ISSN: 2455 - 1597

ijergs
Open Access

International Journal of Engineering Research and Generic Science (IJERGS)
Available Online at www.ijergs.in

Volume - 5, Issue - 2, March – April - 2019, Page No. 134 - 137

An Productive proposal to Content Based Image Retrieval System

Naveen Tiwari, Bhavya Sareen,Vaishali Sharma

Assistant Professor, Department of CSE, Arya College of Engineering and Research Centre, Jaipur

Abstract

Content-based image retrieval (CBIR) systems are capable to use query for visually related images by identifying similarity between a query image and those in the image database. The CBIR systems can be classified broadly into two classes as Low level feature based system and High level Semantic feature based system. The most common contents are color, texture and shape. K-means clustering is used to group similar and dissimilar objects in an image database into k disjoint clusters whereas neural network is used as a retrieval engine to measure the overall similarity between the query and the images. Relevance feedback is a query modification technique in the field of content-based image retrieval to improve the retrieval performance. Goal is to support image retrieval based on *content* properties (e.g., shape, color, texture), usually encoded into *feature vectors*. One of the main advantages of the CBIR approach is the possibility of an automatic retrieval process, instead of the traditional keyword-based approach, which usually requires very laborious and time-consuming previous annotation of database images.This paper aims to introduce the problems and challenges concerned with the creation of CBIR systems, to describe the existing solutions and applications, and to present the state of the art of the existing research in this area.

Keywords: content-based image retrieval, color, texture, shape, image database, applications, Feature Extraction, Relevance Feedback, k-means.

Introduction

Content-based image retrieval has been an active research area in recent years. The interest in this research area has inspired from the need to search and well manage large volumes of Multimedia information [1, 6]. Content-Based Image Retrieval (CBIR) is defined as a process that searches and retrieves images from a large database on the basis of automatically-derived features such as color, texture and shape.CBIR extracts low-level features which is Inbuilt in the images to present the contents of images. Each image has Visual features such as classified into three main classes: color [2, 11], texture [9, 11] and shape [2, 10] features. Color is an important image feature such as used in Content-Based Image Retrieval [2, 9, 13]. These features have potential to identify objects [10] and retrieve similar images on the basis of their contents. These methods do work very efficient in object recognition and Web searching [14].An efficient and effective query reformulation is essential for finding the relevant images from the database. Relevance feedback (RF) is an interactive process which refines the retrieval results to a particular query by utilizing the user's feedback on previously retrieved images [5, 8]. One of the simplest unsupervised learning algorithms is k-means, it solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set using a certain number of clusters [4].

Related Work

Similarity matching is significant issue in CBIR. So many image retrieval applications are based on shape feature and color feature [2, 10]. As well, lots of others have proposed CBIR method in the literature [2, 5, 8, 11, 12]. Estimating local texture based on pixels of the intensity image and a fuzzy index to point out the presence of major colors [11]. It is based on the texture co-occurrence matrix; a few apprehensible features have been proposed to deduct the comparison cost. They also used the relevance and performance cost. Relevance feedback in Content Based Image Retrieval (CBIR) has been an active field for research [5]. Many schemes and techniques of relevance feedback exist with many assumptions and operating criteria. Yet there exist few ways of quantitatively measuring and comparing different relevance feedback algorithms [8]. K-means clustering for the classification of feature set obtained from the histogram refinement method. Histogram refinement provides a set of features for proposed for Content Based Image Retrieval (CBIR) [2]. They used global histograms for image retrieval, because

of their effectiveness and insensibility to minor changes, are broadly used for content based image retrieval [7]. Color feature plays important role in image retrieval system and comparing all the colors in two images would however be time consuming and difficult problem to overcome this problem they introduced a method of reducing the amount of information. One way of doing this is by quantizing the color distribution into color histograms [3, 12]. Using color histogram easier way for color distribution or they used histogram divide in to different classes for matching.

Proposed Work

Each image has three features Color, Shape and Texture.For fast and improve Image retrieval performance we are using color feature extraction. Using color feature extraction firstly we converted color image into grey level, this is containing values from 0 to 255.

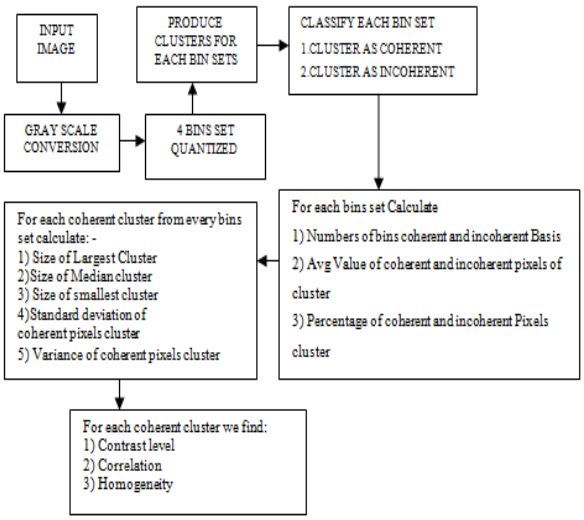

Figure 1: Block diagram of algorithm

After the conversion from RGB to grayscale image, we perform quantization into 4 to reduce the number of levels in the image. We reduce the 256 levels to 16 levels in the quantized image by using uniform

quantization. After the above declared pre-processing stage, we find out the coherent pixels and incoherent pixels. If a pixel is part of a big group of pixels of the same color which form at least 60 percent of the image then that pixel is a coherent pixel and that group is called the coherent group or cluster. Otherwise it is incoherent pixel and the group is incoherent group or cluster.

Then few properties are calculated for each bin. Firstly the numbers of clusters are found for each case, i.e., coherent and incoherent case in each of the bin. Secondly, the Percentages of each bin cluster is

computed. Consequently for each bin, there are six parameters: one each for percentage of coherent pixels

values and Incoherent pixels values, number of coherent pixels values in cluster and incoherent pixels values in cluster, average of coherent pixels values cluster and incoherent Pixels values in cluster.

As we seen in figure 1. Let i is distinguished color, the number of coherent pixels as Ai, the number of coherent bonded components as CAi and the average of coherent bonded component as μAi. Similarly, let the number of incoherent pixels as Aj, the number of incoherent bonded components as CBi and the average of incoherent bonded component as μBi.

For each distinguished color i, the total numbers of pixels bin are Ai+Bi and the color histogram summarizes the image as

$<A_i+Bi,\ldots\ldots\ldots\ldots,An+Bn>$.

After calculation of above features, we consider some additional features based on coherent Pixels clusters only. At this stage, incoherent Pixels clusters are ignored. Four features are selected among the coherent clusters. Few properties are based on the size of the clusters while one is statistical in nature. They are; (i) Size of largest bin value in cluster, (ii) Size of median value bin cluster from bins set, (iii) Size of smallest bin value in cluster from bins set, and (iv) Variance of bins value from bins set. Let us denote the largest cluster in each bin as LAi, the median cluster in each bin as MAi, the smallest cluster in each bin as SAi and Variance of clusters in each bin as VAi.

These features are shown in figure 1.

Again more features are selected bases on above referred features of coherent pixels group or clusters. The following features are selected for retrieval for each of the largest value bin cluster, median of value bins in cluster and smallest value bin in cluster in each of the bin and (i) contrast (ii) homogeneity (iii) correlation. We use set of bins from quantized histogram into four parts and find coherent pixels values and incoherent pixels values distance using Euclidian distance Algorithm and divide into separate cluster which is based on k-means Clustering and next we find out gray level mean, variance, median, standard deviation and various sizes of objects like highest and smallest degree of bins values from image Histogram.

Conclusion

Image retrieval algorithms always use the similarity between the query image and images in image database. The algorithm is based on the features of images such as color and texture. The grayscale values mean variance(in pixel distance), contrast level of image and various sizes of the intensity values are considered as appropriate features for retrieving the grayscale features of image. Proposed work shown that k-means clustering is quite useful for relevant

image retrieval queries. One achievable way for future work is to further improve the cluster picking method by investigating heuristic functions.

References

1. Bing Wang, Xin Zhang, Xiao-Yan Zhao, Zhi-De Zhang, Hong-Xia Zhang ”A Semantic Description For Content-Based Image Retrieval” College Of Mathematics And Computer Science, Hebei University, Baoding 071002, China - 2008.
2. Youngeum An, Junguks Baek AND Minyuk Chang “Classification of Feature Set Using K-Means Clustering From Histogram Refinement Method” (Korea Electronics Technology Institute, South Korea)-2008.
3. Waqas Rasheed, Gwangwon Kang, Jinsuk Kang, Jonghun Chun, and Jongan Park ”Sum of Values of Local Histograms for Image retrieval” - Chosun University, Gwangju, South Korea-2008.
4. Tapas Kanungo, David M. Mount, Christine D. Piatko, Ruth Silverman, AND Angela Y. Wu, “An Efficient K-MeansClustering Algorithm: Analysis AND Implementation” Senior Member, Ieee-2002.
5. Samar Zutshi AND Campbell Wilson “Proto-Reduct Fusion BASED Relevance Feedback IN Cbir”,(Monash University)–-2009.
6. James Z. Wang, Jia Li, Gio Wiederhold, ``SIMPLIcity:Semantics-sensitive Integrated Matching for Picture LIbraries," IEEE Trans. on Pattern Analysis and Machine Intelligence- 2001.
7. C. R. Shyu, et. al, "Local versus Global Features for Content-Based Image Retrieval", IEEE Workshop on Content-Based Access of Image and Video Libraries, 1998.
8. P. Suman Karthik - Jawahar “Analysis Of Relevance Feedback In Content Based Image Retrieval” Center For Visual Information Technology International Institute Of Information Technology Gachibowli, Hyderabad-India, -2006.
9. Zhi-Gang Fan, Jilin Li, Bo Wu, And Yadong Wu “Local Patterns Constrained Image Histograms For Image Retrieval” Advanced R&D Center Of Sharp Electronics (Shanghai) China -2008.
10. Tang li “Developing a Shape-and-Composition CBIR Thesaurus for the Traditional Chinese Landscape” University of Maryland College of Information Science College Park, MD, United States Library Student Journal, July 2007.
11. Sanjoy Kumar Saha, Amit Kumar Das, Bhabatosh Chansa, “CBIR Using Perception Based Texture And Colour Measures”CSE Department; CST Department Jadavpur Univ., India; B.E. College, Unit ISI, Kolkata, India -2003.
12. Greg Pass Ramin Zabih “Histogram Refinement for Content-Based Image Retrieval” Computer Science Department Cornell University-1996.
13. Mohamad Obeid’ Bruno Jedynak Mohamed Daoudi ”Image Indexing & Retrieval Using Intermediate features” Equipe MIRE ENIC, Telecom Lille , Cite Scientifique, Rue G. Marconi, 59658 Villeneuve d’Ascq France-2000.
14. M.S. Lew, N. Sebe, C. Djeraba, And R. Jain, “Content- Based Multimedia Information Retrieval: State Of The Art And Challenges” ACM Transactions On Multimedia Computing- 2006.