Optimized Content Based Image Retrieval System using Colored Features with IWD

Damandeep Singh¹, Nitin Bhagat²

¹Research Scholar at SRI SAI College of Engineering and Technology

²Assistant Professor at SRI SAI College of Engineering and Technology

Abstract: In this paper the main challenge lies in the reduction of the semantic gap between low-level features extracted from the image and high-level user semantics. So, the need of CBIR is motivated. The goal of CBIR is to extract visual content of an image automatically, like color, texture, or shape. This project aims to introduce the problems and challenges concerned with the design and the creation of CBIR systems, which is based on a free Database oriented Image retrieval System. With the help of the existing methods, describe a possible solution how to design and implement a task specific descriptor, which can handle the informational gap between a source and images in a Database, making an opportunity for the efficient search hereby. There are lots of algorithms to optimize space and time for the images that are to be retrieved but they are less efficient and also decreasing the quality of images. The main focus of this thesis is to optimize the images efficiently and effectively. This thesis proposed a new technique for images using Intelligent Water Drop optimization for RGB images.

Keywords: IWD, CBIR, RGB, YCBCR, Histogram

1. Introduction

Content Based Image Retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to retrieve the image in large databases by visual properties.

"Content-based" means that the search will analyze the actual contents of the image. The term 'content' refers to colors, shapes, textures, or any other information that can be derived from the image itself [Jain et al., 2007]. Without the ability to examine image content, searches must rely on metadata such as captions or keywords. Such metadata must be generated by a human and stored alongside each image in the database.

Clustering is a form of unsupervised classification that aims at grouping data points based on similarity. Image Clustering is the process of dividing imageelements into classes or clusters so that items in thesame class are as similar as possible, and items indifferent classes are as dissimilar as possible.

Different measures of similarity may be used depending on the image to place items into classes. Figure 1.2 represents schematic diagram of CBIR using clustering .Similarity measure controls how the clusters are formed for the images. Some examples of measures that can be used as in clustering include distance, connectivity, and intensity.



Figure 1.2: Schematic Diagram of CBIR Using Clustering

Hard Clustering is the clustering in which image is divided into distinct clusters and where each image element belongs to exactly one cluster [O. A. Mohamed Jafar et al.,2013]. On the other hand, Fuzzy Clustering or soft clustering is technique in which image elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that image element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels, and then using them to assign image elements to one or more clusters.

2. Related Work

Iztok Fister Jr, Xin-SheYang and Du san Fister, et al, (2013) "A Brief Review of Nature-Inspired Algorithms for Optimization", They proposed that Nature-Inspired Algorithm is based on 4 categories that is swarm-intelligence-based, bio-inspired, and physics-based and chemistry based, depending on the sources of inspiration. A few algorithms have proved to be very efficient and thus have become popular tools for solving real-world problems.

The purpose of their review is to present a relatively comprehensive list of all the algorithms in the literature, so as to inspire further research.

Navjot Kaur and Parminder Singh, et al, (2013), "An Efficient Edge Detection Approach Based on Pollination Based Optimization" The detection of the edge is one of the important parts in the field of Image Processing. The detection is basically done as per the natural phenomenon of PBO. Edge detection is always study focus in the field of medical image processing and analysis. It is necessary step in medical image processing. Edge detection of noise free pictures is comparatively less complicated, however in most sensible cases the photographs area unit degraded by noise. Edges in photos provide low-level cues, which could be utilized in higher level processes, like object detection, recognition, and classification, furthermore as motion detection, image matching, and trailing. Edges and textures in image are typical samples of high-frequency information. High-pass filters deduct low-frequency image information and therefore enhance high-frequency information like edges. Many approaches to image interpretation measure supported edges. They used bilateral filter to remove noise. And then used pollination based optimization algorithm. This technique achieves better results.

Xin-SheYanga, Mehmet Karamanoglua and Xingshi Heb, et al, (2013). "Multi-objective Flower Algorithm for Optimization" In this paper they proposed a flower pollination algorithm which is a new nature-inspired algorithm, based on the characteristics of flowering plants. They extended the flower algorithm to solve multi-objective optimization problems in engineering. By using the weighted sum method with random weights, they found those multi-objective flower algorithms can accurately find the Pareto fronts for a set of test functions. They then solve a bi-objective disc brake design problem, which indeed converges quickly.

Sanghavi J.B., Bhoyar K.K. and Gawande U.H., et al, (2012) Review of content based image retrieval systems of medical domain. They focused on recent advances in CBIR systems in medical domain like (Magnetic Resonance Imagers) MRI images, (Computed Tomography) CT scan images, (Ultrasound Probes (US), Nuclear Imagers, etc. They also focus on the various feature extraction techniques and algorithm implemented for CBIR systems in different cases of medical domain.

Binitha S and S Siva Sathya, et al, (2012). A Survey of Bio inspired Optimization Algorithms. In this paper they presented a broad overview of biologically inspired optimization algorithms, grouped by the biological field that inspired each and the areas where these algorithms have been most successfully applied. They also provide the Taxonomy and nomenclature of various bio inspired optimization algorithms grouped by the area of inspiration.

Rashmi A. Mahale and Prof. S.D. Chavan, et al, (2012). A Survey: Evolutionary and Swarm Based Bio-Inspired Optimization Algorithms. Theypresented a paper which gives overview of most predominant and successful classes of bio inspired optimization methods involving evolutionary and swarm based algorithms inspired by natural evolution and collective behavior in animals respectively. The algorithm maintains a collection of potential solutions of the problem, which are used to create new potential solutions through the use of operators.

3. Problem Formulation

The increase in computing power and electronic storage capacity has lead to an exponential increase in the amount of digital content available to the users in the form of images and video, which form the bases of many educational, entertainment and commercial applications. Consequently, the search for relevant information in the large space of image and video databases has become more challenging. The main challenge lies in the reduction of the semantic gap between low-level features extracted from the image and high-level user semantics. So, the need of CBIR is motivated. The goal of CBIR is to extract visual content of an image automatically, like color, texture, or shape. This project aims to introduce the problems and challenges concerned with the design and the creation of CBIR systems, which is based on a free Database oriented Image retrieval System. With the help of the existing methods, describe a possible solution how to design and implement a task specific descriptor, which can handle the informational gap between a source and images in a Database, making an opportunity for the efficient search hereby. There are lots of algorithms to optimize space and time for the images that are to be retrieved but they are less efficient and also decreasing the quality of images. The main focus of this thesis is to optimize the images efficiently and effectively. This thesis proposed a new technique for images using Intelligent Water Drop optimization for RGB images.

4. Methodology

4.1 System Architecture



5. Results and Discussions

The proposed system has been designed for the retrieval of images from the database contains a number of images of different classes. In this system images have been retrieved on the basis of different contents available in the image. The features from different images have been extracted using different feature extraction approaches. These approaches provide information about different color components available in the image. The percentage of the particular red, green and blue color has to be measure using RGB model and colorrelorgram. The color moments features have been defined in the required work.

	Content Based Image Retrival System Using IVD	
Operations Select image directory for processing Create DB of image features	03-	 X
Load Dataset	Browse For Folder 25	
Query by sample		
Similarity Vietrica	protemo project Speech_HMM	
Menhaten 💌	 a rahul annitsar a resrch wik_new 	
Num of images returned	BSSI b sanjeli - Copy	
1 💌	Faider: senjol - Copy	
Content Based Retrival	Make New Folder OK Cancel	
Parameters	01-	
Percision	Precision Recall	

Figure 4.1: Feature extraction of database images

Represents the extraction of features from various database images. The color, shape, energy, histogram and transformation features have to be computed. This figure represents that to define the path of the database where that has been stored. After getting path the number of images has been count available in the folder and features have been extracted for a single image and these features extracted have to be computed for all the images available in the database and these features are stored in a file to load the dataset for the matching purposes. This dataset contains the information of different features computed from the each image available in the dataset. Single feature has been computed for single iterations available in the image. All the features have to be stored in a file. File having a single row contains name of the images and the feature values for particular image.

		Content Based Image Retrival 1	System Using	WD		
Operations	1					71
Create DB of image features	0.9					
	A Select the Data	azet			8	
Load Dataset	Look in:	🎉 sanjoli - Copy	×	+ 🗈 💣 📼 •		
Guery by sample	Ca.	Name		Date modified	Туре	
Browse for image	Recent Places	images (201		3/31/2015 7:18 PM	File folder	
Similarity Motrics		an deteset		10/31/2014 2:55 PM	Microsoft	
Washatan 💌	Desktop	31 optimageFeature		3/31/2015 7:22 PM 10/31/2014 10:42	Microsoft Microsoft	
Num of images returned	Libraries	guery@mageFeature		3/31/2015 7:22 PM	Microsoft	
1	1					
	Computer					
Content Based Retrival	Network					
		* [
Parametera		Fie name:		-	Open	
		Hies of type: [MAT-files (* mat)		-	Cancer	 -
Percision						

Figure 4.2: Load Dataset for matching of query image

Fig. 5.4 represents the user interface control window which requires the dataset to be loaded which contains the information for different feature values for the images. Loading of the dataset having optimized feature value is further utilized for similarity matching with the database images of the query image.

	Content Based Image Retrival System U	sing WD	
Operations	1r		
Select image directory for processing			
Create DB of image features	0.9 -		
Load Dataset	0.8 -		
Query by sample	0.7 -		
Browse for image	Dataset loaded successfulyt		
Simlarity Metrics	OK		
Manhatan 💌	0.5 -		
Num of images returned	0.4 -		
1	0.3 -		
Content Based Retrival	0.2 -		
	0.1 -		
Parameters			
	Precision Recall		
Percision			

Figure 4.3: Dataset loaded successfully

Fig. 4.3 represents a message box that display a message when the dataset is loaded successfully for the processing of the content based image retrieval system on the basis of different types of features. These features contain the values for features related to the different types of images. These feature values undergoes for similarity matching of the query image with database images.

	Content Based Image Retrival System Using M/D	
Operations	10	
Select image directory for processing		
	0.9 -	
Create DB of image features		
Load Dataset	0.8 -	
Course in a second	0.7	
Browse for image	Preceed with the query by executing the cbir button	
Endersky Making	0.6 -	
annar ty month	OK	
Manhatan 💌	0.5	
Num of images returned		
	0.4	
1	0.3-	
Control Based Betrical	0.2 -	
	0.1 -	
Parameters		
	Precision Recall	
Percision		

Figure 4.4: Query image for matching

Fig. 4.4 represents that to load a query image for the retrieval of different images from the database image on the basis of the feature values. After loading a query image different features have been computed for the image and these features are stored in the temporary file for matching purpose.



Figure 4.5: Retrieved images on the basis of Manhattan distance

Fig. 4.5 represents the graphic user interface contains choice for the user to select the type of the distance that has to be used for the matching of the query image features with the database image features. In this figure images have been retrieved on the basis of Manhattan distance classifier that computes the distance between different features values of the database image and query image. By selecting the option for the different number of images that have to be extorted

International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

from the database images. Number of images selected for retrieval of the images that has minimum distance with the query image have been selected and extorted from the database. And these extracted images are displayed on the axis.



Figure 4.6: Retrieved images on the basis of Euclidian distance

Fig. 4.6 represents the graphic user interface contains choice for the user to select the type of the distance that has to be used for the matching of the query image features with the database image features. In this figure images have been retrieved on the basis of Euclidian distance classifier that computes the distance between different features values of the database image and query image. By selecting the option for the different number of images that have to be extorted from the database images. Number of images selected for retrieval of the images that has minimum distance with the query image have been selected and extorted from the database. And these extracted images are displayed on the axis.



Figure 4.7: Confusion Matrix for purposed system

This figure represents confusion matrix for the purposed system that has been developed on the basis of different images that has been divided into different classes.

6. Conclusion

The content based image retrieval system is the application of computer vision to retrieve the image in large databases by visual properties. On the basis of content, extracted from an image, various types of image retrieval systems are used. The content used in image can be color, texture, shape, histogram etc. In our proposed system the image is retrieved using such features. In the proposed system, many features like color, texture, shape, correlogram etc. are used to extract image from a particular database. We have used Wang dataset for our CBIR system.After computation of these coefficients of the image the mean and standard deviation of the coefficients has been calculated as feature values. These feature matrix optimization process has been done by using pollination based optimization (PBO). Euclidean dot product and manhattans distance formula computes the distance between different images available in the database. Minimum the distance with a particular image is maximum matched image with the query image. In the CBIR system number of images has to be extracted from the database image. The images with minimum distance can be extracted from these database images. These are maximum matched image on the basis of contents available in query image.

References

- [1] Abbott, R. (2005). Challenges for Bio-inspired, computing. *The Proceedings of the Bio GEC Workshop*, *GECCO*.12-22
- [2] Abbass H.A., Sarker R. (2002). The Pareto differential evolution algorithm. *International Journal on Artificial Intelligence Tools*. 531-552.
- [3] A.Gupta, R.Jain (1997). Visual Information Retrieval. Communication. *ACM*, *no.5*, 70–79.
- [4] Anick, P.G. and Vaithyanathan S. (1997). Exploiting clustering and phrases for context-based information retrieval. 20th ACMSIGIR international conference on research and development in information retrieval. 314–323.
- [5] A.R. Mehrabian, C. Lucas (2006). A novel numerical optimization algorithm inspired from weed colonization, *Ecological Informatics*. 355–366.
- [6] A. Z. Chitade (2010). Colour Based Image Segmentation Using K-Means Clustering. International Journal of Engineering Science and Technology Vol. 2(10), 5319-5325.
- [7] A. Jain, R. Muthuganapathy, and K. Ramani (2007). Content-Based Image Retrieval Using Shape and Depth from an Engineering Database. ISVC 2007, Part II, LNCS 4842, Springer-Verlag Berlin Heidelberg. LNCS 4842. 255–264.
- [8] Borlund, P., P. Ingwersen (1997). The Development of a Method for the Evaluation of Interactive Information Retrieval Systems. *Journal of Documentation, Vol. 53*. 225-250.
- [9] Binitha S, S Siva Sathya (2012). A Survey of Bio inspired Optimization Algorithms.*International Journal*

sr.ne

of Soft Computing and Engineering (IJSCE). ISSN: 2231-2307.137-151.

- [10] B. Chandra Mohan, R. Baskaran (2012). A survey: Ant Colony Optimization based recent research and implementation on several engineering domain. *ELSEVIER Expert Systems with Applications*. 418–462.
- [11] Melas and S. Wilson (2002). Double Markov Random Fields and Bayesian Image Segmentation. *IEEE Trans. Signal Processing, vol. 50, no.* 2, 357-365.
- [12] Derek Hoiem, Rahul Sukthankar, Henry Schneiderman, Larry Huston (2004) .Object-Based Image Retrieval Using the Statistical Structure of Images. Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference. 490-497.
- [13] Dharani, Aroquiaraj, (2013). A survey on content based image retrieval. International Conference 2013 on Pattern Recognition, Informatics and Mobile Engineering (PRIME).485 – 490.
- [14] E. M. Arkin, L.P. Chew, D.P. Huttenlocher, K. Kedem, and J.S.B. Mitchell (1991). An efficiently computable metric for comparing polygonal shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 13, no.* 3. 209-226.
- [15] F.Jing, M.Li, H.-J.Zhang, and B.Zhang (2004). An Efficient and Effective Region-Based Image Retrieval Framework. *IEEE Trans. Image Processing*, 699–709.

Online): 23,9'