Feature Fusion for Image Retrieval using Image Processing.

Amol Potgantwar, Shreyas Deshmukh

Abstract— Image processing and computer vision uses Content-based image retrieval (CBIR) function to solve the issue of image retrieval, which means, solving the issue of image searching in expansive databases. The actual data of the image will be evaluated when a search is performed that refers to content-based. The term content can be any attribute of an image like colour-shade, various symbols or shapes, sizes, or any other data. There are various approaches for image retrieval but the most prominent are by comparing the main image with the subsets of the relatable images whether it matches or not and the other one is by using a matching descriptor for the image. One of the main trouble for huge amount of CBIR is the representation of an image. When a given image is worked upon it is divided into number of attributes in which some are the primary ones and others are the secondary ones. These attributes are checked with the local and MPEG-7 descriptors. All this is then mapped in a single vector which is the same images but in compact form to save the space. Principle Component Analysis (PCA) is used lessen the attribute size. To store the attribute data in similar clusters and to train them to give the correct output the study also uses k-means clustering algorithm. Hence, the proposed system deals with the image retrieval using various algorithms and methods.

Index Terms—CBIR, image retrieval, MPEG-7 descriptors, PCA, k-means clustering.

I. INTRODUCTION

Now a days we have a lot of data in various forms, like documents, videos, images, etc. There are many softwares, apps, methods used to work with a particular kind of data. If the data is in text format then we have softwares like adobe reader, notepad, and so on. Same is with the images because a huge amount of data is transferred to one another using images in which some are open to all but some are private. But, we are continuously dealing with images in today's world.

There are many areas which are benefited by CBIR technology like video management. But it is still a topic under research. The technology is in boom but many things need to be tested and analyzed for efficient working of the system. Proposed System provide compact image representation and relevant result to the user query. As Using MPEG-7 descriptors number of features are extracted from an image which will provide high accuracy in image retrieval. Principle component analysis method provide compression of image features without loss of data.

The usage of images in today's world is in huge quantity over the internet which directly increases the number of users. Like, there are so many different images for a single word. Images play a very important role in the society. Today images are used in every field like, medicine, biology, aeronautics, crime, designing, advertising, archeology.

People have been in the use of photographs, TV and a lot more other things. This contributes to the growth of image usage. Images were always stored manually is past years until there was any technology for storing data. The images were stored according to their similarities giving them the index number in shelves, folders, cupboards. But, as the number of images increased the manual work was hectic and difficult. So, basically there was a need to automate this work for proper storage and efficient image retrieval.

There is a necessity to develop sucha system that will effectively manage the image collection. It is obviously true that if there is a lot of data here are chances of them getting disturbed or the quality depreciation. These kind of thoughts give the rise to build new systems that would make life much easier and simpler.

The quantity of collections of digital images have grown due to increase in the number of online users during this period, for example, in web applications that allow to add digital albums and images.

To convey any kind of information people find it easy to communicate through images rather than text or any other medium of communication. It is also true that visuals are very easy to remember than texts hence ad agency use a set of images into a movie where the customer remember the ad for lifetime. There is a lot to play with the images and they play an important role.

The goal of this system is to study the content-based image retrieval using different MPEG-7 descriptors. Also feature extraction of images, clustering of images, reranking of images. Goal of this system is to show that the advantages of using MPEG-7 Descriptors along with local descriptors and principle component analysis provide a better solution for many number of problems. These kind of scenarios are common in a many domains of data. Hence, the system has tried to find out a new way through which the classification problem has been solved by clustering way, which is providing high quality results due to incorporation of side information.

There is a lot of scope for this technology as it is used in many fields which will always help mankind to lessen their troubles and make the tasks easy. The future world is all about digitalization and the work would be done with one click.

II. LITERATURE SURVEY

In the proposed system [2], RootSIFT method is used instead of SIFT to improve retrieval performance. It is a new method where we get high performance without the need to change the other requirements. They have also introduced a method for expansion of the query where a higher model is used for retrieval of image using the concept of inverted index. The main objective of the system is to yield image retrieval immediately and...
dynamically from large databases, where the image has a query subset to compare and give the output correctly.

In [3], the authors have majorly worked upon VLAD. They have focused on the same problem that of huge requirements for storing the data. Here, three contributions are done: firstly they have used normalization with a little change to it which will ultimately improve the performance, second being the adaptation of vocabulary where if one word is typed incorrectly images are shown only for the correct word but it should also show the other images relating to that word, third is the multi-VLAD which allows the system to localize the data or the image that is only a small part of the whole image which gives a high and fine resolution of the image.

In this study [4] a new data structure is introduced which is the and evaluated a new data structure called the Inverted Multi-Index. They have used databases of SIFT and GIST vectors. A comparative of inverted index and inverted multi-index is also done to check the efficiency. A well known method is used for query image that is the nearest neighbour search.

Here in [5], the concept of bag-of-words is used which means if two or more attributes come under a same zone then it will be counting 1 or else in different zones then count 0. Matching of kernels is the main key here. There will be a set of local attributes and a dictionary where the local attributes will be checked with the dictionary we have. Experiments are done using Caltech. It gives prominent and high efficiency.

L. Chu, in [7] have discussed some problems like huge number of attributes in the non-duplicate zones, small number of prototypical attributes, duplicate zones. Here they have proposed PDIR (Partial Duplicate Image Retrieval) which correctly retrieves the noisy or duplicate images by correctly matching the SIFT attributes. The study is based on COP (Combined Orientation Position) model. It has two parts one being the consistency to accept the proper data and consistency to reject the noisy data. Experiments are performed to check the performance of the developed system.

In the proposed study O. Chum and other authors have proposed an algorithm which is formed from the min-Hash. This model tells us about the dependencies of data where the attributes that occur in high probability [8].

This study presents the objective of retrieving all the occurrences of a given image from a large database. They have also used the concept of bag-of-words but have also mentioned some issues with it. Re-issuing of the documents add into the new query and new terms. The main idea is to retrieve the images that match to the input query image. Then to combine the retrieved image and the original query. This gives a rich and qualified model of images. Continue this again and again to from a high leveled model [9].

In [10], gain the concept of VLAD is used. The system uses indexes that are given to the images using the local descriptors which are closely related to VLAD and FV (Fisher Vector). It is suggested that it is easy to encode an image giving excellent results by combining the mentioned descriptors with some coding technique.

[12] demonstrates video annotation using image query. The images are saved in the database then from the video similar kind of images are segregated. These images are then compared with the images in the database and the output is given. They have proposed a framework of hierarchical video-to-near-scene annotation.

III. PROBLEM STATEMENT

The aim of this system is image retrieval using image fusion. The type of feature that the system is extracting Edge Histogram Descriptor (EHD) and Scalable Color Descriptor (SCD). When a given image is worked upon it is divided into number of attributes in which some are the primary ones and others are the secondary ones. These attributes are checked with the local and MPEG-7 descriptors. All this is then mapped in a single vector which is the same images but in compact form to save the space. Principle Component Analysis (PCA) is used lessen the attribute size. To store the attribute data in similar clusters and to train them to give the correct output the study also uses k-means clustering algorithm. Hence, the proposed system deals with the image retrieval using various algorithms and methods.

IV. PROPOSED SYSTEM

A. System Architecture

![Fig. 1 System Architecture](image)

The proposed system is divided into following three blocks:

1. **Attribute Extraction:**
The term content can be any attribute of an image like colour-shade, various symbols or shapes, sizes, or any other data. In a CBIR system these attributes are represented in vector format. This group of attribute vectors of the images which is stored in the database is known as attribute database. This process is started as soon as the user gives a query as input like a drawing or image of any object. Then the query is compared with the images present in the trained database. The whole image at once is not compared but small paerts of it are compared. These attributes are checked with the local and MPEG-7 descriptors. For example if the query image given by the user is the face of horse then all the attributes will be compared and then in the further steps all the images that includes horse will be shown.

2. **Attribute Space Reduction:**
When attribute extraction is done on image using descriptors, there is an increment in the space where the image is stored. This
happens because every feature is extracted and stored separately which increases the storage space. To overcome this issue we use a method known as Principle Component Analysis (PCA). PCA has its applications in different fields like compression of images, pattern recognition, etc and is a useful technique. This method is used to find out the similarities and differences present in the data. If there is an image that is being checked as it is the input then every attribute from it is extracted and it happens that these attributes cannot be of high dimension hence, PCA represents it high dimension with less storage space and without the loss of data.

3. Cluster Formation:
Cluster formation is done by using a very popular method which is k-means clustering algorithm. This method forms various clusters based upon the similarities among the data. These clusters and the data in it is trained.

This algorithm is basically used to form groups of similar kind of data in k partitions or groups. For instance we are checking a dataset of different colours which has colours like red, blue, green, yellow, black, etc. First colour is taken and placed in the cluster number one which is red. Then the next colour is picked up if it is red then it will be placed in cluster one if it is other colour like blue then a new cluster will be formed. This algorithm goes on until last input. We then finally get all the clusters differentiated in a very good manner which will not cause any confusion. The same goes with the proposed system here all the images matching to the input image will be shown correctly.

The details of steps performed in proposed system are as follows:
Step 1: Get the the image that has to be matched with the images in the dataset.
Step 2: Extract attributes of the images
Step 3: Combining all the attributes
Step 4: Apply PCA on these attributes
Step 5: Matrix calculation from step 4 output
Step 6: Apply K-means clustering algorithm
Step 7: Query Image
Step 8: Extract feature of query images
Step 9: Classify cluster of image
Step 10: Re-rank images,
Step 11: Retrieve relevant images

Algorithm 1 : Feature Extraction
Input : Query Image
Process:

Algorithm 2 : Principle Component Analysis
Input: Image Features
Output: Image Features
Process: This algorithm contain following Stages:
1. Extracting the information from the image and acquire data
2. To construct the covariance matrix from the acquired data
3. Eigenvalues and corresponding eigenvectors are calculated.
4. Short Eigen vectors and find principle Eigen vectors vc = (av1, av2, av3, . . . ,avn) where, vc: Eigen vector
5. Finally the information is obtained which is in the form of matrix including eigenvectors (components) of the covariance matrix and the original image is obtained.
Output: Dimension Reduction of image features.

Algorithm 3 : Image Classification Using k-means

The k-means algorithm is used for clustering, by applying k-means algorithm the clusters of an images are formed and images are trained. The generalized k-means algorithm is explained in the following paragraph.

K-means clustering is a method commonly used to automatically partition a data set into k groups. First the value of k is decided and then the centers are randomly placed to which similar data will be put in the clusters. Then the similar kind of data is placed near the matching centers which forms a group of similar kind of data. Rearrange the sets again to find out whether the sets are the same or there is any change. If it is the same then exit out of the loop and again follow the same process for the new data.

Algorithm 4: Similarity Measure

The distance between any two points in the space is known as euclidean distance. If two points are taken say G and H in 2D space where the coordinates of G are (g1, g2) and the coordinates of H are (h1, h2). Hypotenuse will be formed by joining the two coordinates G and H. This will follow the pythagorous theorem where one will be the hypothenuse which will be the square root of the sum of the other two distances. Euclidean distance method is used to find similarity measure and provide efficient searching. This method is used to compare profiles of variables, To compute similarities and dissimilarities among images.
V. MATHEMATICAL MODEL

Let $S$ be the solution

So,

$$S = \{ \text{IN}, \text{FT}, \text{OT} \}$$

where,

$S$ = Set of inputs, functions, outputs

Now,

IN is the set if inputs

i.e., $\text{IN} = \{ \text{IN}_1, \text{IN}_2 \}$

Let,

$\text{IN}_1$ = set of image dataset

$\text{IN}_2$ = query image

Now,

FT is the set of functions

i.e., $\text{FT} = \{ \text{FT}_1, \text{FT}_2, \text{FT}_3, \text{FT}_4, \text{FT}_5, \text{FT}_6, \text{FT}_7 \}$

Let,

$\text{FT}_1$ = Edge Histogram Descriptor

$\text{FT}_2$ = Scalable Color Descriptor

$\text{FT}_3$ = Color Layout Descriptor

$\text{FT}_4$ = aggregate feature

$\text{FT}_5$ = train image dataset

$\text{FT}_6$ = query image feature extraction

$\text{FT}_7$ = classify image

Now,

OT is the set of output

i.e., $\text{OT} = \{ \text{OT}_1 \}$

Let,

$\text{OT}_1$ = get relevant images

VI. RESULT ANALYSIS

The analysis done here is for huge number of dataset. There are a lot of images from which the system gives the correct output and accuracy level is high. There are some figures that demonstrate the how the system performs.
There are many other methods by which the accuracy is calculated like PCA + Kmeans by cosine similarity, KNN by absolute distance, and many more.

VII. CONCLUSION

The goal of this system is to study the content-based image retrieval using different MPEG-7 descriptors. Also feature extraction of images, clustering of images, reranking of images. Goal of this system is to show that the advantages of using MPEG-7 Descriptors along with local descriptors and principle component analysis provide a better solution for many number of problems. These kind of scenarios are common in a many domains of data. Hence, the system has tried to find out a new way through which the classification problem has been solved by clustering way, which is providing high quality results due to incorporation of side information. There is a lot of scope for this technology as it is used in many fields which will always help mankind to lessen their troubles and make the tasks easy. The future world is all about digitalization and the work would be done with one click.

REFERENCES

12. Chien-Li Chou, Member, IEEE, Hua-Tsong Chen, Member, IEEE, and Suh-Yin Lee, Senior Member, IEEE, Multimedia Video-to-Near-Scene Annotation, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 19, NO. 2, FEBRUARY 2017
Feature Fusion for Image Retrieval using Image Processing.


AUTHORS PROFILE

Dr. Amol Potgantwar has completed Ph.D and is H.O.D of Computer Department, Sandip Foundation, Nashik. The focus of his research in the last decade has been to explore problems at near field communication and its various applications, in particular interested in applications of mobile computing, wireless technology, near field communication, image processing and parallel computing. He has registered 5 patents on indoor localization system for mobile device using RFID and wireless technology, RFID based vehicle identification system and access control into parking, Standalone RFID nad NFC based healthcare system. Since last 5 years he is working with the organization Sandip Institute of Technology and Research Center. His focus has always been on nurturing and guiding the students in different domains and projects.

Shreyas Deshmukh has completed B.E in Computer Engineering from Sandip Foundation, Nashik. Has been a part of WSN based Infrastructural Health Monitoring and Audit System project in B.E. Had patent application on the same project which is also recognized by central government of India. He has been working with multinational company named FIS Global as Quality assurance engineer. He has good skill set in manual and automation testing domain. Along with, he has served many projects within deadline. He has got many recognitions from his team for working hard and efficient on the different projects.