ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

Improved Feature Extraction Techniques in CBIRS Using Log Gaber

Rajiv Bansal, Sukirti Gupta

Assistant Professor, Dept. of C.S, JMIT Radaur, India

M.Tech Student, Dept. of C.S, JMIT Radaur, India

ABSTRACT: Content based image retrieval(CBIR) is also called by its another name as content based visual information (CBVIR) and query by image content(QBIC) .This application is used for searching digital images in huge databases. Traditional approach were very complicated than the CBIR. This technique captures the result on the basis of images rather than the keywords, words etc.CBIR is necessary as searches rely on metadata are dependent on annotation quality. Here we will use KNN along with the Log Gaber filter feature selection algorithm. We have also validated our result using precision, recall and accuracy.

KEYWORDS: CBIR, , Image Retrieval, Feature Extraction, Log Gaber,KNN.

I.INTRODUCTION

CBIR can efficiently provide images information basedon their content. A single visual feature can be considered only as a part of single perception whereas multiple visual can be considered as an image through different perceptions [1]. Initially built CBIR systems were developed for searching relevant images on basis of texture, colourand some other information. With the development of CBIR systems, the need of user-friendly interfaces also arise .It means inclusion of query methods that permits descriptive semantics, queries where user feedback is involved [2]. In the year 1992 systems were developed when system wasused to describe images retrieval experiments from a database on the basis of colors and shapes [3].Associated problem with algorithms istheir dependency on visual similarity in judging semantic similarity. Yet semantic similarity can be thought of as a higher subjective measure [4].

II. RELATED WORK

Dong-Ho Lee, Hyoung-Joo Kim in their paper a fast content based indexing and retrieval technique by the shape information in large image database present an efficient content-based image retrieval (CBIR) system which employs the shape information of images to facilitate the retrieval process.**Henning Müller, David Bandon,** in their paper a review of content based image retrieval in medical applications presents the availability of large and steadily growing amounts of visual and multimedia data, and the development of the Internet underline the need to create thematic access methods that offer more than simple text-based queries or requests based on matching exact database fields[10].**Hung-Yi Chang** in their paper a semantic learning of content based image retrieval using analytic hierarchy process presents a new semantic learning method for content-based image retrieval using the analytic hierarchical process (AHP) is proposed.**S.H. Kwok, J. Leon Zhao** in their paper Content based Object organization for efficient image retrieval in image databases presents a blob-centric image retrieval scheme based on the blob world representation. The blob-centric based on multidimensional indexing.**Da Deng** in his paper Content based image collection summarization and comparison using self-organizing maps presents progresses made on content-based image retrieval have reactivated the research on image analysis and a number of similarity-based methods have been established to assess the similarity between images[11].**HosseinPourghassem, Hassan Ghassemian** in their paper content based image medical classification using a new hierarchical merging scheme presents aAutomatic medical image classification is a technique for assigning a medical image to a class among a number of image categories. Due to computational complexity[12].**Weiguo Fan** in their paper a genetic programming framework for content based image retrieval presents effectiveness of content-based image retrieval (CBIR) systems can be improved by combining image features or by weighting image similarities, as computed from multiple feature vectors. **J. Ferri, Juan**in their

ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

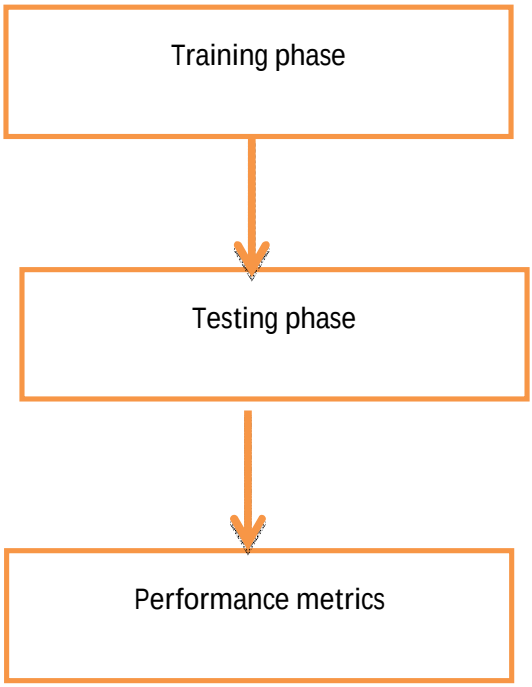
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

paper A naive relevance feedback model for content based image retrieval using multiple similarity measures presents a novel probabilistic framework to process multiple sample queries in content based image retrieval (CBIR)[13].**Marcello Pelillo** in their paper content based image retrieval with relevance feedback using random walks presents a novel approach to content-based image retrieval with relevance feedback, which is based on the random walker algorithm introduced in the context of interactive image segmentation[14].**Reda Alhajj** in their paper Integrating wavelets with clustering and indexing for effective content based image retrieval presents Recent development in technology influenced our daily life and the way people communicate and store data.**M.E. ElAlam** in their paper a new matching strategy for content based image retrieval presents a Adopting effective model to access the desired images is essential nowadays with the presence of a huge amount of digital images.**IDavarGiveki** in their paper content based image retrieval system using clustered scale invariant feature transforms presents a The large amounts of image collections available from a variety of sources have posed increasing technical challenges to computer systems to store/transmit and index/manage the image data to make such collections easily accessible.**Tom Misteli** in their paper A high Content Image Based screening pipeline for the systematic identification of anti-progeroid compounds presents a an early onset lethal premature aging disorder caused by constitutive production of progerin, a mutant form of the nuclear architectural protein lamin A. The presence of progerin causes extensive morphological, epigenetic and DNA damage related nuclear defects that ultimately disrupt tissue and organismal functions[15].

III.PROPOSED WORK

Thousands of images are generated every day, which implies the necessity to classify and access them by an easy and faster way. Classification is an information processing task in which images are categorized into several groups [5]. Image categorization refers to the labelling of images into one of some predefined categories [6].



1. Training Phase

In this step, we have extracted the data from the database. We performs training on 60 % of the images. In this step, we have divided the dataset into two parts; one is train class and another is test class. The test class and train class are divided on the basis of training percentage. The training percentage consists of n-rows from which some values are chosen randomly and then we check with which class it belongs.

ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

2. Testing Phase
In this step user will select the image then extract the features from the selected image. After feature extraction , user will compare the feature with the training phase. After comparison, sorting will be performed, and then user can use euclidian distance. Under Euclidian distance , we will take top 5 members. After doing all these, we will calculate performance metrices (like accuracy, precision, recall). Performance metrices should be tested on multiple images.

3. Performance Metrics: Some of the performance metrics are shown as below:
Accuracy – The accuracy is the proportion of the total number of predictions that were correct. It is determined using the equation:
Acc = $\frac{TP+TN}{P+N}$ **[7]**
Recall - Recall is the proportion of positive cases that were correctly identified, as calculated using the equation :**Recall** = $\frac{TP}{TP+FN}$ **[8]**
Precision – Precision is the proportion of the predicted positive cases that were correct, as calculated using the equation**Precision** = $\frac{TP}{TP+FP}$ **[9]**
where TP – true positive; FN – false negative; TN – true negative; FP – false positive; P – positive; N – negative .

Classifier
In pattern recognition, the **k-Nearest Neighbours algorithm** (or **k-NN** for short) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression: In **k-NN classification**, the output is a class membership. An object is classified by a majority vote of its neighbours, with the object being assigned to the class most common among its k nearest neighbours (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbor.In **k-NN regression**, the output is the property value for the object. This value is the average of the values of its k nearest neighbors.

k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The k-NN algorithm is among the simplest of all machine learning algorithms. Both for classification and regression, it can be useful to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.The neighbors are taken from a set of objects for which the class (for k-NN classification) or the object property value (for k-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. A shortcoming of the k-NN algorithm is that it is sensitive to the local structure of the data. The algorithm has nothing to do with and is not to be confused with k-means, another popular machine learning technique.
The k-nearest neighbour classifier can be viewed as assigning the k-nearest neighbours a weight and all others 0 weight. This can be generalised to weighted nearest neighbours a classifier. That is, where the i^{th} nearest neighbour is assigned a weight $\omega_{ni,}$with $\sum\omega_{nis}= 1$. An analogous result on the strong consistency of weighted nearest neighbour classifiers also holds.

IV.EXPERIMENTS AND RESULT

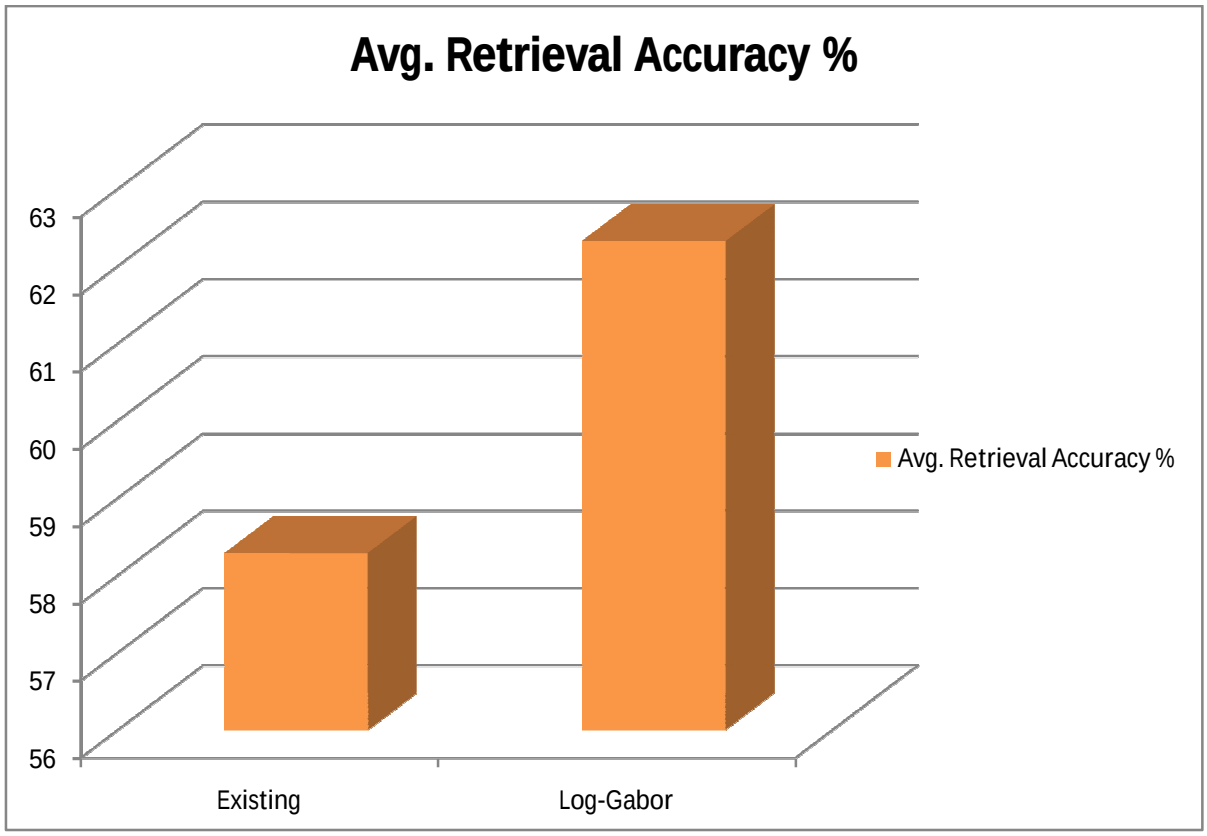
In this we basically compare the retrieval rate of images of Gaber with Log Gaber.Log Gaber giver the better results in comparison of that of Gaber.

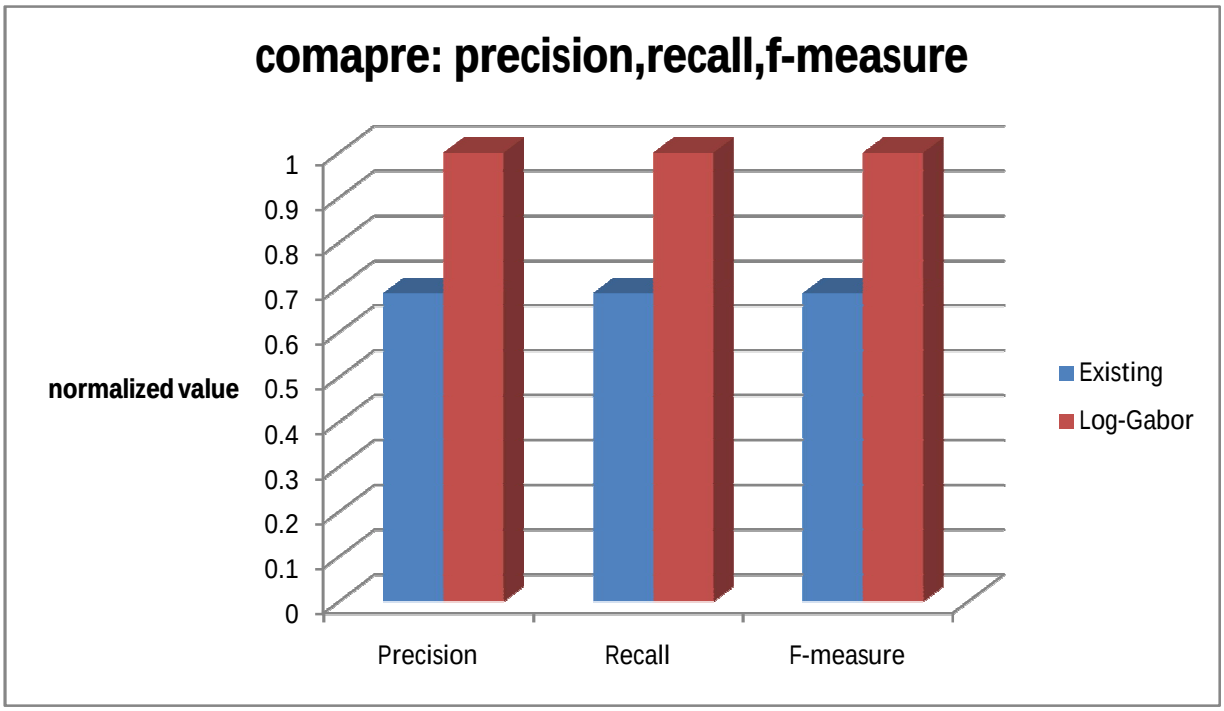
ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016





ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

V. CONCLUSION

Based on the observations during the process of features extraction, training phase, Testing Phase; we concluded that CBIRS images is built through several stages of process. The features are extracted using LOG GABER Filter and images are filtered using K-Nearest Neighbour (KNN) classifier. Based on the analytical results obtained, the amount of accuracy is determined by the number of features matched.

REFERENCES

[1] Science and Technology, an International Journal (2015). Local texton XOR patterns: A new feature descriptor for content based image retrieval
[2]Content-based Multimedia Information Retrieval: State of the Art and Challenges.
[3]Rui, Yong; Thomas S. Huang; Shih-Fu Chang (1999). "Image Retrieval: Current Techniques, Promising Directions, and Open Issues.
[4] Content-Based Image Retrieval - Approaches and Trends of the New Age RitendraDattaJia Li James Z. Wang The Pennsylvania State University, University Park, PA 16802, USA
[5] Zhiwu Lu, Horace H.S Ip. "Image Categorization by Learning with context and consistency", Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, IEEE, 2009
[6] JispaKurian, V. Karunakaran, "A Survey on Image Classification Methods", IJARECE, Volume 1, Issue 4, October 2012
[7]MarcinKorytkowski, LeszekRutkowski, Rafał Scherer, "Fast Image Classification by boosting Fuzzy Classifiers", Institute of Computational Intelligence, Poland, Science Direct, 2015.
[8] Yuxing Hu, LiqiangNie, "An aerial image recognition framework using discrimination and redundancy quality measure", National University of Singapore, Singapore, Science Direct, 2015.
[9] S. Arivazhagan, R. AhilaPriyadharshini, "Generic visual categorization using composite Gabor and moment features", MepcoSchlenk Engineering College, Sivakasi, Tamil Nadu, India, Science Direct, 2015.
10]http://en.wikipedia.org/wiki/Image_retrieval.
[11]BadrinarayanRaghunathan and S.T. Acton,"A content based retrieval engine for remotely sensed imagery", In 4th IEEE Southwest Symposium on Image Analysis and Interpretation, Austin, TX, USA, pages 161-165, April 2000.
[12] BadrinarayanRaghunathan and S.T. Acton,"A content based retrieval engine for circuit board inspection", In Proceedings of the International Conference on Image Processing, Kobe, Japan, pages 104-108, October 1999
[13]. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang and B. Dom, s"Query by Image and Video Content: The QBIC System", IEEE Computer, vol. 28, (1995), pp. 23-32.
[14] A. Pentland, R. Picard and S. Sclaroff, "Photobook: Content-based Manipulation of Image Databases", Proc. SPIE Storage and Retrieval for Image and Video Databases II, SanJose, CA, (1994), pp. 34–47.
s[15] M. Stricker and M. Orengo, "Similarity of Color Images", in Proc. SPIE Storage and Retrieval for Image and Video Databases, (1995), pp. 381-392.