International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 06 | June-2018 www.irjet.net p-ISSN: 2395-0072

CBIR Using Support Vector Machine in OpenCV

Piyush A. Dahake[1], S. S. Thakare[2]

[1]*Electronics and Telecommunication Department, GCOE, Amravati, India*
[2]*Assistant Professor, Electronics and Telecommunication Department, GCOE, Amravati, India*

Abstract: *Digital image processing is an active research area for image retrieval since the last two decades. This is having an importance as increasing the multimedia content over the internet and data storage. In Content Based Image Retrieval there are some low level descriptor of image like Color, Texture and Shape. These image descriptor are used for image representation and analysis. This paper presents a content based image retrieval (CBIR) using support vector machine system. In CBIR the image features extracted by color moment like HSV, RGB and Grey Image, for shape used HOG (Histogram of Oriented Gradients) and finally for Texture feature used GLCM (Grey Level Co-occurrence Matrix) and HOG this all descriptor leads to more accurate results for image retrieval. In this system SVM classifier is used to classify the data with input query image and also to classify the images into different classes to measure the similarity only with the images belongs to same class. This system has more accurate retrieval scenario and reduced retrieval time.*

Keywords:

Content Based Image Retrieval, Image Retrieval, SVM Classifier, HOG.

.1. INTRODUCTION

The system in image processing which retrieves the desired images from a bunch of collection of images automatically having the visual features is called a Content Based Image Retrieval (CBIR) system. Content based image retrieval becomes an important area in digital image processing for research with increasing demand and also use in digital images in various fields [1]. A typical performance of CBIR technique has two major functions. First one is the feature extraction, all features of images are extracted from each image in database and it describes the content of the image. The second function is to reckon the similarity measurement of features between features of the query image and the images in the database. In CBIR the user input are in the form of query image and the features are extracted from the input image and it is stored as a features vector. These all features of query image are compared with the previously extracted feature and stored feature vectors of the database images. The comparison is done by comparing the distance values and these vector values are used to rank the images at output according to their similarity with the query (input) images. The most similarly matching images are finally displayed to the user. In the system of CBIR the main challenge is to represent each image in particular way to gives accurate identification of the image. Therefore, the perfectly successful retrieval system depends on the choosing the right image descriptors which will represent that image accurately and size of feature vector. Here in this system the initial classifier is used to identify the class label of the query image. Then measure the distance between the query image and the images in the database which falls under class, computed to retrieved most similar. Then distance of query image and the images in the database i.e. class is computed for retrieve the most similar images.

2. METHODOLOGY

Proposed system is designed to reduce the complexity of feature extraction and make image retrieval much more efficient (Figure1). Before extracting features, images must be filtered out to reduce noise in the image. Then feature extraction is done for the resulting image. As the output of feature extraction it produces a data file. This file is a complex when compared in terms of similarity. So the proposed support vector machine algorithm is used to train input feature file for efficiency, correctness and reducing the time duration of image retrieval. Results produce by this module are sent to the Clustering and Store Module. In most of the existing systems, this is the most time consuming and the most significant phase. So it is important to compress and simplify feature extraction process for effective component based image retrieval. Proposed system is implemented to improve the efficiency of feature extraction and time. This system is different from other existing approaches in two ways. First is improving the efficiency of the system by introducing a new hybrid methodology in feature extraction module and a new methodology of SVM to train data and produce module file.

3. SYSTEM DEVELOPMENT

The basic processing steps of proposed CBIR techniques depicted in figure1. It is divided into three parts:

1) Query unit 2) Database unit and 3) Retrieval unit.

In query unit any one image from database is selected as an input image and the feature of that image get extracted and stored as a feature vector. The block diagram of proposed system is given below.

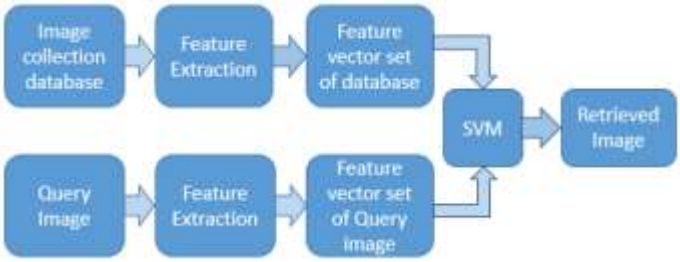

Fig -1: Block Diagram of Proposed system

3.1 Flow of design:

The CBIR techniques has some following steps that are given below,

Create a Database: Create or store some image database to prepare for own database for testing purpose or as a inbuilt database.

Query Image: The query image is nothing but an input image which is we giving to the system as input and according to that images system will find out similar images.

Features Extraction: Extract the important features from query image and database image like color, texture and shape etc.

Feature Matching: Measure the content of query image and database images and based on that it checked and the feature which are closed to the input image it retrieved the corresponding images from database.

Retrieve Image: Based on the image features content it will retrieve the images.

3.2 Feature Extraction:-

The feature extraction has some processes to extract the feature image. Below is the description of feature extraction.

Here is the 1) Image Acquisition

2) Superpixel Segmentation

3) Feature Extraction

Fig -2: Describe form of Feature Extraction

The working of image acquisition is to apply the filter to query and database image. The image before going to next operation it will remove the noise from image.

A superpixel can be defined as a group of pixels which have similar characteristics. It is generally color based segmentation. Superpixels can be very helpful for image segmentation. There are many algorithms available to segment superpixels. Superpixels are becoming increasingly popular for use in computer vision applications. However, there are few algorithms that gives output with a desired number of regular, compact superpixels with a low computational overhead. We introduce a novel algorithm that clusters pixels in the combined ve-dimensional color and image plane space to efficiently generate compact, nearly uniform superpixels. The simplicity of our approach makes it extremely easy to use a lone parameter species the number of superpixels and the efficiency of the algorithm makes it very practical. Experiments show that our approach produces superpixels at a lower computational cost while achieving a segmentation quality equal to or greater than four state-of-the-art methods, as measured by boundary recall and under-segmentation error. We also demonstrate the benefits of our superpixel approach in contrast to existing methods for two tasks in which superpixels have already been shown to increase performance over pixel-based methods. In this system the SLIC algorithm is used for superpixel segmentation.

3.2.1 Color Feature:

Color is most important content in the color image and it is most widely visual content. In this type of feature identifiers the proportion of pixels of specific color in the image. It is a three dimensional and as vector in it. In this system used RGB, HSV, Lab and grey image color format various color descriptors based on the color coherence vector, color histogram, color moments and color correlogram.

3.2.2 Image retrieval Based on Texture features:

The texture is an important feature in the image which is used in pattern recognition. The similarity in texture can be useful in make difference between areas of images with similar color. The texture is like sky, leaves and sea etc. In this system we used Histogram of Oriented Gradients (HOG) and Grey Level Co-occurrence Matrix (GLCM) algorithms are used.

3.2.3 Image retrieval Based on shape features:

The shape is nothing but edge in image, edge indicates sudden change in the pixel density of the image. There are various types of edge detection techniques Canny, Sobel, Prewitt and Robert edge detection techniques. But in this system we used Histogram of Oriented Gradients (HOG) algorithm is used to describe the shape feature.

3.3 Support Vector Machine (SVM):-

Classification performed by SVM is by generating a hyper plane. The optimal hyper plane separates the required data in two categories. SVM are close cousin multilayer perception neural network. Generally support vector

machine can be defined as the training method if polynomial radial basis function where the weights are calculated by solving QP problem [2]. According to the support vector machine predictor value is called an attribute and transformed attributes which is used to define the hyper plane is known as feature.

The method of selecting the appropriate representation is called feature selection. Now the vector can be defined by set of features that describes predicted values. T

he aim of support vector machine is to find the optimal hyper plane that is used to distinguish group of vectors in a way that one category of the required variables is on one side of hyper plane and the other class of variables are on the other side of plane[5]. Vectors nearer to the hyper plane are called as support vectors. The optimum hyper plane can be defined as the linear classifier with the maximum margin for a given set of variables.

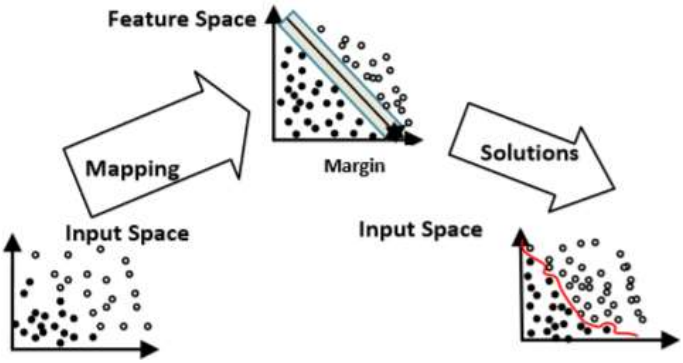

Fig -3: Flow of SVM Algorithm

4. RESULT

This paper provides an overall performance investigation of the experimental model based on different algorithms. Here, Content based image retrieval using support vector machine is implemented using software tools. Results for software implemented system are observed and compared.

Resource utilization for algorithm is observed with the help of Design Summary Overview which allows a quick access to design overview information, reports, and messages.

At the end in this CBIR using Support vector machine results techniques are observed and based on various parameters their comparative analysis is done.

Also, comparison of proposed design with earlier designs have been done successfully.

Fig -4: Output of System for face images

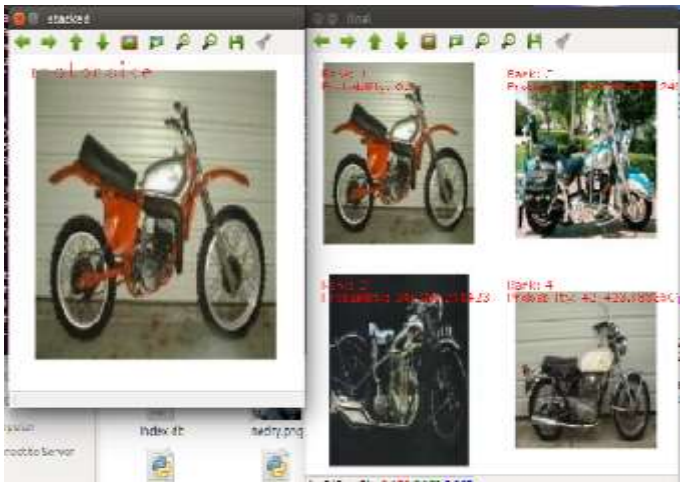

Fig -5: Output of System for Bike images

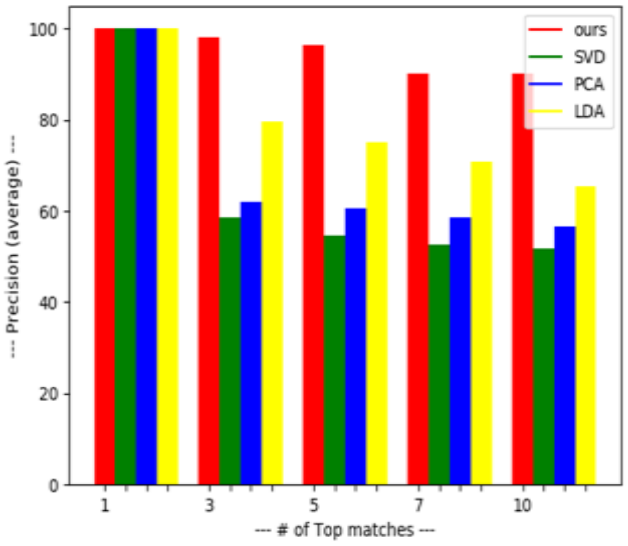

Fig -6: Graph of Comparative Analysis

The comparative analysis of SVM with some of the previous algorithms is shown below in form of table as well as tabular graph format. In CBIR technique if the number of retrieval images are increasing so it will decrease the pricision of retrieve images and will increase the retrieval time of image.

	No. of top recognized matches				
Methods	1	2	4	8	10
Single Value Decomposition	100	58.5	54.5	52.5	51.9
Principal Component Analysis	100	62	60.5	58.5	56.5
Linear Discriminant Analysis	100	79.5	75	70.9	65.5
Support Vector Machine	100	98	96.5	88	79.5

Table -1: Comparative Analysis for number of Images retrieval

In this proposed model the Support Vector Machine (SVM) is compared with different existing method such as SVD, principal component analysis PCA, and LDA. The recognition rate for top ten images to single value decomposition technique PCA, LDA and SVM are 51.9%, 56.5%, 65.5%, 79.5% and the computational time for all existing methods are 40.95 sec, 42.56sec, 58.6sec and for SVM it took 6.03356sec.

5. CONCLUSION

The proposed system of image retrieval using SVM with python in OpenCV is implemented. Along with the comparative analysis of various image retrieval algorithms have been done. From the comparative analysis it is observed that the retrieval time for a particular image have been reduced and the precision is increased. In this system color, texture and edge features for a typical input image and predefined data based have been extracted. The primary aim of this paper is to represent the significance of Support Vector Machine in efficient retrieval of image. Furthermore, it possible to improve this retrieval performance by using other machine learning algorithms along with reduction in the retrieval time.

REFERENCES

[1] K.B.A.B. Chathurika, P.K.S.C. Jayasinghe, "A Revised Averaging Algorithm For An Effective Feature Extraction In Component-Based Image Retrieval System", International Advance Computing Conference (IACC). IEEE, 2015

[2] AnushaYalavarthi, K. veeraswamy, K. AnithaSheela, "Content Based Image Retrieval Using Enhanced Babor Wavelet Transform", International Conference On Computer, Communications And Elelctronics. IEEE, 2017

[3] Ms. Sandhya R. Shinde, Ms. SonaliSabale, Mr. Siddhant Kulkarni, Ms. Deepti Bhatia, "Experiments On Content Based Image Classification Using Color Feature Extraction", International Conference On Communication, Information And Computing Technology (ICCICT), Jan. 16-17, IEEE, 2015.

[4] Datta, Ritendra, et al. "image Retrieval: Ideas, influences, and trends of the new age" ACM computing Surveys (CSUR)40.2 (2008)

[5] Cortes, corinna, and Vladmir Vapnik, "Support Vector networks." Machine learning20.3 (1995)

[6] Marouane Ben Haj Ayech, Hamid Amiri, "Content Based Image Retrieval In The Topic Space Using SOM And LDA",3rd, International Conference On Content Engineering And Information Technology (CEIT), 2015

[7] kanwalpreetkaur, "On Comparative Performance Analysis Of Color, Edge And Texture Based Histograms For Content Based Color Image Retrieval",International Conference On Reliability Infocom Technologies And Optimization (ICRITO) (Trends and Future Direction), IEEE. 2014

[8] Shujin Zhu And H.R. Tizhoosh, "Radon Features And Barcodes For Medical Image Retrieval Via SVM", International Joint Conference on Neural Networks (IJCNN), IEEE. 2016

[9] Ammar Huneiti, Maisa Daoud, "Content-Based Image Retrieval Using SOM and DWT", Journal of software Engineering and Applications, 2015

BIOGRAPHIES

Piyush A. Dahake was born in 1992. He received diploma in (Electronics and Telecommunication) from Government Polytechnic Amravati in 2013, also holed B.E. in (Electronics and Telecommunication) from PRMCEAM, Badnera, Amravati in 2016. His research interest into DIP, Digital Electronics, Embedded System. He is currently pursuing M. Tech in (Electronics System And Communication) at Government College of Engineering Amravati (MH).