Command Based Image Retrieval

Sunil Bhutada, Professor, IT Department Sreenidhi Institute of Science and Technology Yamnampet, Hyderabad. (sunilb@sreenidhi.edu.in)

Veekshith Kajeepeta, B.Tech IV year, IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad (veekshith.k@gmail.com)

Venkat sai kamalammagari, B.Tech IV year, IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad (talktovenkatsai@gmail.com)

Sreelekha Komma, B.Tech IV year, IT Department Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad (sreelekhakomma24@gmail.com)

Abstract:

As the years are passing we are going on searching for the images that we want which consume a lot of time and effort, at last, we get rid of searching for similar images. But here comes a solution to our problem by using a single command which helps us to save our time and effort, so here commands are helped to get wanted similar images by giving a single image in our cmd and it helps to show the similar or related image to our main image. The main image is also known as the query image. So here comes the command based line argument where single-line command helps us a lot in this process here we used python and computer vision together based on the single command. The image search engine we are going to build based on a command-line argument, where this search engine helps to analyse the features of the main image which is given as a query and find out the images that we want.

1. Introduction:

Due to the fast pace of advancement in the technology, usage of modern digital devices has been increased at a higher rate, which constitutes an increase in the stored and shared databases. The increase has become a challenge for the systems to find relevant pictures from given lot of big data. The general use of image extraction is mainly based on the query is to search for the images according to the features with the main image given to the computer vision. Image searchers on the internetwork on a text-based search that is matched with the main words from the images on databases.

Hence the output will get generated according to the match of keywords, and in this process, we will also come across some images that are not relevant. It's the point that it is unable to proceed with the method of labelling for a larger database which gets the non-relevant images along with the relevant ones. The next type would be an automatic annotation method that can be labelled image following the image details using a vector space model. It helps a system to detect colour, texture, edges, shape, etc. The perception of annotation of the automatic models can also misguide this process. CBIR is a model that will help us to save time and effort, as it is based on feature extraction of the image that is part of the main image.

When a user gives the actual image to the CBIR system, its most important function is to compare all the features of the actual image to the images in the stored data, and according to the visual similarity,

will find the outputs. In CBIR, having considered the features, the sorted list of images will be computed according to the match and given as output. According to the subject history. The real-world applications of this CBIR are such in medical, textile, traffic, crime investigation fields etc. The overall aspect of the image retrieval system is to get similar images. The selection of the features is dependent on the image given to the system.

The process of finding the features mainly depends on the image taken by the user according to their requirement. For any image retrieval application, the discriminative feature representation is the second most important aspect after considering the required feature selection. The features are considered based on their three important parts of the image as we go on looking for every feature of the image. As the features are of both high and low level. Where the high-level features help us to find out every small part in the picture in the low level it will extract the main feature.

In this review, we mentioned the recent trends in detail to be enlighted the important contribution, and further actions to be proceeded by keeping the actual detail on the CBIR and the classification through feature extraction. We have also started a structure to addressed and simplify the mentioned details.

2. Literature survey:

There will be several methods present to retrieve similar images such as using Matlab, based on vector spacing, on image processing in this way there are many methods present but my project is based on the command-line argument.

Firstly we will go with the retrieving of similar images based on Matlab and SQL it is a big process is based on system GPU so it takes so long time. Here when we open system a window will open for GUI so based on that we need to search for our query image[3] it is long and effort process that other techniques where some system doesn't even support MatLab. In this, we can calculate precision and recall. Here we have a bag of features which can be collected using shape, texture and colour here we even train the classified image and store the image detail in the file for further use[1]. The CBIR system here always works using GUI so it is a long process than other techniques as we feel but in the Matlab, we can extract features in many ways it is a very good process using the Matlab.

Now we will see with the KNN algorithm where here we train the sample data or image first later we will apply the weighted mean value to get the accurate result from our data this can be worked accurately when the class labels are present then later we should train the data and test the data to get results but this whole process is depended mainly on the k validated data points[2]. To get this validated data points we need to train the nearest neighbours of the trained dataset. In this, after getting the weighted value we used to find the distance using Euclidian distance now these weights are added to choose the biggest weighted class. It is a modifies KNN so it can work on both labelled and unlabelled data[4].

The spatial features of the image can be described using vector space modelling were for every image there is an important point of extracting the spatial distance in the image it will help in all the three details will be extracted.

The support vector image modelling is an ML algorithm[5] which helps to find the spatial distance accuracy to retrieve the similar images accurately, this support vector machine always tries to derive the training data which will help us to derive all the features of different data that to be trained in a predefined way as the image has some of the hidden [2] data using spatial distance frim each pixel we can retrieve that hidden data and store in the data. Where that hidden data helps a lot to retrieve similar images.

3. Proposed Methodology:

In the current generation of digital communication, the usage of digital images for various things like expressing, sharing and interpreting information has drastically increased. As the usage of digital images has increased, how it can be retrieved are simultaneously improving. Quite often, while working on these images, we might require to retrieve similar images for different purposes based on the visuality of the image. It might look easy when you are dealing with a few tens of images but the task gets tougher when the count of images is in thousands and millions.

In such cases, it becomes highly impossible for a human alone to complete the task. To overcome this situation, a more efficient and accurate system for image retrieval, known as Content-based Image Retrieval System is required to search and collect relevant images.

It searches pictures basing on the optical appearance of the query input image, such as colour, texture, shape, and return similar images. There are many other retrieval systems which collect images based on the text input, query input, annotation-based etc. But out of all these, Content-Based Image Retrieval is most advisable because searches that rely on the metadata depending on the quality of the text and its completeness. Having humans notations of the images by entering keywords would consume more time and may not be able to grab the keywords properly required to brief about the image leading to less efficiency and accuracy.

CBIR is being used by various companies in various sectors for designing in Agricultural and other Engineering fields, various Artistic collections, Prevention of crimes such as theft, murder etc, Geographical Information and remote sensing systems, Medical purposes, Face detection, Military purposes etc.

3.1. Architecture:
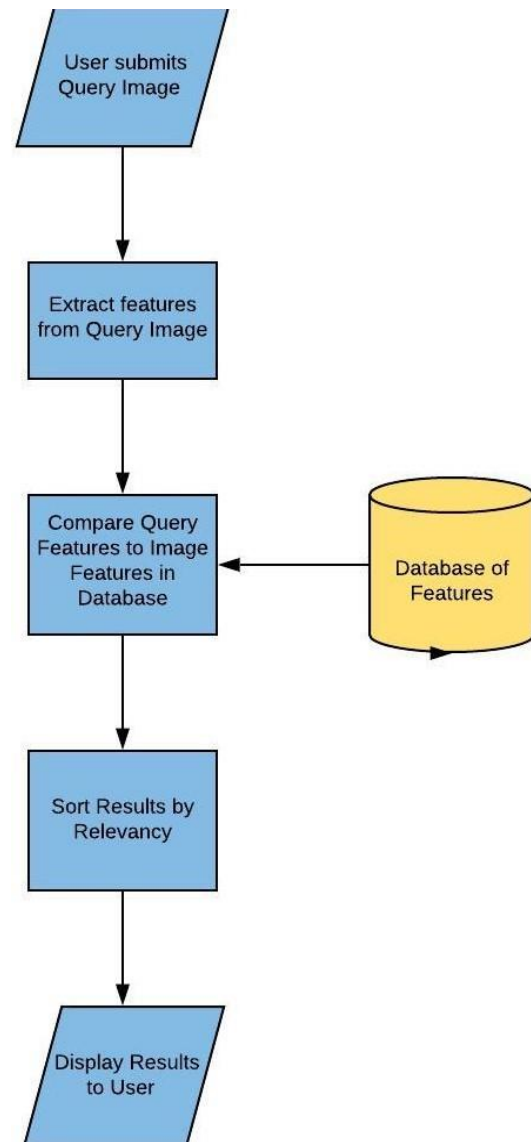
First, the user decides what images he/she wants. Based on the images, the user gives a Query Image as an input. Since we are performing the operations on command prompt, we give the input in Command Line Interface. The input consists of the python file which consists of the code for the algorithm followed by the name of the query image and respective options to derive specific output. This image is taken as the input and the computer extracts features from the Query Image based on the factors like colour, texture, shape etc. These Query image features are gathered in a database and then check with each image f in the database (which contains a large collection of images stored in a proper and

retrievable format). The main image every detail are compared with the features of every picture from the database. After comparison, based on the same features, the images with the same features are retrieved and sorted. Finally, the results are displayed on the screen as output to the user. There can be any number of similar images and there is no specific limit to it. The images retrieved using this method are more accurate and efficient when compared with other Image Retrieval Algorithms.

3.2. Pseudocode:

3.2.1. Input:

We will take image as an input. Here the image is considered from a query. Suppose if we require similar images of houses, we give the name of the input image as house in a query.
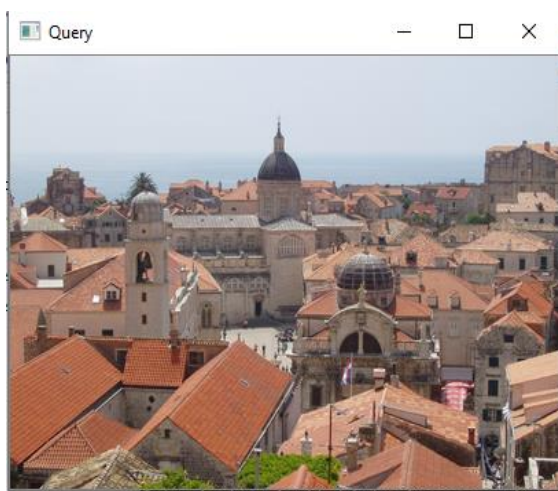
From this query image, features of houses like colour, shape and texture are extracted. After these features are extracted, they get stored into a file named index and with an extension .csv. We also give the database path, through which the features of images in that database will also be extracted and stored into the same index file.

3.2.2. Output:

The features of the input image will be compared with those of the of the images of that database stored in the index file. After comparision, according to the relevance the sorted order of images as output can be seen. After the query image is a view within few seconds result of all the houses are shown below. Based on the name in the command like houses, trees, the water we can retrieve all the images on the name as we stored our pictures according to the names so it helps us to retrieve the images easily.


4. Experimental setup:

4.1. Collection of Data:

First we need to collect the image data. The data can be collected from a different source of the internet as our data includes a set of images with similarities

because based on those features in the query image we can also retrieve other similar images easily and it is a time-saving process. So data collected should be stored in our database. The database should consist of different kind of data collected from different sources of internet. Now we should one image as the main image and store that query image in the other folder to get to know easily in the same way because that query image helps us to retrieve the other similar images. As the same way pick other required image from database and retrieve their similar images. As data collection plays the main role for us to get the output very fastly in a few seconds.
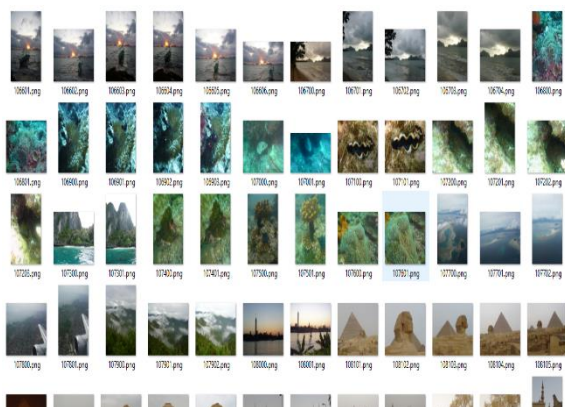
4.2. Features Extraction:

In features extraction, to extract features from an image, we divide the image into five parts according to their respective co-ordinates. Now computer vision will analyse each part of the divided five parts starting with colour analyses. In general, the image pixel will contain of three colours Red, blue, green (RGB). Now it converts RGB colour space into Hue Saturation Value (HSV) colour space.

We convert it to get the picture intensity value because RGB fails to identify pre-received colours. In order to identify shapes and texture in an image, OpenCV will recognize the visual features of it accordingly. Using all these features, a histogram vector will be formed.

4.3. Storing Features:

Indexing is nothing but storing the features of an image. From the query image, features of like colour, shape and texture are extracted. After these features are extracted, they get stored into a file named index and with an extension .csv. We also give the database path, through which the features of images in that database will also be extracted and stored into the same index file.

4.4. Searching:

The search engine is used to find out the images. Here we built a search engine to find out the images based on the colour, texture and shape. Where the images can be found by using the single query image.

Now comes to the colour, texture and shape. The image is firstly converted into the pixel form where the pixel consists of red, blue and green colours now we used to convert that RGB into hue saturation value colour space by using this we can form the histogram for texture and shape to be recognized by the computer vision.
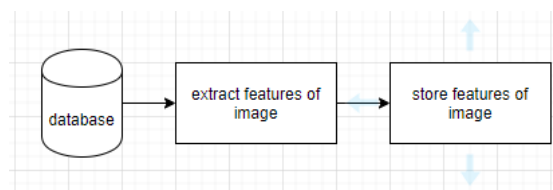
4.5. Command based input:

This undermentioned command is used to take a query image, through which the database is searched with the similarities in features with a path set for the database to search for.

```
python search.py --index index.csv --query queries/houses.png --result-path dataset
```

5. Conclusion:

So far, CBIR technology had a little use on the normal purposes of image extraction such as for home entertainment or journalism. On the other hand, it had shown a significant impact on areas such as crime prevention, Medical diagnosis etc. With rising demands of multimedia applications

over the world wide web, the importance of image retrieval has also increased. The dramatic rise in the size of the images has led to the creation of more efficient and effective retrieval systems. The development of these systems started with textual annotations and later developed into image retrieval based on visual content. This came to know as CBIR or Content-Based Image Retrieval. However, all these techniques have their benefits and limitations. Different Systems are used for different purposes to satisfy different user requirements. In this project, we have developed a system which searches relevant images based on the user given query input image. Here, we have taken a sample of images for demonstration purposes. Similarly, any number of images can be stored and retrieved from the database based on the relevancy of images. We have also emphasized the need for advanced image retrieval techniques. The primary goal of Image Retrieval is to give an effective and efficient purpose for organising, finding and gathering a large collection of pictures from the database. Therefore, new techniques and methodologies of Image Retrieval Systems keep flourishing according to the requirements of the users. It is an inevitable process and keeps developing throughout.

REFERENCES

1. Anum Masood, Muhammad Alyas Shahid, Muhammad Sharif, "Content-Based Image Retrieval Features: A Survey", Int. J. Advanced Networking and Applications Volume: 10 Issue: 01 Pages: 3741-3757 (2018) ISSN: 0975-0290

2. T. Dharani, Laurence Aroquiaraj, "Content-Based Image Retrieval System using Feature Classification with Modified KNN Algorithm", Minor Research Project (No.F.41-1361/2012 (SR)) for the financial support

3. Ramakrishna Reddy.Eamani, G.V.Hari Prasad, "Content-Based Image Retrieval Using Support Vector Machine in digital image processing techniques", International Journal of Engineering Science and Technology (IJEST), ISSN: 0975-5462 Vol. 4 No.04 April 2012.

4. M. Arundhathi Reddy, Dr Linganagouda Kulkarni, Dr M. Narayana. "Content-Based Image Retrieval using Color and Shape Features", International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 6, Issue 6, June 2017.

5. MS. PRAGATI ASHOK DEOLE, PROF. RUSHI LONGADGE, "Content-Based Image Retrieval using Color Feature Extraction with KNN Classification", International Journal of Computer Science and Mobile Computing A Monthly Journal of Computer Science and Information Technology ISSN 2320–088X IJCSMC, Vol. 3, Issue. 5, May 2014, pg.1274 – 1280.

6. https://www.kaggle.com/cbir/Data