e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

DESIGN AND IMPLEMENTATION OF CONTENT BASED IMAGE RETRIEVED SYSTEM BASED ON FEATURE EXTRACTION

Deepika Chouhan*1, Mukta Sahu*2, Deepak Sharma*3

*1M.tech Scholar, Department Of EC Engineering, LNCT (Bhopal) Indore Campus, Indore(M.P.), India.

*2,3Assistant Professor, Department Of EC Engineering, LNCT (Bhopal) Indore Campus, Indore(M.P.), India.

(Old Name:- Lord Krishna College Of Technology, Indore)

ABSTRACT

Content based image retrieval from large resources has become an area of wide interest now a day in many problems. In this work, 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. To start with, utilize the Gray-level co-event matrix to extract texture features from arbitrarily formed regions isolated from an image after segmentation to increase the framework viability. This procedure is performed disconnected before query handling, along these lines to answer an inquiry our framework does not have to look through the whole database images; rather only various hopeful pictures are required to be hunt down image similarity. Third, to additionally build the recovery accuracy of our system, consolidate the content based features extracted from picture locales, with global features separated from the entire image, which are texture utilizing GLCM, Shape utilizing Fourier Descriptor and color utilizing Domain color Descriptor. In our analysis it has the advantage of Continual increasing the retrieval accuracy and decreasing the retrieval time. 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 we have received explains that for a particular image which are in accordance with the semantic contents and by using a combination of feature extraction gives efficient retrieval results for all semantic classes.

I. INTRODUCTION

1.1 Content Based Retrieval System

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, which uses visual contents for searching images from large scale image databases according to the ones interest; it has been an active and fast advancing research area since the 1990s. Since many years, a revolutionary progress has been done in both theoretical research and system development. However, there are many challenging research obstacles that will continue to attract innovators and researcher among many 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.

There are wide and expanding surveys on the text-based Image retrieval methods can be found[7]. Text-based image retrieval uses traditional database techniques to a manage images. Through text descriptions, images can be organized by typical or semantic hierarchies to facilitate easy to navigation and browsing based on the standard Boolean queries. However, since automatically generating descriptive texts for a wide spectrum of images it is not feasible, in most text-based image retrieval systems it requires a manual annotation of images.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

1.2 CBIR FRAMEWORK

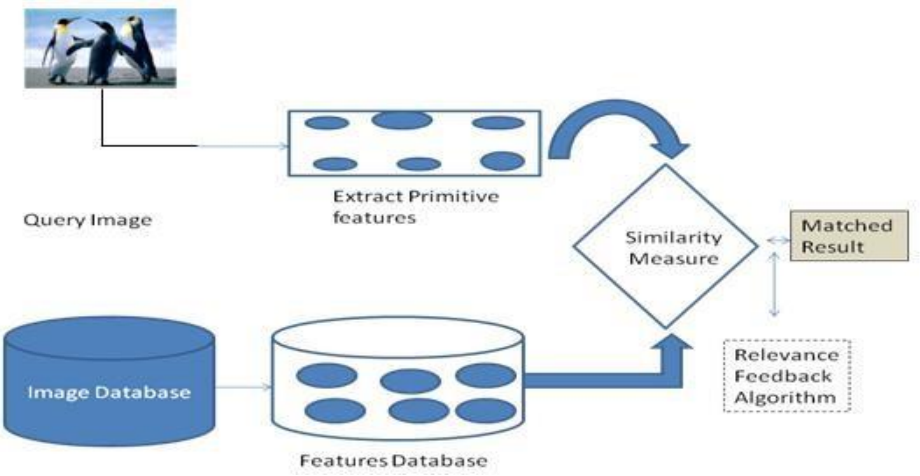

Figure 1.1: A Conceptual Framework for content-based image retrieval system

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 featured vectors of the images which have stored in the database form a feature database. To retrieve the images, users provide the retrieval system with example images or sketched figures, and the visual contents of the images stored in the database are extracted and analyzed by the multi-dimensional features vectors database.

1.2.1 CBIR CONCEPT

The CBIR systems (Figure 1.2), demonstrates the visual contents of the images in the database which are extracted and described by the multi-dimensional feature vectors. The system then changes these particular examples into its internal representation of feature vectors. A feature vector of the images in the database forms a feature database.

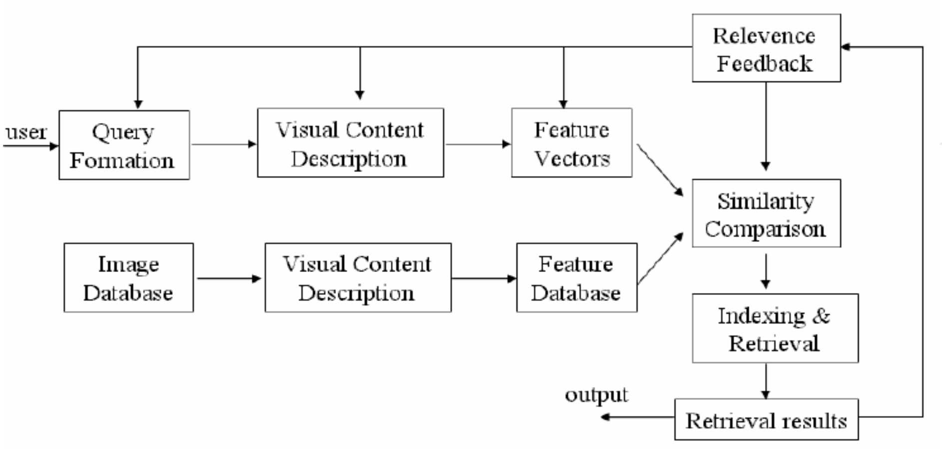

Figure 1.2: Content based image retrieval system

The differences and similarities among the featured vectors of the query examples or sketched 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. In the proposed retrieval system that have incorporated users' relevance feedback has modified the retrieval process.

Following are the problems to be solved in a CBIR system:

Feature extraction (Visual content description) How to extract low level features from an image and store in logical database.

1. **Similarity measure** How to design a similarity measure algorithm?
2. **User interface** How to design a mathematically description of a query, how to select a feature subset for a

specific query?

3. **Retrieval** How to retrieve the relevant images for a given query images according to similarity measure
4. **Evaluation** How to evaluate the performance of a CBIR system?

1.3 REGION BASED IMAGE RETRIEVAL (RBIR)

Early in CBIR methods used global feature extraction to obtain the image descriptors.

For example, QBIC [2, 51] developed at the IBM Alma den Research Center extracts several features from each image, namely the color, texture, and shape features. These descriptors are obtained globally by extracting an information by using a means of color histograms for color features; global texture information on with coarseness, contrast, and the direction.

1.4 FIELDS OF APPLICATION

Image retrieval based on visual content is widely adoptable and very useful in a variety of applications such as publishing media and content research writing, historical research, fashion designing, graphical and visual implantation and design, architectural and engineering design, etc. [11]. In hospitals and health care sectors, some ailments require the medical practitioner to search and review similar X- rays or scanned images of a patient before pro ffering a solution.

II. SURVEY ON CBIR

Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape [1]. The most important reason for its development is that for many large image databases, a traditional method of image indexing have proved to be an insufficient, very hard working, and is time consuming. These traditional methods of that are image indexing, and ranging to storing an image in the large image database and associate it with a phrase, keyword or number, to associate it with a categorized description, which is becoming obsolete task. CBIR, each image that is stored in the database has its features extracted and compared to the features of the quer y 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, bothby a people and computers, the difficulty o f formulating a query, and the issue of evaluating results properly. A number of general-purpose image search engines have been developed. In the commercial do main, 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. However, a more important reason for using the signature is to gain an improved correlation between image representation and semantics. The signature serves as an image representation in the “view” of a CBIR system. The components of the signature are called features. One advantage o f a signature over the original pixel values is the significant compression of image representation. 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].

2.1 IMAGE CLASSIFICATION

Contrary to clustering, the process of assigning a data item to a predefined set of categories which is called as classification [4]. Images are grouped into semantically more meaningful image categories based on low- level visual features that have formed using statistical classification methods. Retrieval systems assume that sets of images that have similar visual characteristics based on their low- level features are also semantically related [25]. Assigning images into different semantic classes is the main of image classification..

2.2 GLOBAL FEATURE BASED CBIR SYSTEMS

Some of the existing CBIR systems extract features from the whole image not from certain regions in it; these features are referred to as Global features. Histogram search algorithms [26] characterize an image by its color distribution or histogram. Too Many distances have been used to define the similarity o f two color histogram representations. A Euclidean distance and its variations are the most commonly used.

Image retrieval using color features often it gives disappointing results, because in many cases, images with similar colors do not have similar content. This is due to the fact that global color features often fails to capture color distributions or textures within the image. D. Zhang &Y. Deng [36] proposed a method combining both color and texture features to improve retrieval performance. By computing both the color and texture features

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

from the images, the database images are indexed using both types of features. During the retrieval process, given a query image, images in the database are firstly ranked using co lor features. In a second step, a number of top ranked images are selected and re- ranked according to their texture features. Two alternatives are provided to the user, one is the retrieval based on color features, and the other is retrieval based on combined features. When the retrieval based on color fails, the user will use the other alternative which is the combined retrieval. Since the texture features are extracted globally from the image; they are not an accurate description of the image in some cases, which degrades the system performance.

III. PROBLEMFORMULATION

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.

3.1 COLOR FEATUREEXTRACTION

For extracting the color information's, images are preprocessed by k-means as and a clustering algorithm for reducing color and then dominant color feature vector is extracted as described in following subsections.

3.1.1 k-means clustering algorithm

The K-means is an easy unsupervised learning algorithms which can solve the well-known clustering problem. The next procedure which had followed a very simple and most promising way to classify a given set of data through a numbers of clusters which have a is fixed a priority. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. At this point we need to re-calculate k-new centroids as bare centers of the clusters resulting from the previous step. After we have these So, the better choice is to place them as much as possible far away from each other. In our next step is to take each point belonging to a given data set and associate it to the nearest centroid. Then no point is pending, the first step is completed and an early group age is done.

The K-means is a easy to use and simple algorithm that has been used for adopting a certain solution to many problem domains. K- Mean's algorithm is used to reduce the number of color in the image. In all the colors are grouped in to k cluster such that mean value of colors in each cluster issame.

3.1.2 Dominant Color Extraction

Images are converted in to a HSV color space before extracting color features. HSV color space is more perceptually uniform than other color spaces [7, 59]. The images are then divided in to layout of size 3 x3, containing 9 non-overlapping blocks of uniform size. From each block one dominant color representing the block is extracted using k-means algorithm as explained above. For this pixel in each block are grouped in to k cluster with respect to similarity of their color.

Figure 3.1: The Original Image and Its Corresponding Color Reduced Image

The color representing the centroid of the cluster containing the highest number of pixels is taken as the dominant color of that block. As a result, we obtain 9 dominant color representing color of each of the nine blocks respectively. Figure 4.1 shows the original image and the color reduced image containing 9 colors, obtained after applying k-means algorithm. Now the color feature vector F_c can be given as

$$F_c=(D_i,P_i) \quad i=1 \ldots\ldots M \quad 3.1$$

Where, D_i is the dominant color and P_i is the percentage of that color.

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

IV. RESULTS

We had present an evaluation of the proposed CBIR systems: that we introduced in the previous chapters. We also compare their performance with two other recent CBIR systems.

4.1 WANGDATABASE

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. The images are subdivided into 10 classes, such that it is almost sure that a user wants to find the other images from a class if the query is from one of these 10 classes. This is a major advantage of this database because due to the given classification it is possible to evaluate retrieval results. The images are of size 384 × 256 or 256 × 384 pixels,

IMPLEMENTATIONENVIRONMENT

The image retrieval system is implemented using MATLAB image processing tools and statistical tools. During the implementation, we use a platform of Intel Core i3 of 3.2 GHz CPU with 3 GB RAM. 1000 image database went through image segmentation algorithm to obtain more than 5800 objects. These images are manually divided into 10 classes such as Africans, buses, buildings, and flowers.

Usually, precision and recall scores are not discussed in isolation. Instead, either values for one measure are compared for a fixed level at the other measure (e.g. precision at a recall level of 0.75) or both are combined into a single measure, such as precision/recall graph. Precision-recall pair is a good standard of performance evaluation. It provides meaningful result when the database type is known and has been effectively used in some earlier research. For other data sets, especially those that have been created by collecting user generated images, the result may vary due to different human concepts of image classification.

4.2 EVALUATION OF PROPOSEDSYSTEM

In this section, we test the main part of our system. We had evaluate the system regarding in 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.

The Figure shows the retrieval of images that is stored in the system for query number 2 i.e. images from category of Beaches. 18 out of 20 images are correctly retrieved. Confusion matrix represents performance of retrieval system in form of true and false positive rate.

Similarly the simulation is done on 10 different images of Wang data set, and the confusion matrix is drawn. Precision and Recall can be find out by the confusion matrix. Table 4.1 shows the various features i.e. Recall and Precision used in the proposed work in terms of Accuracy rate implemented in Wang image Database.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

Table 4.1: Value of Precision and Recall

S.No	Class	Image Number	Recall	Precision	Accuracy
1	Africans	1	0.17	0.85	82 %
2	Beach	102	0.18	0.9	68 %
3	Building	200	0.13	0.65	60 %
4	Bus	330	0.2	1	90 %
5	Dinosaur	401	0.2	1	100 %
6	Elephant	501	0.13	0.65	86 %
7	Flower	610	0.19	0.95	94 %
8	Horse	701	0.19	0.95	94 %
9	Mountain	802	0.4	0.2	72 %
10	Food	942	0.18	0.9	84 %
Average			**0.197**	**0.805**	**83**

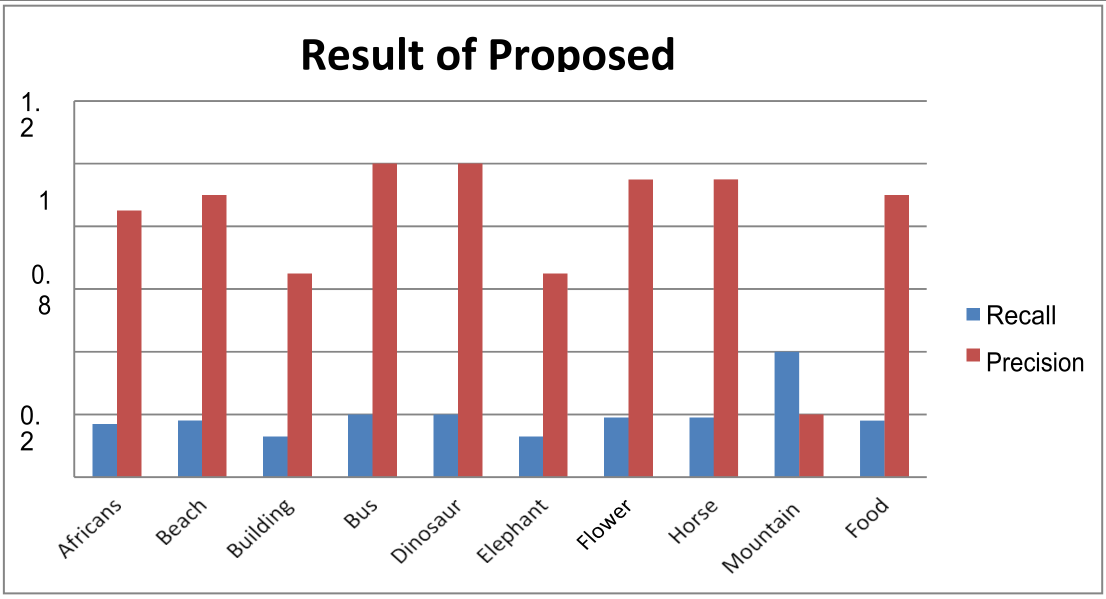

Figure 4.1: Graphical representation of Precision and Recall

Table 4.2 shows the comparison of Precision used in the proposed work with two different methods used in Literature review for the Wang Database

Table 4.2: Comparison of Precision with base paper

S.No	Class	Proposed CBIR	Anandh [25]	Chathurani [26]
1	Africans	0.85	0.76	0.72
2	Beach	0.9	0.83	0.39
3	Building	0.65	0.75	0.44
4	Bus	1	0.87	0.85
5	Dinosaur	1	1	1
6	Elephant	0.65	0.72	0.56
7	Flower	0.95	0.94	0.75
8	Horse	0.95	0.84	0.89

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

9	Mountain	0.2	0.64	0.48
10	Food	0.9	0.63	0.58
Average		**0.805**	**0.798**	**0.666**

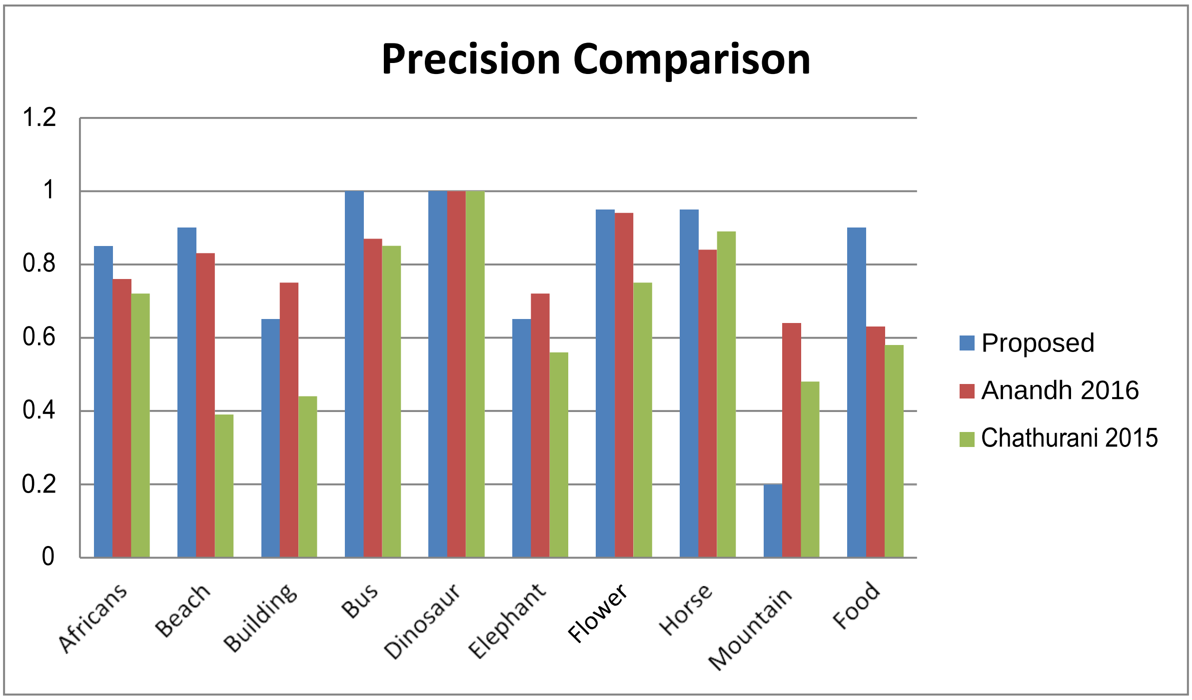

Figure 4.2: Graphical representation of Precision Comparison

4.3 RESULT

According to the simulation and values of precision, it is clearly indicated that our proposed method provides better results than existing retrieval methods for the same database

V. CONCLUSION

Content based image retrieval is a tough but mere challenging method of capturing relevant images from a large image storage space. Although this area has been explored for many years, there is not an any technique has achieved the accuracy of human visual perception in exploring and differentiating images. Whatever the size and content of the image database is, a human being can easily recognize images of same category.

The work that is suggested by us has presents a easy and effective combination of features for efficient content based image retrieval system. The experimental results have shown that the proposed method is quite effective and is better than some of the traditional methods in retrieving user intended images. In future the validity of the proposed technique will be tested on larger image database and work will be done to further improve the method by utilizing more effective feature set in the process of retrieval.

Our system, have good retrieval results and high precision/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 there retrieval results in case of not satisfying the user.

VI. FUTURESCOPE

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. To obtain better performance, the system can automatically pre-classify the database into different semantic

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:04/April-2022 Impact Factor- 6.752 www.irjmets.com

images (such as outdoorvs. indoor, landscape vs. cityscape, texture vs. non texture images) and develop algorithms that are specific to a particular semantic image class.

VII. REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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, December2008.

[8] Ryszard S. Chora´s" Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems" international journal of biology and biomedical engineering,2007.

[9] R. C. Veltkamp and M. Tanase, "Content-Based Image Retrieval Systems: A Survey", http://aalab.cs.uu.nl/cbirsurvey/cbir-survey2004.

[10] Ritendra Datta, Dhiraj Joshi, Jia Li, And James Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age", ACM Computing Surveys, Vol. 40, No. 2, Article 5, Publication date: April2008.

[11] A. Pentland, R. Picard, and S. Sclaroff , "Photobook: Content Based Manipulation of Image Databases," International Journal of Computer Vision, vol.18, no 3, pp.233– 254, June 1997.

[12] G. Pass and R. Zabih, "Histogram Refinement for Content Based Image Retrieval", 3rd IEEE Workshop on Applications of Computer Vision, pp. 96-102,1966.

[13] J.Z. Wang, "Wang Database," [Online], Available at: http://wang.ist.psu.edu/ , last visited August 12th2010.

[14] Suchismita Das, Shruti Garg and G. Sahoo, "Comparison of Content Based Image Retrieval Systems Using Wavelet and Curvelet Transform," The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.4, pp.137-154, August2012.

[15] Nidhi Singhai and Prof. Shishir K. Shandilya, "A Survey On: Content Based Image Retrieval Systems," International Journal of Computer Applications (0975 – 8887) Volume 4 – No.2, pp. 22-26, July2010.

[16] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image Retrieval Using BDIP and BVLC Moments", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 13, No.9, pp. 951-957, September2004.