

A Study on Image Retrieval by Low Level Features

Nareshkumar .S
VIT University, Vellore, India

Vijayarajan .V
VIT University, Vellore, India

ABSTRACT

In Today's digital world vast amount of digital images are shared in online. When the number of images increased day by day the image retrieval based upon user perception is decreasing. In order to retrieve the relevant image we should need high efficient algorithms as well as retrieval methods. Image indexing plays an important role in image retrieval. Because by means of indexing only we can correlate the images and based on that only we can retrieve relevant images. But indexing also having problem when we had a very huge image database. Usually we are having two kinds of image retrieving methods Text Based IR and Content Based IR. In this paper we are going to see in detail about the methods as well as efficient algorithms for retrieving relevant image.

Keywords

Image Retrieval, low level features, CBIR.

1 . INTRODUCTION

Images are used to convey different aspect of important issues of life as well as makes the user to understand very clearly the information in single term. Images are definitely reducing work of an person in simple way. And also it is an effective way to remove the errors which are occurring in day to day life. But every minute lots of images are added to internet makes failure of the present image retrieval systems. Because traditionally the retrieval systems are based on the metadata of the image. Based on that only it going to retrieve the images from the database. But in reality if we are going to retrieve the images based on the metadata or keyword it may result in irrelevant images too. Keywords may not exactly mentioning an particular image and also same keyword can referred to one or more images in practical. So traditional approach is failed to meet the user needs. So here comes the problem that the images retrieved are not fully relevant or fully different result [1][2]. Here comes the content based image retrieval which was fully based on the low level features of images like Color, Texture and Shape etc. CBIR is getting the query as an image itself where in previous case we will get keyword from user. So in this first we need to extract the features of image based on the color, shape and texture. Then we need to mapping the features with the features of the images in the database. By this we can retrieve the relevant image but we had an some problem in this also. That is it will take more time to produce the output because the indexing of the images is not effective and features are not extracted correctly.

2. TEXT BASED IMAGE RETRIEVAL:

In this images are retrieved with the metadata related of an image which clearly describes about the image. But the problem with the textual based image retrieval is the users query keyword is based on user perception regarding the image to retrieve and also a single image can be referred by one or more keywords and vice versa. This is very easy to implement but in practical it will leads to lots of irrelevant information retrieval [2]. It will lead to overhead for the system to find the difference keywords which are processed in the natural language. This method is not fully producing the output what user needs because the image features are not taken care in this. But this problem is solved in the content based image retrieval because in CBIR we are going to consider the features of the images and without the consideration of the keyword or description of the image.

3. CONTENT BASED IMAGE RETRIEVAL

The most widen used method is CBIR system [6,7,8,9] because of the accuracy of retrieving the relevant images. In this method user is going to give the input itself as an image such that we can easily avoid the ambiguity due to description or keyword. This will increase the accurate images from the database. First the features of the input image is gathered using efficient algorithms and methods in various from by means of color, texture and shape. Based on the features values the image in the database is correlated and find the exactly matching images from the database and finally produce the output. Recently number of system are proposed using the CBIR system some of them are, MIT Media Lab's Photobook [10], NETRA [11] and Virage [12].

3.1 Image Retrieval based on Color

Using color we can retrieve the images. This was the first approach in retrieval of images using the features. Because based on color we can easily identify the images with similar colors by means Indexing based on color histogram [4].

Here we can see the color based image retrieval results. Query image is given at the top left corner and all other result images are mentioned respectively. The main



Fig: 1 Color based Image Retrieval

problem with the image retrieval with color is different images or objects with same color will be retrieved instead of accuracy. Our method CB is using the effective algorithm to retrieve the images from the database using the color intensity. But this method is not effective when the combination of color is present in the image. There is no effective methods to identify the exact color combinations in order to achieve the high accuracy.

3.2 Image Retrieval based on texture:

Texture based image retrieval is second technique which was employed in the CBIR. In this the pattern or similar texture alignments are extracted from the query image by means of converting the image into gray image and find the pixel intensities to find any patterns available. The major problem with this method is the main object in the image is not considered because we are going to find the pattern forms only. So the objects in the image may vary and thus result in the irrelevant images.

TB representing the images retrieved using the texture. Here we can see the result for the textural based image retrieval. The pattern recognition requires effective algorithm in order to achieve better result. Here the query image is represented at the top left corner and all other images are the results. Due to pattern finding the object in query image and retrieved image is same but the color is different, which makes the method ineffective. And also objects may be similar but the background patterns will be vary such that the system will be produce the output as irrelevant images. This is because of the

similar texture patterns available but the objects contained in the image may be vary. This makes the algorithm very ineffective and overhead over the system.

3.3 Image Retrieval based on Shape:

Shape based image retrieval is done by extract the shape of the object in the image and finds the same in database. It is not preferred by most of the users because there is no exact algorithm to find the shape of the image and also we need to find the exact object in the image for retrieving user query. This is done by means of finding the edges of the objects in the image. The major problem is different objects interacts and changes the border in an image but at the same time the primary object may present in the database image but it doesn't have the different objects such that the shape of the two is totally different while retrieving. In such case the result will not be same and irrelevant images will be retrieved.

SB representing the images retrieved based on the shape. If two different objects having similar shapes means the algorithm is failed to differentiate the two and retrieve both of the images.

So when we are using all the three technique separately we can't achieve high accurate result. And also we can combine the features means we can achieve moderate result.



Fig:2 Texture based Image Retrieval

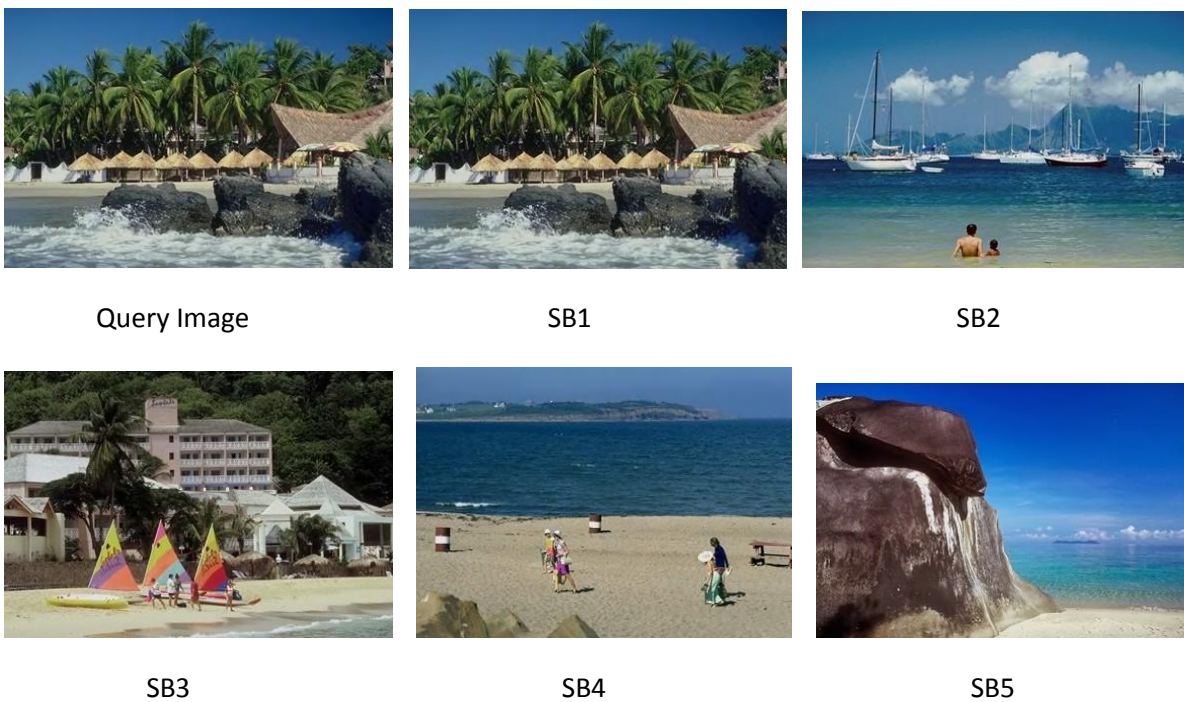


Fig: 3 Shape Based Image Retrieval

4. PRECISION AND RECALL:

In Pattern recognise and information retrieval, **precision**[4] is the fraction of retrieved instances that are relevant, while **recall**[4] is the fraction of relevant instances that are retrieved. Both precision and recall are therefore based on an understanding and measure of relevance. Both precision and recall are therefore based on an understanding and measure of relevance . In even simpler terms, high **recall** means that an algorithm returned most of the relevant results. High **precision** means that an algorithm returned more relevant results than irrelevant.

Precision:

$$P(a) = N(a) / N(a) \text{ and } W(a)$$

Recall:

$$R(a) = N(a) / T(a)$$

N(a)- Number of Images retrieved relevant to query image

W(a)- Number of images retrieved irrelevant to query image

T(a)- Total number of images in database, which is used to compare with the query image

Table 1: Precision and Recall for IR based on Colour

Input	Output	Precision	Recall
1(5)	5 of same colour	1	0.1
1(10)	7 of same colour and 3 other colour	0.7	0.14
1(15)	12 of same colour and 3 of different colour	0.8	0.24

Table 2 : Precision and Recall for IR based on Texture

Input	Output	Precision	Recall
1(5)	3 same texture and 2 different texture	0.6	0.1
1(10)	6 same texture and 4 different texture	0.6	0.2
1(15)	8 same texture n 7 different texture	0.533	0.3

Table 3 : Precision and Recall for Image retrieval based on shape

Input	Output	Precision	Recall
1(5)	3 same and 2 different shape	0.6	0.1
1(10)	6 same and 4 different shape	0.6	0.2
1(15)	8 same and 7different Shape	0.5	0.3

5. CONCLUSION

From the experiments we can conclude that if we combined all the features also there will be chances of irrelevant images. This is because when we are having the same shape, same color and texture pattern of different images leads to irrelevant image retrieval. In future the low level features can be combined with the high level features like ontology semantics(tagging) of images in achieving the high accurate relevant image retrieval.

6. REFERENCES

- [1] *The SMART Retrieval System: Experiments in Automatic Document Processing*. Prentice-Hall, 1971.
- [2] Gerard Salton and Christopher Buckley. Term-weighting approaches in automatic text retrieval, *Information Processing and Management*, 1998.
- [3] C. Faloutsos, W. Equitz, M. Flickner, W. Niblack, D. Petkovic, and R. Barber. *Journal of Intelligent Information Systems*, 3(3/4):231–262, July 1994.
- [4] Michael J. Swain and Dana H. ballard. Indexing via Color Histograms, University of Rochester, NY 14627, USA.
- [5] [www.wikipedia.com/precision and recall](http://www.wikipedia.com/precision%20and%20recall)
- [6] C. Faloutsos, W. Equitz, M. Flickner, W. Niblack,
- [7] D. Petkovic, and R. Barber. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, 3(3/4):231–262, July 1994.
- [8] M. Stricker and M. Orengo. Similarity of color images. In *Storage and Retrieval for Image and Video Databases SPIE*, volume 2420, pages 381–392, San Jose, CA, Feb. 1995.
- [9] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. In B. Furht, editor, *Multimedia Tools and Applications*, chapter 2, pages 43–80. Kluwer Academic Publishers, 1996.
- [10] J. R. Smith. Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression. PhD thesis, Columbia University, 1997.
- [11] A. Pentland, R. W. Piwd, and S. Sclaroff. Photobook Tools for content-based manipulation of image databases. In *Srorage and Retrieval for Image and V i 0 Databases II*, volume 2185 of *SPIE Proceedingsseries*, San Jose, CA, USA, 1994.
- [12] [11] W. Y. Ma and B. S. Manjunath. A toolbox for navigating large image databases. In *IEEE International Conference on Image Processing (KIP)*, Santa Barbara, California, October 1997.
- [13] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jab. and C.-F. Shu. The Wage image search engine: An open framework for image management. In I. IC. Sethi and R. J. Jain, editors, *Storage and Retrieval for Image and V i 0 Databases N*, volume 2670 of *Proceedings of SPIE*, pages 76-87, 1996.