

## Categorization and Searching of Color Images Using Mean Shift Algorithm

<sup>1\*</sup>Prakash PANDEY, <sup>2</sup>Uday Pratap SINGH and <sup>3</sup>Sanjeev JAIN

*Lakshmi Narain College of Technology, Bhopal (India)*

E-mails: <sup>1</sup>[prakash24pp@gmail.com](mailto:prakash24pp@gmail.com), <sup>2</sup>[usinghiitg@gmail.com](mailto:usinghiitg@gmail.com), <sup>3</sup>[dr\\_sanjeevjain@yahoo.co.in](mailto:dr_sanjeevjain@yahoo.co.in)

### Abstract

Now a day's Image Searching is still a challenging problem in content based image retrieval (CBIR) system. Most CBIR system operates on all images without pre-sorting the images. The image search result contains many unrelated image. The aim of this research is to propose a new object based indexing system Based on extracting salient region representative from the image, categorizing the image into different types and search images that are similar to given query images.

In our approach, the color features are extracted using the mean shift algorithm, a robust clustering technique, Dominant objects are obtained by performing region grouping of segmented thumbnails. The category for an image is generated automatically by analyzing the image for the presence of a dominant object. The images in the database are clustered based on region feature similarity using Euclidian distance. Placing an image into a category can help the user to navigate retrieval results more effectively. Extensive experimental results illustrate excellent performance.

### Keywords

Content-based, Segmentation, Category, Indexing, Searching.

### Introduction

With the advances in multimedia technologies and the increasing emphasis on multimedia applications, the production of image and video information has resulted in large

volumes of images and video clips. This trend is likely to continue, and to cope effectively with the explosion of multimedia information, image-based systems have been proposed to properly organize and manage this information for rapid retrievals [1]. The use of low-level visual features to retrieve relevant Information from image and video databases has drawn much research attention in recent years. Color is perhaps the most dominant and distinguishing visual feature. Tools available for searching an arbitrary image collection, such as in the Internet, are still far from satisfactory. This is because the range of images is wide and the content of the images is complex. Most well known Internet image searching tools (e.g. Google Image Search - <http://images.google.com>) use image filename as the primary means of indexing image attributes. This type of image indexing inevitably fails as it is based on the flawed assumption that image content is always reflected correctly by the image filename. The remainder of the paper is organized as follows. In Section II, architecture of the proposed system is discussed. Section III image segmentation of the proposed system, in section IV, category generation, section V, query processing, section VI and VII, concludes the paper and future extension of the present work.

### ***Previous Work***

To index an image using its content, there are currently three general approaches: object recognition, statistical analysis and image segmentation. Current CBIR systems such as IBM's QBIC [14], allow automatic retrieval based on simple characteristics and distribution of color, shape and texture. But they do not consider structural and spatial relationships and fail to capture meaningful contents of the image in general. Also the object identification is semi-automatic. The Chabot [15] project integrates a relational database with retrieval by color analysis. Textual meta-data along with color histograms form the main features used. Visual seek allows query by color and spatial layout of color regions. Text based tools for annotating images and searching is provided. Object recognition techniques are limited to specific domains, e.g., images containing simple geometric objects. This approach has been used to retrieve images of tools and CAD geometric objects [12] and also medical images [13]. For most other types of images, such as images containing people, sceneries, etc, Object recognition techniques are infeasible. To pursue the complexity of these images, researches then employed statistical indexing based on color and texture [16], [17]. Images can be retrieved by specifying a combination of RGB color values, textural measures, and more

recently using other features such as shapes and spatial relations between regions in the image [18], [5], [19]. A new image representation which uses the concept of localized coherent regions in color and texture space is presented [2], [3]. Images can be retrieved by specifying a combination of RGB color values, textural measures, and more recently using other features such as shapes and spatial relations between regions in the image [5]. Color is a low level feature, and by itself cannot adequately describe objects in images. To enable objects to be extracted and indexed within the image, image segmentation techniques are used [4]. However, segmentation results of general images are noisy and contain too many regions. Thus, this approach is still limited to simple objects, and thus for general images currently do not provide a meaningful object based representation.

### *Architecture of Proposed System*

Figure 1 shows architecture of a content-based image retrieval system using mean shift algorithm. Three main functionalities are supported: Data insertion, Categorization and Query processing. The data insertion subsystem is responsible for extracting appropriate features from images results from this will be classified into four different general and two semantic categories and storing them into the image database. This process is performed off-line. The query processing, intern, is organized as follows: the interface allows a user to specify a query by means of a query pattern and to visualize the retrieved similar images. The query-processing module extracts a feature vector from a query pattern and applies a metric as the Euclidean distance to evaluate the similarity between the query image and the database images. Next, it ranks the database images in a decreasing order of similarity to the query image and forwards the most similar images to the interface module. The database images are indexed according to their feature vectors to speed up retrieval and similarity computation. Note that both the data insertion and the query processing functionalities use the feature vector extraction module.

The dialog box shows the features extracted from an image and the measured category for that image. The category is obtained automatically by analyzing the composition of color, texture and structure from the main regions. The regions are produced using the perception-based image segmentation system [6]. By implementing perceptual grouping, the results achieved are clean and only containing significant regions.

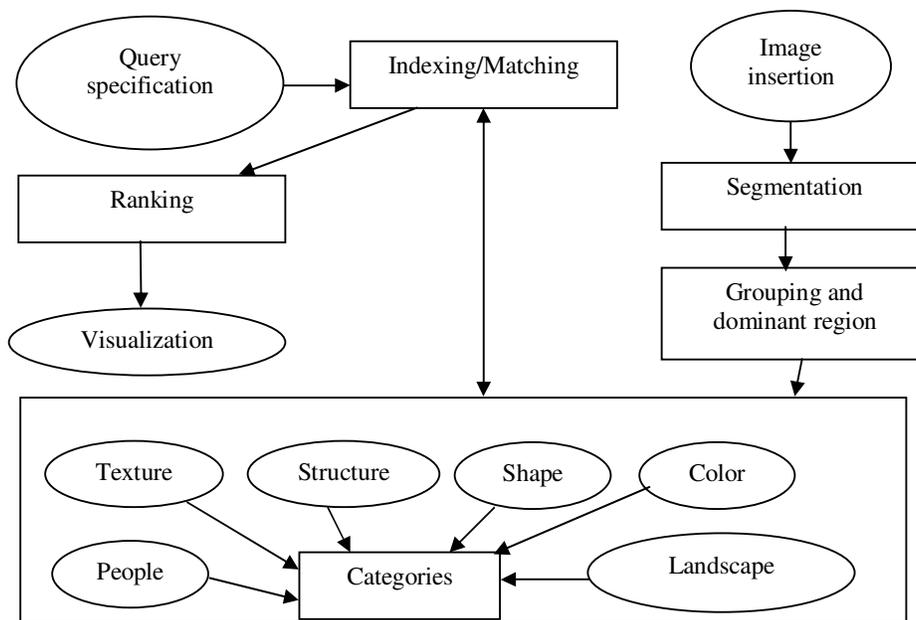


Figure 1. Architecture of new system

The proposed categories were chosen to provide sufficient grouping of images with similar themes, using general features. The domain assumed is photographic images, but with an unlimited range of themes nature, people, building, etc. The semantic categories are provided to detect most common types of images, such as those exist on the internet. The single object category is for images containing only two regions, conforming to the Gestalt figure/ground principle.

Table 1. Image categories

| Category Name      | Feature Characteristics   |
|--------------------|---|
| Landscape          | Colors green & blue<br>Spatial relation in vertical layers            |
| People             | Human skin hue  |
| Shape dominant     | 2 regions foreground/background image<br>Shapes are non-complex       |
| Color dominant     | More than 2 regions<br>Color distribution smooth (small variance)     |
| Texture dominant   | More than 2 regions<br>Color distribution non-smooth (large variance) |
| Structure dominant | More than 2 regions<br>Shapes complex<br>Contains geometric objects   |

**Image Segmentation**

The local color feature extraction starts with color image segmentation. For image segmentation, we use mean shift algorithm [7]. Here, color clustering is performed on each

image to obtain regions. After segmentation, only small number of color remains. Information like number of regions, time taken to segment an image, boundary points, region points, and region numbers can be extracted. Large classes of image segmentation algorithms are based on feature space analysis. In this paradigm the pixels are mapped into a color space and clustered, with each cluster delineating a homogeneous region in the image. Pixels were directly associated with the mode to which the path converged. The approximation does not yield a visible change in the filtered image. Recursive application of the mean shift property yields a sample mode detection procedure. The modes are the local maxima of the density and they can be found by moving each iteration the window by the mean shift vector, until the magnitude of the shifts becomes less than a threshold. The procedure is guaranteed to converge. The number of significant modes detected automatically determines the number of significant clusters present in the feature space.

### *Mean Shift Algorithm*

1. Choose the appropriate radius  $r$  of the search window.
2. Choose the initial location of the window.
3. Compute the mean shift vector

$$MSV = \frac{r^2}{p+2} \frac{\nabla p(x)}{p(x)}$$

Where  $x$  is the  $p$ -dimensional feature vectors,  $p(x)$  is the probability density function of  $x$  and  $\nabla p(x)$  is the gradient of  $p(x)$ .

4. Translate the search window by the shift amount.
5. Repeat till convergence.



Figure 3. (a) Original image. (b) Mean shift filtering result.

### *Region Grouping and Dominant Region Extraction*

To implement Gestalt laws [6] into our proposed grouping scheme, we separate them into two categories: local and global measure rules. Local measure rules are based on Gestalt

principles of proximity, similarity and good continuation. These rules will be used in grouping initial segments, to produce significant regions in the image. The first grouping performed is the size grouping. The aim of this grouping is to merge noisy areas (region whose size is less than 100) using the similarity of region. The next grouping stage is the color histogram grouping. This is performed by comparing the similarity of two region color histograms. This is then followed by line continuation grouping. Regions are grouped based on comparing line continuation surrounding regions.

The aim of dominant region extraction is to eliminate background, non-important regions, producing the most essential region. The reduction of non useful regions is required to reduce the matching and expensive structural analysis. The background will be eliminated by applying the figure/background principle of Gestalt laws. By analyzing that the largest region surrounding other objects entirely can conclude that this region is the background thus eliminated. In the proposed system, dominant region will be extracted automatically by analyzing the size (largest), location (center) of the regions and appearance of interesting shapes and structures. Interesting shapes will be judged on the region shape regularity and its geometric properties.

A five-step process is performed in the segment merging and dominant region extraction stage:

1. The image map is a two dimensional array, with the same proportions as the image. It contains the value representing the segment each pixel belongs to, at the corresponding point in the array.
2. Segment information is calculated from the image map. This includes information about the average color of the segment and its size. This information is required before any segment merging can begin.
3. A list of neighboring segments is recorded for each segment in the image. This list is used to determine the possible segments that the segment can be merged with.
4. Size merging is performed.
5. The dominant region is extracted from the image.

### ***Category Generation***

Category generation is responsible for assigning the image to one of the categories, with the category aiming to provide sufficient grouping of images with similar characteristics. These descriptors are based on a number of features that can be easily detected in images, but

are often best found via the application of a non-generic image analysis technique (i.e. texture dominant images are handled most effectively with robust texture matching, whereas shape dominant images are compared with good shape matching).

The shape dominant category is for images containing only two regions, with simple, regular shapes and for those conforming to the Gestalt figure/background principle. The color dominant category is for images containing regions with a smooth color distribution and less regular shapes. The texture dominant category is for images that contain highly textural regions (using the procedure proposed in [11]). The structure dominant category is for images that include straight lines, geometric shapes or conform to a structural template. The people category is for images with prominent regions with a hue in the human skin range. The landscape category is for pictures of landscapes that obey the landscape template and have regions with colors in the 'sky', 'land' and 'sea' color ranges.

### *Query Interface*

Here, we describe cluster-based indexing method aiming to speed-up the evaluation of range queries. This is carried out by reducing the number of candidate images, i.e. images on which the optimal region-matching problem has to be solved.

The procedure is as follows:

1. Given  $n$  the number of query regions, for each query region  $q_j$ , find the regions belonging to cluster  $c_j$ , where  $j=1..n$ .
2. For each region  $r_i$  in the image database
  - a. Find the feature vector  $f_i$ , for region  $r_i$ .
  - b. For each query region  $q_j$ , in the query set,
    - i. Find query feature vector  $f_j$ , for  $q_j$ .
    - ii. Find the Euclidean distance between  $f_i$ , and  $f_j$  using:

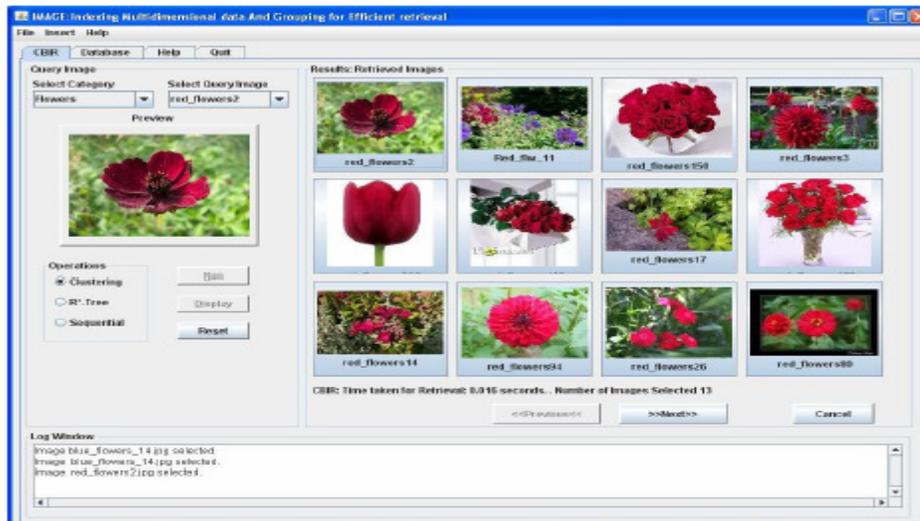
$$d_{ij} = \sqrt{\sum_{k=1}^m (f_{ik} - f_{jk})^2}$$

where  $m$  is the dimension of the feature vector. This score is zero if the regions features are identical, it increases as the match becomes less perfect.

- iii. Measure the similarity between  $f_i$  and  $f_j$  using  $\mu_{ij} = d_{ij} - \tau$ , where  $\tau$  is the search range limit set by user.

- iv. If  $\mu_{ij} \leq 0$ , then  $f_i$  belongs to cluster  $c_j$  and go to step 2.

After the completion of the above procedure, we index the images. Once the user selects the query, we apply range search on the tree and the selected regions belonging to image are retrieved as resultant set. Members of resultant set is ranked according to overall score and return the best matches in decreasing order of similarity along with their relative information.



*Figure.5. Result after comparison*

## Conclusion

This research aims to develop an image retrieval system that extracts the dominant region in an image, placing the image into one or more categories and search image with the help of Euclidean distance. These categories are developed from an understanding of how psychological principles apply to computer vision, extension of the proposed work some results such as people classifier are inaccurate. This research will provide a new image retrieval system that provides users with the ability to further classify the content of an image and search images. The impact from this method is more accurate retrieval results. An image will be represented by rich descriptions that relate directly to the content of the image. However, the idea of categorization and image search has been fulfilled in this paper.

## Future Work

In future by this work database can be classified. In this classified database retrieval will be efficient. In this method many classes of database will be created corresponding to

each basic color. With the development of CBIR, however, there are many problems. One is that since different researchers use different feature spaces, or use the same feature space but different description, then the measure is different, it is difficult for retrieval to be universal, especially in WEB. In future it can be enhanced to retrieval on internet or WEB Browsing. Currently, some results such as people classifier are inaccurate. Many images of people encountered have poor color quality and noisy. Thus, more robust classifier is required. We are using weighting method for categorization we can use better weighting, priority and hierarchical system within the proposed categories. Investigate a better classifier for different categories. Implement object based searching and queries. With clean and meaningful segmentation produced, this idea can now be realized.

### References

1. Kian-Lee Tan, Beng Chin Ooi & Chia Yeow Yee, *Multimedia Tools and Applications*, Vol 14, pp. 55–78, 2001.
2. Zaher Aghbari, & Akifumi Makinouchi, *Semantic Approach to Image Database Classification and Retrieval*, NII Journal, No. 7, 2003.
3. Shu-Ching, Chen Stuart H. Rubin, & Mei-Ling, *A Dynamic User Concept Pattern Learning, Framework for Content-Based Image Retrieval*, *IEEE transactions on systems, man, and cybernetics—part c: applications and reviews*, vol. 36, no 6, Nov. 2006.
4. A.Mohan, C. Papageorgiou, & T. Poggio. *Example based object detection in images by components*. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 23(4), pp. 349–361, April 2001.
5. S.Sclaroff, L. Taycher, & M. La Cascia. *Imagerover: A content based browser for the World Wide Web*. In *IEEE Workshop on Content-based Access of Image and Video Libraries*, Vol. 6, pp. 2–9, 1997.
6. A.Wardhani. *Application of psychological principal to automatic Object identification for CBIR*. PhD thesis, Information technology, Griffith University, 2001.
7. D.Comanicu & P. Meer. *Mean shift: A robust approach toward feature space analysis*. *IEEE Trans. Pattern Anal. Machine Intell.* Vol 24(5), pp. 603–619, 2002.
8. V.Bruce & P.R.Green, *Visual Perception: Physiology, Psychology and Ecology*, 2nd Ed., Lawrence Erlbaum Assoc., Hove and London, 1990.

9. Breiteneder, C., Merkl, D., & Eidenberger, H., Merging Image Features by Self-organizing Maps in Content-based Image Retrieval, Proc. of European Conference on Electronic Imaging and the Visual Arts, Berlin, 1999.
10. W.Y. Ma, B. Manjunath, NeTra: A toolbox for navigating large image databases," Proc. IEEE Int. Conf. Image Processing, Vol 5, pp. 568-71, 1997.
11. A.W. Wardhani & R. Gonzalez, "Using high level information for region grouping," in Proc. of the IEEE Region 10 Conference (Tencon'97): Speech and Image Technologies for Computing and Telecommunications, Vol 10, pp. 339-342, 1997.
12. G.Srinivas, E. Fasse, & M. Marefat. Retrieval of similarly shaped parts from a cad database. Systems, Man, and Cybernetics, Vol 3, pp. 109-118, 1998.
13. B.Mojilovic & J. Gomes., Semantic based categorization, browsing and retrieval in medical image databases. In Proceedings International Conference Image Processing, Vol. 3, pp. 145-148, 2002.
14. <http://www.qbic.almaden.ibm.com/>
15. <http://http.cs.berkeley.edu/~ginger/chabot.html>
16. H.Lu, B. Ooi, & K. Tan., Efficient image retrieval by color contents. In Proceedings of 1994 international conference on applications of databases, pp. 95-108, 1994.
17. T.S. Chua, K.-L. Tan, & B. C. Ooi., Fast signature-based color-spatial image retrieval. In Proceedings of IEEE conference on multimedia computing and systems, pp. 662-669, 1997.
18. M.Flickner, H. Sawhney, H. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, & P. Yanker., Query by image and video content: The QBIC system. IEEE Computer, Vol 28, No.9, pp. 23-31, 1995.
19. J.Smith & S. Chang. Visualeek: a fully automated content-based image query system. ACM Multimedia '96, 1996.