iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** III **Month of publication:** March 2018

DOI: http://doi.org/10.22214/ijraset.2018.3691

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

Review for CBIR System for Retrieving Images from Enormous Databases using Dissimilar Algorithms

Rohan Verma[1], Pooja Verma[2], Varun Barthwal[3]
[1, 2, 3] HNB Garhwal University Uttarakhand Technical University HNB Garhwal University Central University

Abstract: In order to improve the speed and accuracy of image retrieval. More and more images have been generated in digital form around the world. There is a growing interest in finding images in large collections or from remote databases. In order to find an image, the image has to be described or represented by certain features. The shape is an important visual feature of an image. Searching for images using shape features has attracted much attention. There have been introduced many shape representation and description techniques. This paper is presenting different algorithm which originates expected results. Images have always been considered an effective medium for presenting visual data in many applications of industry and academia. With the development of technology, a large number of images are being generated every day. Therefore, managing and indexing of images become essential in order to retrieve similar images effectively. After the completion of pre-processing, it will compare with the query image. In this paper the study of different approaches is discussed.

Keywords - CBIR, Color model, shape feature, PSO, k-mean clustering methods, classification distance and neighbor density, SVM.

I. INTRODUCTION

A. Content-based image retrieval

CBIR with relevance feedback a major task in the CBIR systems is the similarity matching between the query image and the retrieved images. Unfortunately, the gap between high-level concepts and low-level features, as well as the subjective perception for the visual content by the human beings, result in a significant mismatch between the retrieval results judged manually and by the computers. To improve the retrieval precision, human interactions are usually involved. Relevance feedback is an interactive process which integrates users' evaluation of the retrieval results. Typically, the relevance feedback technique includes an interactive scoring system to evaluate the past retrieval results to improve the subsequent content retrieval. Image retrieval is the field of the study that concerned with looking, browsing, and recovering digital images from an extensive database. CBIR is viewed as a dynamic and quick advancing research area in image retrieval field. It is a technique for retrieving images from a collection by similarity. The retrieval based on the features extracted automatically from the images themselves. Many of CBIR systems, which is based on features descriptors, are built and developed. A feature is defined as capturing a certain visual property of an image. A descriptor encodes an image in a way that allows it to be compared and matched to other images. In general, image features descriptors can be either global or local. The global feature descriptors describe the visual content of the entire image, whereas local feature describes describe a patch within an image (i.e. a small group of pixels) of the image content. The superiority of the global descriptor extraction is the increased speed for both feature extraction and computing similarity. However, global features still too rigid to represent an image. Particularly, they can be oversensitive to location and consequently fail to identify important visual characteristics. The term CBIR refers to the process of retrieving similar images from a large collection of image database. The image retrieval is done on the basis of similarity matching between query image and database images. Content-based image retrieval (CBIR) is a technique for retrieving images on the basis of automatically-derived features such as color, texture and shape. The architecture of a CBIR system can be understood as a basic set of modules that interact. Content-based image retrieval (CBIR) is a technique in which images are indexed by extracting their low-level features and image retrieval is only based upon these indexed image features. In an effective image retrieval system, the user poses a query and the system should find images that are somehow relevant to the query. Thus, a way of representing the query, a way of representing images, and a way of comparing a query and an image are needed. This kind of approach is known as querying by content. With the development of the Internet, and the availability of image capturing devices such as digital cameras, image scanners, the size of digital image collection is increasing rapidly.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

Efficient image searching, browsing and retrieval tools are required by users from various domains, including remote sensing, fashion, crime prevention, publishing, medicine, architecture, etc. The basic way of image retrieving is as shown in figure 1.

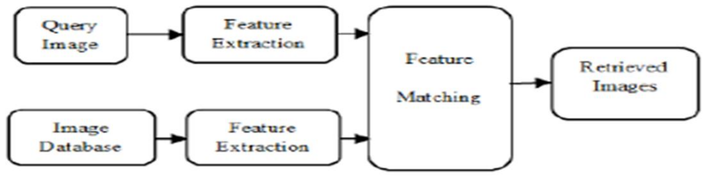

Figure 1: The CBIR system

II. CONTENT-BASED IMAGE RETRIEVAL APPROACHES

A. Color Model

The system is more sensitive to color information than grey values of images. Color is one of the most widely used low-level visual features and is invariant to image size and orientation. Generally color features are extracted using the color histogram technique. The main issues regarding the use of color histograms for indexing involve the choice of color space and quantization of the color space. When a perceptually uniform color space is chosen uniform quantization may be appropriate.

$h_{A,B,C}(a, b, c) = N.Prob(A=a, B=b, C=c)$

In these studies the developments of the extraction algorithms follow a similar progression:

1) Selection of a color space
2) Quantization of the color space
3) Computation of histograms,
4) Derivation of the histogram distance function
5) Identification of indexing shortcuts.

There are used many more color models like RGB, CMY, YIQ, YUV, HSI, HSV. In RGB, pixel color information is R, G, B. When mean color for RGB is found, that is (sum of that component for all pixels) divided by a number of total pixels. Color Histogram is a standard demonstration of the color characteristic in CBIR systems. It is very efficient in the description of both local and global features of colors. This computes the chromatic information and invariant of the image along the view axes for translation and rotation, when the large-scale image database computes histogram, its efficiency is not satisfactory and to overcome this conflict joint histogram technique is introduced. A color histogram is frequency count of each color. Color histograms are a fundamental technique for retrieving images and extensively used in CBIR system. The color space has segmentation, for every segment the pixels of the color within its bandwidth is counted, which demonstrates the relative frequencies of the counted colors. We use the RGB color space for the histograms. Only minor differences have been observed with other color spaces for the histogram. Color Histogram H(m) is a distant probability function of the image color.

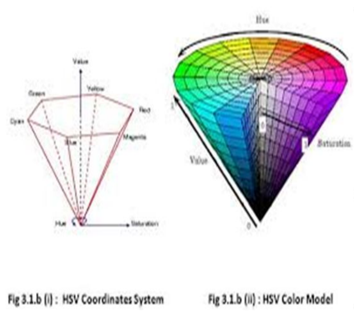

Figure 2: HSV color model

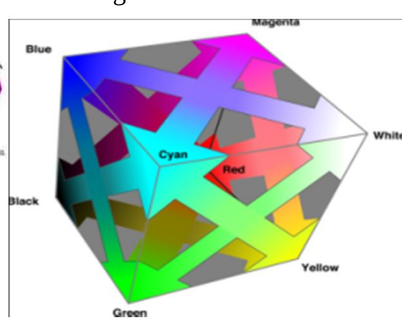

Figure 3: CMY color model

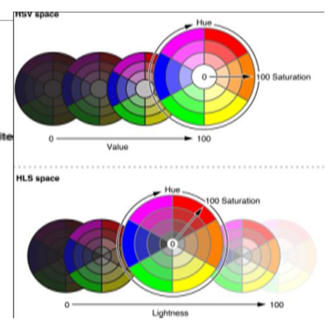

Figure 4: HSV color model

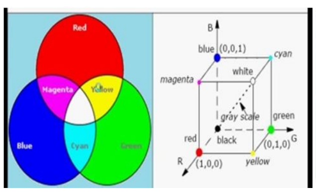

Figure 5: RGB color model

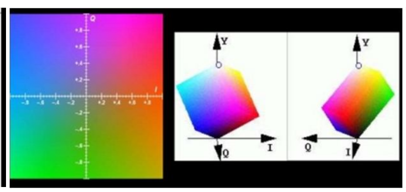

Figure 6: YIQ color model

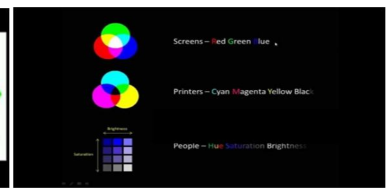

Figure 7: color models

B. Shape

The shape is an important and most powerful feature used for image classification, indexing and retrievals. Shape information extracted using a histogram of edge detection. The edge information in the image is obtained by using the Canny edge detection. In shape, the two edge detection techniques are used in that the comparison between both of these techniques has done. Shape representations can be generally divided into two categories:

Boundary-based

Region-based.

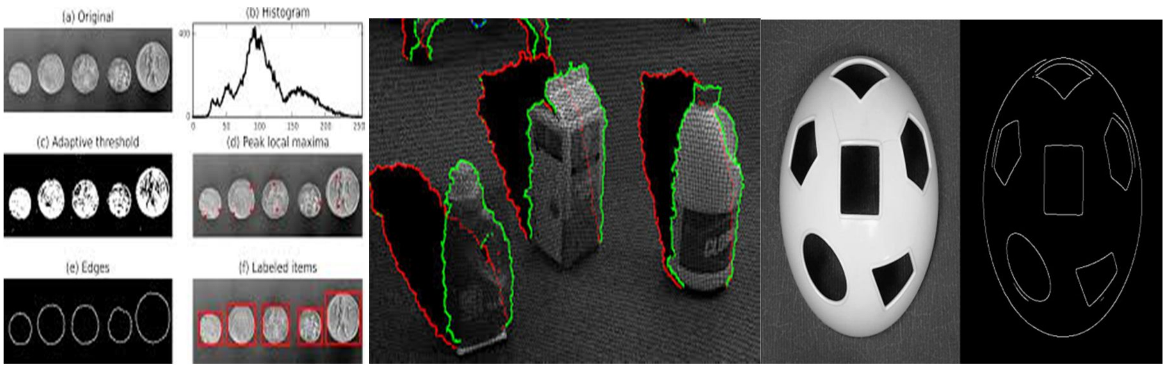

Figure 8: Boundary-based and Region-based Images

C. PSO

It became obvious during the simplification of the paradigm that the behavior of the population of agents is now more like a swarm than a flock. The term swarm has a basis in the literature. In particular, the authors use the term in accordance with a paper by Millonas, who developed his models for applications in artificial life, and articulated five basic principles of swarm intelligence. First is the proximity principle: the population should be able to carry out simple space and time computations. Second is the quality principle: the population should be able to respond to quality factors in the environment. The third is the principle of diverse response: the population should not commit its activities along excessively narrow channels. Fourth is the principle of stability: the population should not change its mode of behavior every time the environment changes. Fifth is the principle of adaptability: the population must be able to change behavior mode when it’s worth the computational price. Note that principles four and five are the opposite sides of the same coin. The particle swarm optimization concept and paradigm presented in this paper seem to adhere to all five principles. Basic to the paradigm are n-dimensional space calculations carried out over a series of time steps. The population is responding to the quality factors pbest and gbest. The allocation of responses between pbest and gbest ensures its diversity of response. The population changes its state (mode of behavior) only when gbest changes, thus adhering to the principle of stability. The population is adaptive because it does change when gbest changes. The term particle was selected as a compromise. While it could be argued that the population members are mass-less and volume-less, and thus could be called “points,” it is felt that velocities and accelerations are more appropriately applied to particles, even if each is defined to have arbitrarily small mass and volume. Further, Reeves discusses particle systems consisting of clouds of primitive particles as models of diffuse objects such as clouds, fire and smoke. Thus the label the authors have chosen to represent the optimization concept is particle swarm. Particle swarm optimization is an extremely wimple algorithm that seems to be effective for optimizing a wide range of functions. We view it as a]mid-level form of A-life or biologically derived algorithm, occupying the space in nature between evolutionary search, which requires eons, and neural processing, which occurs on the order of milliseconds. Social optimization occurs in the time frame of ordinary experience - in fact, it is an ordinary experience. In addition to its ties with A-life, particle swarm optimization has obvious ties with evolutionary computation. Conceptually, it seems to lie somewhere between genetic algorithms and evolutionary programming. It is highly dependent on stochastic processes, like evolutionary programming. The adjustment toward pbest and gbest by the particle swarm optimizer is conceptually similar to the crossover operation utilized by genetic algorithms. The term particle was selected as a compromise. While it could be argued that the population members are mass-less and volume-less, and thus could be called ‘points’, it is felt that velocities and accelerations are more appropriately applied to particles, even if each is defined to have arbitrarily small mass and volume. Further, Reeves discusses particle systems consisting of clouds of primitive particles as models of diffuse objects such as clouds, fire and smoke. Thus the label the authors have chosen to represent the optimization concept is particle swarm.

D. k-Mean clustering Method

Clustering is a way of grouping together data samples that are similar in some way according to some criteria that we pick its form of unsupervised learning So, it's a method of data exploration–a way of looking for patterns or structure in the data that are of interest. Clustering algorithms are generally used in an unsupervised fashion. They are presented with a set of data instances that must be grouped according to some notion of similarity. The algorithm has access only to the set of features describing each object; it is not given any information as to where each of the instances should be placed within the partition. K-means clustering is a method commonly used to automatically partition a data set into k groups. It proceeds by selecting k initial cluster centers and then iteratively refining the results. The algorithm converges when there is no further change in assignment of instances to clusters.

E. Classification nearest-Neighbor Velocity Matching

A satisfying simulation was rather quickly written, which relied on two props: nearest-neighbor velocity matching and "craziness." A population of birds was randomly initialized with a position for each on a torus pixel grid and with X and Y velocities. At each iteration a loop in the program determined, for each agent (a more appropriate term than the bird), which other agent was its nearest neighbor, then assigned that agent's X and Y velocities to the agent in focus. Essentially this simple de created asynchrony of movement. Unfortunately, the flock quickly settled on unanimous, unchanging direction. Therefore, a stochastic variable called craziness was introduced. At each iteration some change was added to randomly chosen X and Y velocities. This introduced enough variation into the system to give the simulation an interesting and "lifelike" appearance, though of course the variation was wholly artificial.

F. SVM

SVM is supervised machine learning technique. It shows the image database as two sets of vectors in a high or infinite-dimensional space. It relies on a fundamental principle, which is called a maximum margin classifier. A maximum margin classifier is a hyperplane, which separates two 'clouds' of points at equal distance. The margin between the hyperplane and the clouds is maximal. SVM built a hyperplane or set of hyperplanes that increases the margin among the images that are relevant and not relevant to the query. The goal of SVM classification technique is to find an ideal hyperplane to separate the irrelevant and relevant vectors using maximizing the size of the margin between both classes.

III. CONCLUSION

From the literature survey it is concluded that a wide variety of CBIR algorithms have been proposed in different papers. The selection feature is one of the important aspects of Image Retrieval System to better capture user's intention. It will display the images from the database which are the more interest to the user. The purpose of this survey is to provide an overview of the functionality of content-based image retrieval systems. Most systems use color and texture features, few systems use shape feature, and still less use layout features. CBIR, Color model, PSO, k-mean clustering methods, Classification nearest-neighbor velocity matching, SVM have been used extensively in various areas to improve the performance of the system and to achieve better results in different applications. Content-based image retrieval is currently a very important and active research in the field of multimedia databases. Since the explosive growth of image data in the large image archives need a more précised retrieval techniques to find the similar images. In this paper the intensity scale measurement is used for the color histogram to retrieve the similar images from the database. The edge detection techniques are used to detect the edges of the image database. And many more are used for better results. So if it is required to get the perfect results sometimes multiple approaches should be used for more perfection.

REFERENCES

[1] D. Patidar, N. Jain, and A. Parikh, "Performance analysis of artificial neural network and K Nearest neighbors image classification techniques with wavelet features," in Proceedings of IEEE International Conference on Computer Communication and Systems, 2014, pp. 191-194.
[2] Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: ideas, influences, and, trends of the new age. ACM Comput. Surv. **40**(2), 1–60 (2008).
[3] D. L. Olson and D. Delen, "Advanced data mining techniques," Springer Science & Business Media, 2008, pp. 138-145.
[4] Y. Chen, J. Z. Wang, and R. Krovetz, "Content-based image retrieval by clustering," in Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval, 2003, pp. 193-200.
[5] Gurpreet Kaur and Mnupreet Kaur (2015), Content-Based Image Retrieval using BPNN and K-Mean Algorithm, International Journal of Science and Research (IJSR), Volume 4, Issue 1.
[6] Ramadan, Rabab M., and Rehab F. Abdel-Kader (2009), Face recognition using particle swarm optimization-based selected features, International Journal of Signal Processing, Image Processing and Pattern Recognition, 2.2: 51-65.
[7] Canny, J., "A Computational Approach to Edge Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, 1986.

[8] Gaurav Jaswal Amit Kaul, " Content-Based Image Retrieval ", National Conference on Computing, Communication and Control, A Literature Review, National Institute of Technology, Hamirpur-177001, Himachal Pradesh(India).

[9] Nidhi Singhai, Prof. Shishir K. Shandilya, "A Survey On: Content-Based Image Retrieval Systems ", International Journal of Computer Applications (0975 – 8887) Volume 4 – No.2, July 2010.

[10] Pooja Verma, Manish Mahajan, "Retrieval of better results by using shape techniques for content-based retrieval", IJCSC, Vol. 3, No.2, January-June 2012, pp. 254-257, ISSN: 0973-7391.

[11] Mit Patel, Keyur Brahmbhatt, Kanu Patel, " Feature-based Image retrieval based on clustering and non-clustering techniques using low-level image features", International Journal of Advanced Engineering and Research Development (IJAERD) Volume 1, Issue 3, April 2014, e-ISSN: 2348 - 4470, print-ISSN:2348-6406.

[12] Heng Chen, Zhicheng Zhao "an effective relevance feedback algorithm for image retrieval" 978-1-4244-6853- 9/10/ 2010 IEEE.

[13] Xiang-Yang Wang, Hong-Ying Yang, Dong-Ming Li "A new content-based image retrieval technique using color and texture information", Computers & Electrical Engineering, Volume 39, Issue 3, April 2013, Pages 746-761.

[14] Priti Maheswary, Dr. Namita Srivastava (2009), Retrieval of Remote Sensing Images Using Colour & Texture Attribute, International Journal of Computer Science and Information Security, Vol. 4, No. 1&2.

[15] Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, vol. IV, pp. 1942–1948. IEEE, Piscataway (1995).