Analysis and Implementation of Effective Content Based Image Retrieval Based on Color, Texture And Shape Feature Extraction

Ajay Sancha[1], Dr. Sanjay Chouhan[2]
[1, 2] Jawaharlal Institute of Technology, Borawan

Abstract- *Content based image retrieval from large resources has become an area of wide interest now a day in many problems. In this thesis it is introduced that Contend-based image retrieval system that uses color, shape and texture as visual features to describe the content of an image region. Our contribution is of three directions. Our proposed system has the advantage of increasing the retrieval accuracy and decreasing the retrieval time. The experimental evaluation of the system is based on a 1000 WANG color image database. From the experimental results, it is evident that our system performs significantly better and faster compared with other existing systems.*

In our simulation analysis, it is provided a comparison between retrieval results based on features extracted from the whole image, and features extracted from some image regions. The results demonstrate that each type of feature is effective for a particular type of images according to its semantic contents, and using a combination of them gives better retrieval results for almost all semantic classes.

Keywords- Image Processing, Content based Image Retrival, GLCM, Fourier Descriptor, Wang Database.

I. INTRODUCTION

The Content based image retrieval has been an active research area in computer vision and image processing. In CBIR system it can search, browse and navigate similar images from large image databases based on visual content of the images. Visual content of an image can be described formally in terms of color, texture and shape features. Traditional CBIR system makes use of these features to index and retrieval similar images from the database [1]. Some of the existing popular CBIR systems are QBIC [2], Simplicity [3] and Visual-seek [4].

Content-based image retrieval, in our technique which uses visual contents to search images from large scale image databases according to the users' interests; it has been an active and fast advancing research area since the 1990s. During a past decade, remarkable progress has been made in both theoretical research and system development. However, there are remaining many challenging research problems that continue to attract researchers from multiple disciplines. It's before introducing the fundamental theory of content-based retrieval they will take a brief look at its development. Early the work on image retrieval can be traced back to a late 1970s. In 1979, a conference on Database Techniques for Pictorial Applications [5] is held in Florence. Since then, these application potential of an image database management techniques has attracted the attention of researchers [6]. The Early techniques were not generally based on a visual feature but on the textual annotation of images. In other words search images were first annotated with text and then searched using a text-based visual content approach from traditional database management systems.

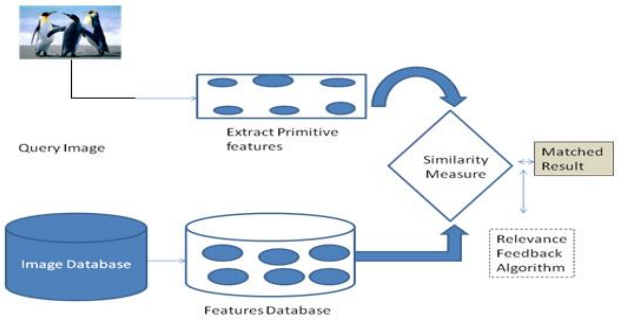

Figure 1.1: A Conceptual Framework for content-based image retrieval system

Content-based image retrieval uses the visual contents of an image such as color, shape and texture, and a spatial layout to represent an index on image. In typical content-based image retrieval systems (Figure 1.1), the visual contents of the images in the database are extracted and described by the multi-dimensional features vectors database. The feature vectors of the images in the database form a feature database. To retrieve the images, users provide the retrieval system with example images or sketched figures. This system then changes these examples into its internal representation of feature vectors. The similarities /distances between the features Vectors of the query example or sketch

and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing schemes.

The indexing scheme provides an efficient way to search for the image database. A Recent retrieval system has been incorporated users' relevance feedback to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results.

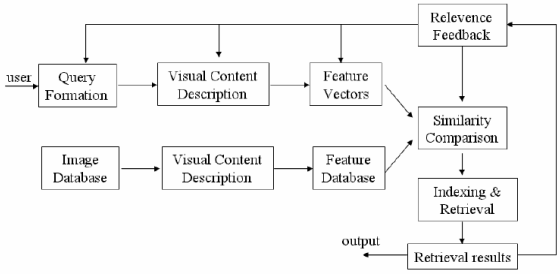

Figure 1.2 Content based image retrieval system

II. LITERATURE REVIEW

Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape [1]. Reasons for its development are that in many large image databases, a traditional method of image indexing have proven to be an insufficient, laborious, and is extremely time consuming. These old methods of that are image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete. CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image [5]. Several CBIR systems currently exist, and are being to constantly develop.

CBIR for general-purpose image databases is a highly challenging problem because of the large size of the database and the difficulty to understanding images, both by a people and computers, the difficulty of formulating a query, and the issue of evaluating results properly. A number of general-purpose image search engines have been developed. In the commercial domain, QBIC [2] is one of the earliest systems..

The common ground for CBIR systems is to extract a signature for every image database on its pixel values and to define a rule for comparing images. The signature serves as an image representation in the “view” of a CBIR system. The components of the signature are called features. Actually, the main task of designing a signature is to bridge the gap between image semantics and the pixel representation that is to create a better correlation with image semantics [5].

The Existing general-purpose CBIR systems roughly fall into three categories depending on the approach to extract signatures: histogram, color layout, and region-based search. There are also systems that combine retrieval results from individual algorithms by a weighted sum matching metric [10], or other merging schemes [11].

After extracting signatures, the next step is to determine a comparisons rule that is including querying scheme and the definition of a similarity measure in between images. For most image retrieval systems that query is specified by an image to be matched. We refer to this as global search since similarity is based on the overall properties of images. In contrast there are also “partial search” querying a systems that retrieve results based on a particular region in an image [10].

Efficient retrieval system of similar image documents based on their features requires categorizing and therefore clustering them based on some criteria. An in Image clustering and categorization is a means for high-level description of image content. The key to the retrieval process is similarity among low-level visual features.

The Color Selection of exploited CBIR system [12] facilitates query-by-color. It is based on 11 color categories, used by all people, while thinking of a perceiving color. Then the low frequency DCT coefficients that are transformed from YUV color space as feature vectors are used for retrieval of images [9] This system allows users to select its dominant feature of query images so as to improve the retrieval performance. Since the technique is sufficient for performing effective retrieval by introducing users' opinions on the query images.

III. PROBLEM IDENTIFICATION & ALGORITHM

1. Feature extraction (Visual content description): How to extract low level features from an image and store in logical database.
2. Similarity measure: How to design a similarity measure algorithm?
3. User interface: How to design a mathematically description of a query, how to select a feature subset for a specific query?
4. Retrieval: How to retrieve the relevant images for a given query images according to similarity measure
5. Evaluation: How to evaluate the performance of a CBIR system?

From all these point we have to work on Image Retrieval based on content because existing system are having less accuracy according to literature survey.

Retrieval Algorithm

1. Step 1: Uniformly divide each image in the database and the target image into 8-coarse partitions as shown in Color Feature Extraction.
2. Step 2: For each partition, a centroid of each partition is selected as its dominant color.
3. Step 3: Obtain texture features (Energy, Contrast, Entropy and inverse difference) from GLCM.
4. Step 4: Obtain Fourier descriptor (FD) as shape features
5. Step 5: the construct a combined feature vector for color, texture and shape.
6. Step 6: Find the distances between feature vector of query image and the feature vectors of target images using weighted and normalized Euclidean distance.
7. Step 7: Sort the Euclidean distances.
8. Step 8: Retrieve first 20 most similar images with minimum distance.

IV. SIMULATION RESULTS

The database we used in our evaluation is WANG database [17]. The WANG database is a subset of the Corel database of 1000 images, which have been manually selected to be a database of 10 classes of 100 images each.

Figure 1.3: Example Images from each of the 10 Classes of WANG Database

We evaluate the system regarding two metrics: the effectiveness in terms of precision and recall, and the efficiency in terms of the time the system take to answer a query.

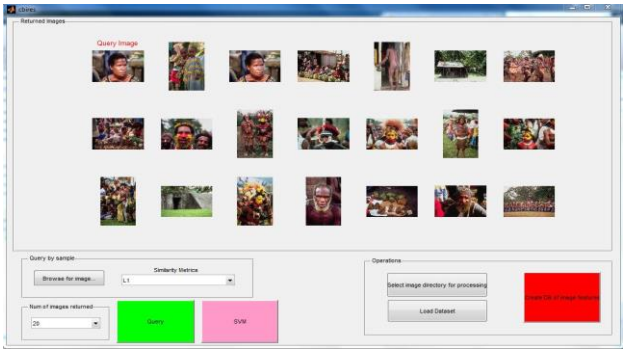

Figure 1.4: Image Retrieval for Query 1 - Africans

Figure 1.4 shows the retrieval of images for query number 1 i.e. images from category of Africans people. 17 out of 20 images are correctly retrieved.

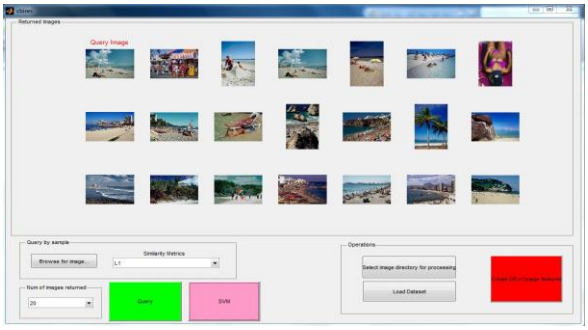

Figure 1.5: Image Retrieval for Query 2 - Beach

Figure 1.5 shows the retrieval of images for query number 2 i.e. images from category of Beaches. 18 out of 20 images are correctly retrieved.

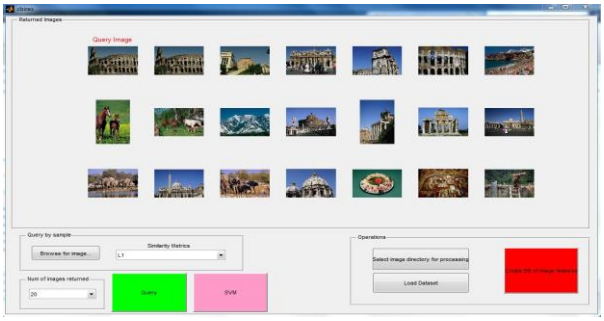

Figure 1.6: Image Retrieval for Query 3 - Building

Figure 1.6 shows the retrieval of images for query number 3 i.e. images from category of Buildings. 13 out of 20 images are correctly retrieved.

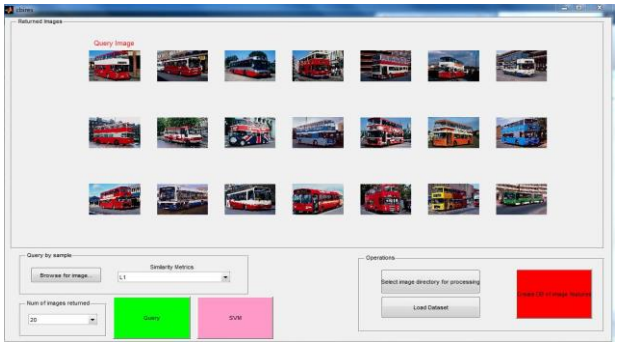

Figure 1.7: Image retrieval for query 4 - Bus

Figure 1.7 shows the retrieval of images for query number 4 i.e. images from category of Bus. 20 out of 20 images are correctly retrieved.

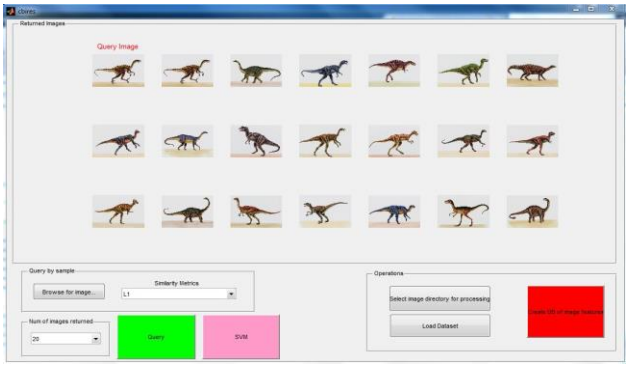

Figure 1.8: Image retrieval for query 5 - Dinosaur

Figure1.8 shows the retrieval of images for query number 5 i.e. images from category of Dinosaur. 20 out of 20 images are correctly retrieved.

Table 1.1 shows the various features i.e. Recall and Precision used in the proposed work in terms of Accuracy rate implemented in Wang image Database.

Table 1.1: Value of Precision and Recall

S.No	Class	Image Number	Recall	Precision
1	Africans	1	0.17	0.85
2	Beach	102	0.18	0.9
3	Building	200	0.13	0.65
4	Bus	330	0.2	1
5	Dinosaur	401	0.2	1
6	Elephant	501	0.13	0.65
7	Flower	610	0.19	0.95
8	Horse	701	0.19	0.95
9	Mountain	802	0.4	0.2
10	Food	942	0.18	0.9
Average			**0.197**	**0.805**

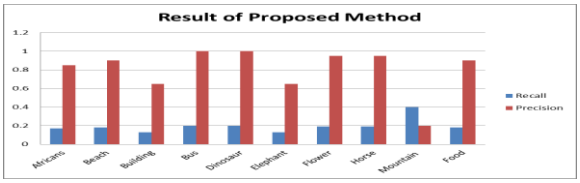

Figure 1.9: Graphical representation of Precision and Recall

Table 1.2 shows the comparison of Precision used in the proposed work with two different methods used in Literature review for the Wang Database

Table 1.2: Comparison of Precision with base paper

S.No	Class	Anandh [1]	Chathurani [2]	Singh [3]	Proposed
1	Africans	0.76	0.72	0.8	0.8
2	Beach	0.83	0.39	0.5	0.75
3	Building	0.75	0.44	0.55	0.8
4	Bus	0.87	0.85	1	0.65
5	Dinosaur	1	1	1	1
6	Elephant	0.72	0.56	0.9	0.75
7	Flower	0.94	0.75	0.85	1
8	Horse	0.84	0.89	0.75	0.8
9	Mountain	0.64	0.48	0.6	0.5
10	Food	0.63	0.58	0.75	0.9
Average		**0.79**	**0.67**	**0.77**	**0.80**

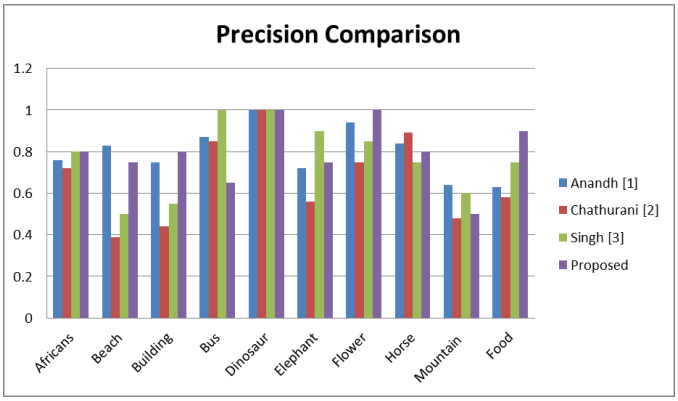

Figure 1.10: Graphical representation of Precision Comparison

V. CONCLUSION

Our proposed system, have good retrieval results and high precession/recall values. According to our simulation results, the CBIR system can be used as the first option in our retrieval system, since it gives accepted results and avoids the complex computations of the segmentation process and region comparison that are present in the other system, which can be used next to further improve the retrieval results in case of not satisfying the user.

The following developments can be made in the future:

1. Region based retrieval systems are effective to some extent, but their performance is greatly affected by the segmentation process. Development of an improved image segmentation algorithm is one of our future works.
2. To further improve the performance of the retrieval system, the study of taking shape features into account during similarity distance computation can be considered.
3. Demonstration of using different color and texture weights in Equation and their effect on the retrieval results.
4. As further studied, the proposed retrieval methods is to be evaluated for more various DBs and to be applied to video retrieval.

REFERENCES

[1] A.Anandh, Dr.K.Mala, S. Suganya," Content Based Image Retrieval System based on Semantic Information Using Color, Texture and Shape Features", International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), IEEE 2016

[2] Chathurani N. W. U. D., Geva S., Chandran V., Cynthujah V."An Effective Content Based Image Retrieval System Based on Global Representation and Multi-Level Searching" International Conference on Industrial and Information Systems, ICIIS IEEE 2015

[3] Jyotsna Singh, Ahsaas Bajaj, Anirudh Mittal‡, Ansh Khanna and Rishabh Karwayun, "Content Based Image Retrieval using Gabor Filters and Color Coherence Vector", 8th International Advance Computing Conference (IACC), IEEE April 2019

[4] Jun-Yue and Zhenbo li "Contend-based image retrieval using color texture fused Features" mathematical and computing Modeling 54, 1121-1127, 2011.

[5] The QBIC web site: http://wwwqbic.almaden.ibm.com.

[6] J. Wang, J. Li, G. Wiederhold, "Simplicity: semantics-sensitive integrated matching for picture libraries," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, pp. 947–963, Sep. 2001.

[7] J.R. Smith, C.S. Li, "Image classification and querying using composite region templates," International Journal of Computer Video Database, vol. 75, pp. 165–174, 1999.

[8] N. Vasconcelos, "From Pixels to Semantic Spaces: Advances in Content-Based Image Retrieval", Computer Volume: 40, Issue: 7, pp. 20-26, 2007.

[9] Asmita Deshmukh & Leena Ragha " An Effective CBIR using Texture" International Conference & Workshop on Recent Trends in Technology, (TCET) Proceedings published in International Journal of Computer Applications (IJCA) 2012.

[10] Xiang-Yang Wang & Yong-K=Jian Yu "An Effective image retrieval scheme using color, texture and shape feature" Elsevier computer standard & interface 33, pp. 59-68 2011.

[11] F. Long, H. Zhang, H. Dagan, and D. Feng, "Fundamentals of content based image retrieval,"Multimedia Signal Processing Book, Chapter 1, Springer-Verlag, Berlin Heidelberg New York, pp.1-26.2003.

[12] J. Mao, and A. Jain, "Texture Classification and Segmentation using Multi- Resolution Simultaneous Autoregressive Models," Pattern Recognition, vol. 25, no.2, pp. 173-188, 1992.

[13] F. Li, Q. Dai, W. Xu and G. Er, "Multi label Neighborhood Propagation for Region-Based Image Retrieval", IEEE Transactions On Multimedia, Vol. 10, No. 8, pp. 1592-1604, December 2008.