Knowledge Engineering*, **32**:271–289, 2000.

[14] Rafael C.Gonzalez,Richard E Woods Digital Image Processing *chapter 6 color image processing.*

BIOGRAPHY

Ch. Ganapathi Reddy received the degree in electronics and Communication Engineering, From RVR & JC College of Engineering, Nagarjuna University, Guntur, India, in 1989. He is a research student of JNTU Hyderabad. Currently, he is Professor & Head at ECE Department GNITS Hyderabad. His interests are in Image Processing, Signal Processing.



Dr.G.R. Babu received the Ph.D, DIC degree in Microwave Engineering and Microstrip Lines from he London University, London, in 1969. Currently, he is Professor & HOD at Keshaava Memorial College of Engineering, Hyderabad.

Dr.PVD Somasekhar Rao obtained his B.E. Degree in Electronics & Communication Engg. From Sri Venkateswara University College of Engg., Tirupati, and M.Tech. Degree in Microwave and Radar Engg., from IIT, Kharagpur. He earned his Ph.D. Degree in Electronics and Communication Engg., from I.I.T. Kharagpur. Currently working as Professor of ECE, JNTU Hyderabad, India.