

Analysis of an Effective CBIR Image Extraction using P2PN Networks

¹V. S. N. Kumar Devaraju, ²v. Subba Ramaiah, ³sirasani Srinivasa Rao

¹Asst.Professor, Department of ECE, Mahatma Gandhi Institute of Technology Hyderabad, Telangana, India.

²Assistant Professor, Department of CSE, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India.

³Associate Professor, Department of ECE, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India.

¹dvsnkumar_ece@mgit.ac.in , ²vsubbaramaiah_cse@mgit.ac.in , ³ssrinivasarao_ece@mgit.ac.in

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract: Content-based image retrieval (CBIR) in peer-to-peer (P-P) framework utilizes visual substance of image like shape, color, & spatial layout, & texture to signify & list the image. The disseminated nature of these methods, whereas nodes have been commonly placed across networks, inherently hinders proficient data recovery. We deliberate the retrieval & searching of data, which will be dispersed on network peers. Our method constructs on unstructured P2P frameworks & utilizes local information. The cause for utilizing unstructured P2P frameworks will be that they execute very small requests on distinctive nodes & might simply accommodate nodes of fluctuating power active research in CBIR is equipped towards improvement of approaches for interpreting cataloging, examining, & indexing image database. The response quality is intensely reliant on decision of strategy utilized to produce similarity measures & feature vectors for examination of features; we suggested a method that incorporates benefits of diverse other methods to enhance the accuracy & presentation of retrieval. In this manuscript, we suggested the diverse image properties.

Keywords: Content-Based Image Retrieval, Peer to Peer system, Color lookup table, support vector machine (SVM), machine learning (ML), Artificial Intelligence (AI)

1 Introduction

The P2P networks that have been formed by similarly privileged nodes linking to another in self-organizing method is the main significant architecture for data sharing. The famous P2P file-sharing networks like eDonkey1 count a large number for clients [1] and many files. Dissimilar to web pages that basically comprise of textual documents like blog articles, news, multimedia files performs a principal role in many P2P networks [2]. The global infrastructure of computers and networks makes exciting chances for accumulating much data for sharing resources & computers in original scale. In past few years, P2P method has become a much commanding for increasing sharing resources & Internet scale file systems over large-scale geographical regions. We have been utilizing the CBIR concept that describes a method that utilizes visual contents to search images from large scale image databases as per user interests.

The image retrieval will be procedure of retrieving the image from vast dataset. Retrieving images in our peer to peer network is an important challenge. So the overlay network of nodes (peers) is constructed on top of heterogeneous networks. Overlays have been deployable & flexible. The initial wave of P2P methods executed unstructured P2P overlays in that no global structure is handled. For searching the messages, data have been sent over numerous hops from one peer to another with every peer answering to queries for data, it has stored locally. The unstructured P2P methods are better to store & share documents, due to their decentralized nature permits simple additions, upgrades, enhanced storage, & provide fault-tolerant assets through utilization of caching & replication. The P2P infrastructure has diverse communication & processing abilities, located across diverse areas. As an outcome, retrieving images over such Internet-scale environments is subject to better differences because of excessive resource consumption, unpredictable communication latencies, & varying resource accessibility.

2 Related Work

A.W. M. Smeulders, et al, [1] depicts the content-based such that search will examine actual contents of image rather than metadata like tags, keywords, explanations related with image. The word content mentions to textures, shapes, color, etc might be resultant from image itself. Moreover, humans physically enter keywords for images in large database might be incompetent, costly. Therefore, a system might filter images based on their content and offer best indexing and return accurate outcomes.

Chun et al. [2] suggested a CBIR technique relied on effective mixture of texture features & multi resolution color. The auto-correlograms features & colors of saturation component images in HSV color space have been utilized.

S. Tong and E. Chang [3] depicts that SVM have been supervised learning methods connected with learning methods, which examine data & identify patterns utilized for regression & categorization examination. The main SVM takes a group of input data & calculates for every provided input that of 2 probable classes creates the output, making it non-probabilistic binary linear classifier. Provided a group of training instances, every marked as belonging to 1 of 2 groups, SVM training method builds a method, which assigns novel instances into 1 classification or alternate. The SVM method will be depiction of instances as points in space, which instances of differentiate classifications have been partitioned by clear gap, which is as wide as probable.

GulfishanFirdose Ahmed, et al, [4] describes the fundamental modules of content-based image retrieval method. The image retrieval models have been based on texture, color, shape are analyzed and compared. In numerous zones of government, commerce, & hospitals, extensive groups of digital images have been being generated. Some of these groups have been digitizing product existing groups of drawings, analogue photographs, prints, diagrams, & paintings. Typically, the only method of searching these groups was by keyword indexing. The digital image databases however, open approach to content-based searching.

X. Zhou and T. et al, [5] examine the nature of relevance feedback issue in continuous illustration space in context of multimedia data recovery. The idea behind relevance feedback may be to take outcomes, which have been firstly returned from a provided query & to utilize data if or not the outcomes have been relevant to execute a novel query.

3 Methodology

In our proposed approach rather than searching the image based on the tag, keywords. Here we are giving image as an input query image. Based on input query image similar type of images is being retrieved. Image features are extracted based on color, shape. SVM is used as a classifier which results in a higher accuracy. As we are using peer to peer system once the input query image is selected, we can get the output images from the other peer. In particular, peers are used to diminish the network distance from their neighboring nodes.

The Properties for extraction the image features is as follows:

3.1 Color Feature Extraction

A color image might be signified utilizing 3 primaries of a color space. Since RGB color space cannot be seen directly by the human eye so HSV color space has been used in this method. The HSV is an intuitive color space, which can be seen directly to visual perception, and it will be general for image retrieval methods. Hue is utilized to represent a pure color, whereas saturation represents the number of white light included to a pure color.

The subsequent 2 features to signify global properties of image might be utilized. The mean of pixel colors depicts principal color of image, & standard deviation of pixel colors signifies difference of pixel colors in image. The deviation degree of pixel colors in image will be named the color complexity of image. Color correlogram gives the information about the features of colors. It comprises spatial color correlations that defines global distribution of local spatial correlation of colors and very easy to compute. The color moment feature will be utilized to discriminate images relied on their color features and it will be provides similarity of color measurement among images. Then similarity values have been compared with images values of stored in the database for image retrieval.

3.2 Texture Feature extraction

Texture will be the visual patterns, which have homogeneity assets that do not outcome from presence of intensity or single color. The texture descriptors are used to calculate the properties like coarseness, smoothness, & regularity. The statistical methods yield texture classification of course, smooth, & grainy etc. The texture gives significant data in image classification as it depicts content of numerous real-world images like trees, fruit skin, bricks, clouds, & fabric. Therefore, texture is a significant feature in describing high-level semantics for image retrieval determination. The texture feature extraction methods are

- Texture boundary detection, Gabor Wavelet Transform, Color texture, Texture classification, Fourier descriptors.

3.3 Shape Feature Extraction

Shape features of images are extracted by using connected component method. The shape feature gives semantic data about an image. The shape features have been normally depicted utilizing area of an image. The

shape features accuracy relies upon segmentation method utilized to separate an image into expressive objects. Nevertheless, robust & quick separation is problematic to attain. This limits the shape features only to retrieval applications whereas area or objects of images have been accessible. The shape descriptors have been separated into 2 kinds: region based & boundary based descriptor. Few of boundary based description methods are polygonal approximations, finite element method, chain codes, & Fourier descriptor. The region based descriptors are area & statistical moment. A good shape feature must be invariant to rotation, translation, and scaling.

4 System Design

The figure1 describes that the system comprises of 2 phases that is training & testing phase. In training phase, there is a creation of the dataset and in the testing phase we are giving image as an input query image. Here first the user select the image from the database then there is a generation of the dataset once the dataset is generated then the dataset is stored in database. Next when user goes to the testing phase here the user gives the input query image then there is a selection of the dataset once the dataset is selected then there is feature extraction.

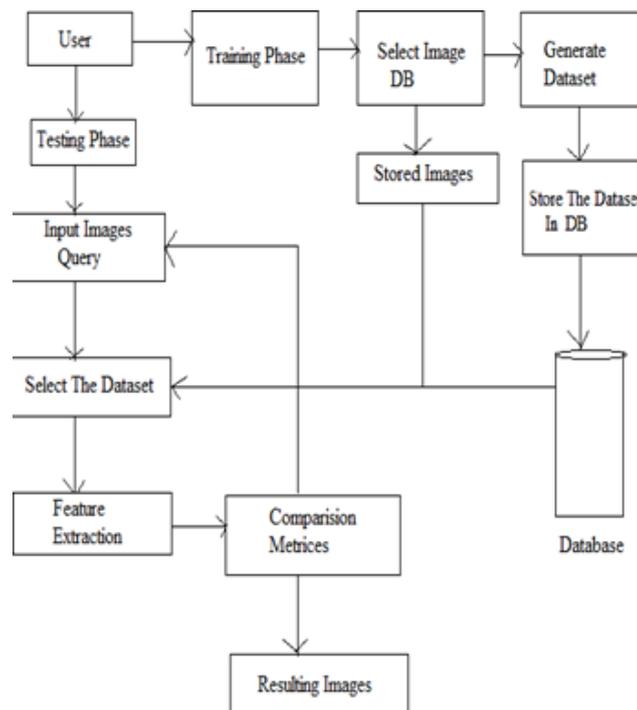


Fig 1: System Architecture

The features extracted from images directly lead to outcomes. Few basic features of images are color features, shape and texture features. Once the features has been selected then there is a metrics selection here we are using the support vector machine classifier that predicts target value of data instances in testing set that have been provided only the attributes. Next step is to retrieve the same type of images from the given input query image. As we are using peer to peer network the above same steps are followed but the output images retrieved are on another peer.

4.1 Classification by SVM

The SVM is a supervised ML technique, which evaluate the data & recognize the patterns, utilized for categorization. The benefit of this method will be categorize the input query object relies on training instances & feature vectors. The features trained and testing images have been categorized by utilizing SVM classifier to provide the outcome to user SVM will be beneficial method for data classification. The ML will be deliberated as subfield of AI & it will be concerned with improvement of methods that allow computer to learn. Over period of time, numerous methods are established for ML tasks. The statistical learning theory gives a system for surveying the issue of gaining information, making decisions, making predictions, from group of data.

SVM will be beneficial method for data categorization. Even though it will be deliberated, which its Neural Networks are simpler to utilize than this, nevertheless, occasionally unacceptable outcomes have been attained. Nevertheless, the classification task typically includes with testing & training data that comprise of few data

instances. Every instance in training set comprises 1 target values and numerous attributes. The aim of SVM will be to generate a method that calculates target value of data cases in testing set that have been provided only the attributes.

5 Implementation

Here the image database consists of 500 images. In which 150 image are utilized for testing and remaining 350 image are utilized for training. First of all images are stored in database. Then the user has to select input image from testing database i.e. the query image. Query image are display on screen once the query image is displaced. Further step will be to get the similar type of image from training database. The features from which images are retrieved depends on color, shape and, texture. Color feature are extracted by utilizing color correlogram, color moment and hsv histogram. SVM method classify query image to relevant image in image database which results in higher accuracy. As peer to peer system is being used .We are using the two systems and connecting this system by giving the IP address of both .And the system is made online we can give input query image on system A and load the dataset on same system. Once the system is being loaded we can get the same type of images on the system B. This is how the images are retrieved in peer to peer. The method is found to be effective and robust, in terms of accuracy and variety of images considered.

The accuracy of method is found to be 92.4% which is effective and is comparable with other methods.

The figure2 describes that for example beach image is given as an input query image .the output image obtained is that the similar type of the beach input images.

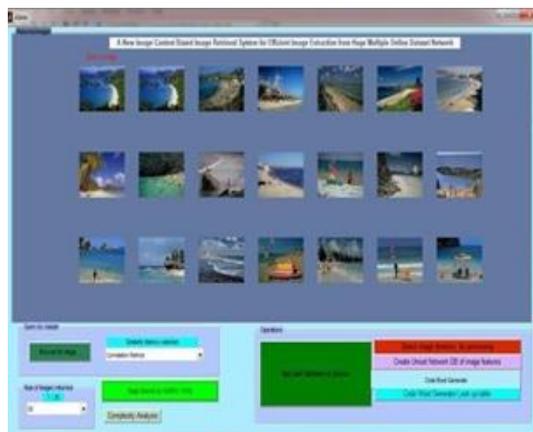


Figure2: Retrieving similar type of images from the given input query image

6 Result Analysis

The figure3 describes the comparative accuracy analysis for 100 images here we are taking five dataset that is the Tribe, Beach, Bus, Buildings, Dinosaurous images. The accuracy analysis for Tribe images is 92%,then the Beach images is 90%,then for the Bus images is 89%,then for the Buildings is the 90%,then last for the Dinosaurous is the 93%.This is how the accuracy analysis for all the five dataset used.

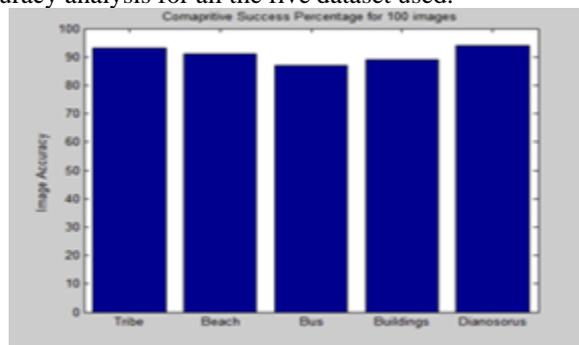


Figure3: Comparative accuracy analysis for 100images

7 Conclusion

This paper examined the primary parts of the content based image retrieval system traits depiction. Many further studies in this area have been recognized and also usage of this technique is shown in the same. It is concluded that to attain the extraction at the high speed as well as for making it so flexible that it can also adjust with the images of large size, it is required to have an impressive multidimensional approach. The main goal of this manuscript will be to signify the importance of SVM in effective retrieval of image. In this manuscript, SVM will be utilized as classifier that is execution the task of categorizing image and this procedure of categorization will be provided to all the traits of image that have been extracted after feature extraction procedure. It will be typically utilized for calculating the highest margin hyper planes within feature space that is also a high dimensional feature space.

References

1. W. M. Smeulders, M. Worring, S. Santini, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380.
2. Y. D. Chun, and I. H. Jang, "Content-based image retrieval using multiresolution color and texture features," *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 1073–1084.
3. S. Tong and E. Chang, "Support vector machine active learning for image retrieval," in *Proc. ACM Multimedia*.
4. GulfishanFirdose Ahmed, "A Study on Different Image Retrieval Techniques in Image Processing", in September 2011.
5. X. Zhou and T. Huang, "Relevance feedback for image retrieval: A comprehensive review," *Multimedia Syst.*, vol. 8, no. 6, pp. 536–544.
6. D. Tao, X. Tang, and X. Li, "Which components are important for interactive image searching?," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 3–11, Jan. 2008.
7. Gupta, A., Jain, R.: *Visual Information Retrieval*. Communications of the ACM, vol. 40, 70–79. (1997)
8. Faloutsos, C., Flickner, M., Niblack, W., Petkovic, D., Equitz, W., Barber, R.: *Efficient and Effective Querying by Image Content*. (1993)
9. Rui, Y., Huang, T.S., Mehrotra, S.: *Relevance feedback techniques in interactive content-based image retrieval*. In *Storage and Retrieval for Image and Video Databases*.
10. Smith, J. R., Chang, S.: *Automated binary texture feature sets for image retrieval*. *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Proc*, May. (1996).
11. Ono, A., Amano, M., Hakaridani, M.: *A flexible content-based image retrieval system with combined scene description keyword*. *Proc. IEEE Conf. on Multimedia Computing and Systems*. (1996) Ray-I Chang et al. *ComSIS Vol. 9, No. 4, Special Issue*, December 2012.
12. K. A. Hua, K. Vu, and J. H. obsamatch, "a flexible and efficient sampling based image retrieval technique for large image databases."
13. Rajshree S. Dubey, RajnishChoubey, JorBhattacharjee, "Multi feature Content Based Image Retrieval", *International Journal on Computer Science and Engineering (IJCSSE)*, volume: 2, No. 06, 2010, pages: 2145- 2149.
14. Jianhua Wu, Zhaorong Wei, Youli Chang, "Color and Texture Feature For Content based image retrieval", *International Journal of Digital Content Technology and its Applications*, 2010.
15. D. Joshi, R. Datta, and J.Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age", *ACM Comput. Surv.*, April-2008.
16. Tianzhong Zhao, Lu, Yafei Zhang1, "Feature Selection based on Genetic Algorithm for CBIR", *Congress on Image and Signal processing, IEEE*, volume: 02, May-2008, pages: 495- 499.
17. Jonathan E. Rowe, "Genetic Algorithm theory", *ACM GECCO '07: Proceedings of the 9th annual conference companion on Genetic and evolutionary computation*, July- 2007.