Image And Video Retreival System

Mohammed Emad Mahmoud
Mohamed Hussien Mostafa
Mohamed Ahmed Abd Alazeem
Mohamed Amr Mohamed
Mohamed Amr Ahmed
Mohamed Khaled Mohamed
Mohamed khaled rashad
Mohamed gamal Talaat

Supervisor: Dr.Gamal A. Ebrahim
Department of Computer Systems Engineering
Faculty of Engineering at Ain Shams University
May 2021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The quantity of virtual content, withinside the shape of snap shots and video, has been growing exponentially in current years. With growing computing strength and digital storage capacity, the capability for big virtual photo/video libraries is developing rapidly. In particular, the World Wide Web has seen an expanded use of virtual snap shots and video, which shape the bottom of many entertainment, educational, and commercial applications. As a result, it has grow to be extra and extra difficult for a consumer to look for the applicable records amongst a big quantity of virtual snap shots or video. Image and video libraries consequently want to offer smooth informational get entry to and the retrieval records ought to be smooth to locate, manipulate and display. As the dimensions of reachable photo and video collections grows to lots of hours, capability visitors will want abstractions and era that assist them browse effectively and efficiently. Text-primarily based totally seek algorithms provide some help in locating precise snap shots or segments amongst big video collections. In maximum cases, however, those systems output many beside the point snap shots or video segments to ensure retrieval of pertinent records. Intelligent indexing systems are crucial for gold standard retrieval of photo and video data. A standard content-primarily based totally photo/video retrieval gadget consists of 3 principal aspects: characteristic extraction, excessive dimensional indexing and gadget layout [24]. Among the 3 aspects, excessive dimensional indexing is vital for speed overall performance; gadget layout is crucial for look overall performance; and characteristic extraction is the important thing to accuracy overall performance. The accuracy overall performance of a retrieval gadget is very subjective and consumer-dependent. To a consumer, the similarity among items is frequently excessive-degree or semantic. However, capabilities we are able to extract from items are frequently low-degree capabilities, as maximum of them are extracted immediately from virtual representations of items withinside the database. The hole among low-degree capabilities and excessive-degree semantics has been the principal impediment to higher retrieval overall performance. In this bankruptcy we discover the today's technology in photo and video retrieval. We describe numerous techniques for extracting capabilities which can be used to degree photo and video similarity in multimedia databases. We additionally describe strategies to bridge the space among low-degree capabilities and excessive-degree semantics.

Chapter 2

Project Description

2.1 Detailed project description

Our project implements content-based multimedia retrieval system, the multimedia techniques to analyze, design, where the two media that we are going to use are the images and the videos.

2.1.1 Image:

We implement three techniques (Histogram - Color Layout - Mean Color)

Histogram:

Convert the image to simple histogram by imageHistogram () function where change 256 bins to 5 bins to filter the sorted images in data base, we get some image which like the simple histogram then we use accurateHistogram () function to get real histogram, when insert the new input image to data base save simple histogram in DB.

We use comparator class which contains some functions such as (add_search_hist () – compare_hist ()), compare_hist () to compare between input image histogram which add_search_hist () save it in class and image histograms in DB.

Color Layout:

In this mode we divide the image into 25 sub-blocks by colorLyout () fn., this function returns histogram for every sub-block, and we deal with every sub-block as image.

Mean Color:

We get average of each channel in input image and get average for them by meanColor() function and search in DB by this average.

2.1.2 Video:

Histogram:

We convert the video into a group of images, every image is keyframe where keyframe is a frame where we define changes in animation. This occurs by KeyFrameEX() where take video and return keyframes then we deal with keyframe as image and read it by getImage() ,in data base we store video as keyframes.

We implement Histogram technique in this multimedia, and deal with video as group of images and take keyframe by videoHisograme () fn. and returns real histogram fer each keyframe.

2.2 Beneficiaries of the project

- Search engines applications will be the main target of the applications as it will enables searching by multimedia contents and develop their searching algorithms.
- End-users can use the application on surfing the internet and search for similar videos and images instead of descriping the content.
- The media can use the project to archive their material and help them dealing with huge amounts of data.

2.3 Detailed analysis

We analyze the project into following points:

1. Objective:
   - Apply different retrival Techniques and deal with new problems.
   - Deal with multimedia systems and how these systems are designed and implemented.
2. Concept of project:
   - The basic concept of this project is how to get retrive matching data for input image or input video.
   - We apply this concept using some retrival techniques.

3. Classification of the project:
   - It is a content based video/image retrival for multimedia systems.
4. Project life cycle:
   - Run the GUI.
   - Insert input image or video.
   - Select retrival technique.
   - Get matching results if exist.
5. Project characteristics:
   - Can deal with image and video inputs.
   - Provide a database to store iamges and videos.
   - Support different retrival techniques.

2.4 Techniques description

2.4.1 Images techniques

Mean color

This technique depends on computing the distance between images based on the color similarity between them. For RGB images the mean color of pixels is computed by finding the average color of the pixels in each channel separately then finding the average between the three values that result from each channel. To get the most similar images to the input image ,the difference between the mean color the input image and each image in the database is computed then we apply a reasonable threshold to execlude the images with large distance.

Mean color is one of the most techniques used in image retreival systems because it can be completed without regard to image size or orientation and it needs less computational power than other techniques.

Histogram

Histogram search algorithms , characterize an image by its color distribution or histogram. A histogram is nothing but a graph that represents all the colors and the level of their occurrence in an image irrespective of the type of the image.

Few basic properties about an image can be obtained from using a Histogram. It can be used to set a threshold for screening the images. The shape and the concentration of the colors in the histogram will be the same for similar objects even though they are of different colors.

Identifying objects in a grey scale image is the easiest one as the histogram is almost similar as the objects have the same colors for same objects. In order for identifying the objects in the images or generating the histogram the system has to obtain the array values of the frequency of occurrence of each color value -from 0 to 255- in the image.

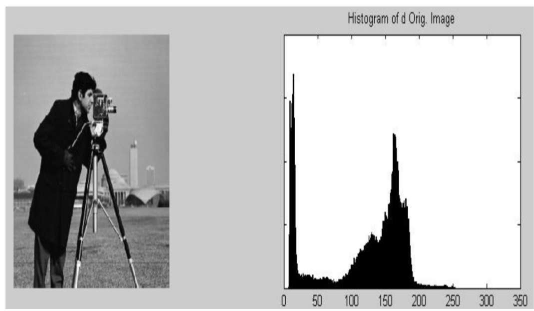

Figure 2.1: example for image histogram

To calculate the distance between two image histogram we calculate the sum of the smallest bin for each corresponding bins in the two histograms for input image I and the model image M normalized to the number of pixels in the model image.

$$\sum_{J=1}^{n} \frac{min(Ij, Mj)}{\sum_{J=1}^{n} Mj}$$

Figure 2.2: histogram distance equation

Color Layout

This technique is similar to histogram based technique except it solves the problem of getting results of images with a low histogram distance value but with a different contents.

In this algorithm we divide each image into 5x5 array of blocks so we get 25 sub-image then we calculate the histogram form each block. To find the distance between to images we get the distance between the histograms of each two corresponding block, then we calculate the total distance by finding the summation of all blocks distance.

$$d_{gridded\_square}(I,Q) = \Sigma\, d_{color}(C^{I}(g),C^{Q}(g))$$

Figure 2.3: color layout distance equation

2.4.2 Videos techniques

Feature extraction technique

To make a content based video retrieval system, we have to deal with videos as frames and compare frame with each others as images. But applying this technique would consume a lot of memory to store the whole frame of each video. So the key frames role appear here.

A key frame (or keyframe) in animation and filmmaking is a drawing or shot that defines the starting and ending points of any smooth transition. These are called frames because their position in time is measured in frames on a strip of film or on a digital video editing timeline.

So we make a key frame extraction operation for each video before inserting it to the database. then we deal with key frames as the same manners of images, We caculate the histogram for each key frame and divide the range into five regions then we get the averge of each region resulting the five values that we store in the data base, after that we use these values as a initial filtering to get the most similar frames of the input ,then we use the accurate -255 values-histogram to make a second step filtering.

Distance calculation

To get the distance between two videos we use the naive video similarity technique. In this technique we calculate how many key frames in the query video is similar to one or more key frame in the model video, then we calculate the ration between these key frames to the total number of the key frames, after that we compare the distance with a threshold valueand if the distance is lower than this threshold, the two videos are similar otherwise they are not similar.

Chapter 3

Project planning

3.1 Task breakdown structure

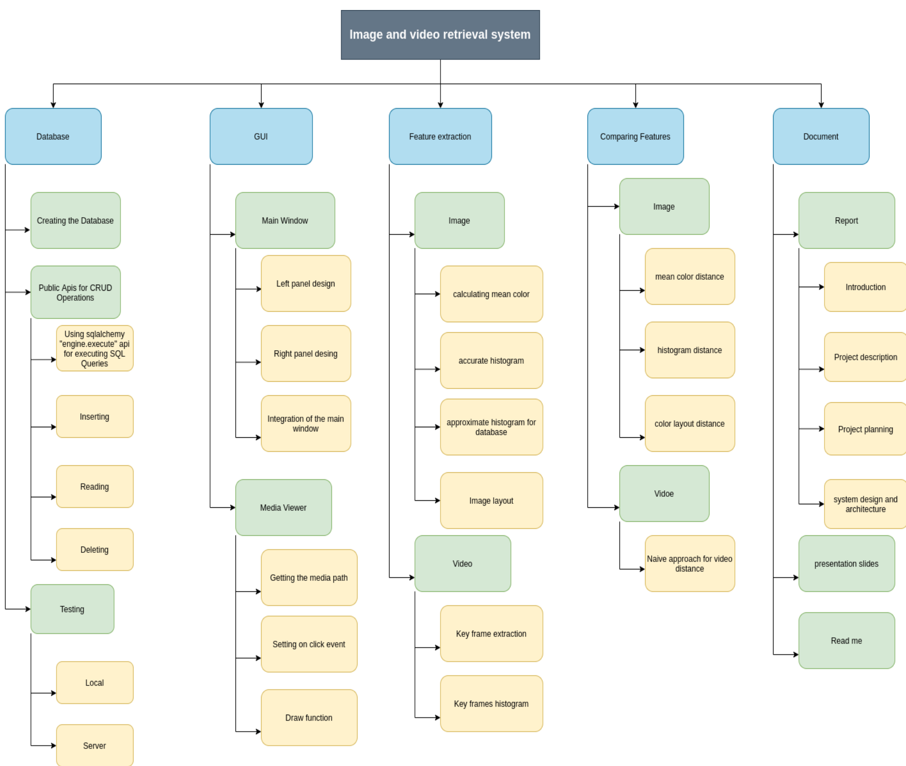

Figure 3.1: Tasks break down structure

3.2 Time plan and Gant chart

Firstly, we split all project team members into 3 basic groups in according to simplify the overall problem into smaller ones that we can deal with. We decided to start with designing the database and find best modelling ways to fit our storage requirements. The second group mission is to implement software functions that do the following:

1. extract data from videos/images.
2. store the different representations of input data into the database.
3. impelemnt the required techinques that fit project targets.

According to third group, it provides the project with a well-interactive graphical user interface. Finally, the last group provides the project full documentation.

We divided our team into the following:

1. Mohammed Khaled Rashad and Mohammed Hussien.
2. Mohamed Amr Ahmed, Mohamed Amr Mohamed, Mohamed Khaled Elkhawas and Mohamed Ahmed Abdel Azem.
3. Mohammed Gamal Talaat and Mohammed Emad.
4. Mohammed Hussien Mostafa, Mohammed Ahmed Abd el Azeem and Mohammed Emad.

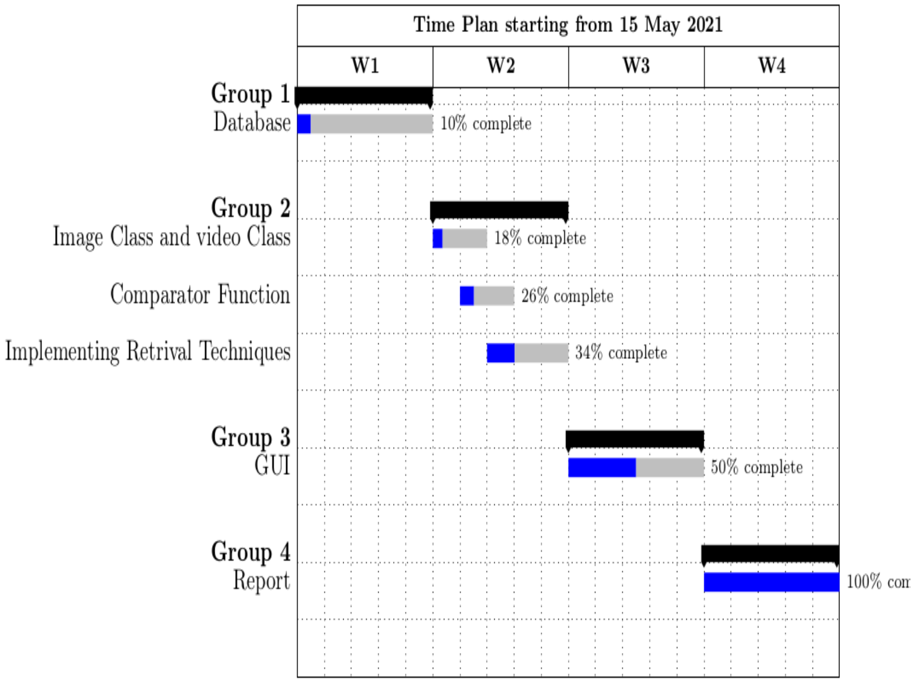

Figure 3.2: Time Plan starting from 15 May 2021

Chapter 4

System Design And Architecture

4.1 System architecture

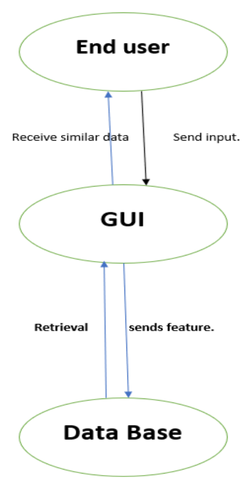

Figure 4.1: system design architecture

4.2 Multimedia database design

4.2.1 The purpose of our data base

The puspose of our data base is to work as images and videos search engine server.This is achieved by storing images and videos with their pre computed extracted features then takes a an image or video as input form the user and query the data base storge based on the extracted features of the input material and outputs the most relevant matrial to the input model.

The project database supports three filter techniques for images retreival operation which are mean color, histogram, and color layout and one technique for video retreival operation base on histogram of the key frames.

4.2.2 Required information

To achieve the required techniques, we need our database to hold some information about each stored element.

1. Images features
    - We store the mean color value for each image. Mean color is calculated by finding the average of the pixels values for each channle separatley, then finding the averge between the three values resulting from each channel.
    - Histogram information is needed to be stored,but if we store the whole 255 values of each histogram ,we will consume huge amount of storage. The solution we have used is to divide the histogram into five regions and get the averge of each region so we convert the 255 values into 5 values only and store them in the database. we need these five values for only first step in filtering and then we compare using the complete histogram -255 values- for the filtered images.
2. Video features
    - Instead of storing the whole video frames we only store the information about the key frames of the video. Each key frame is related to its video through "ID", then we deal with key frames in the same manner of the images so we store five values for the histogram of each key frame.

4.2.3 Database Schema

The data base consists of three main tables :

- The first tables is "Image" which holds the information for each image.
- The second is "Video" which holds only the ID and Path of each video.

- The third table is "Key Frame". this table holds all the key frames of all stored videos and relate each key fram to its video using the video ID.

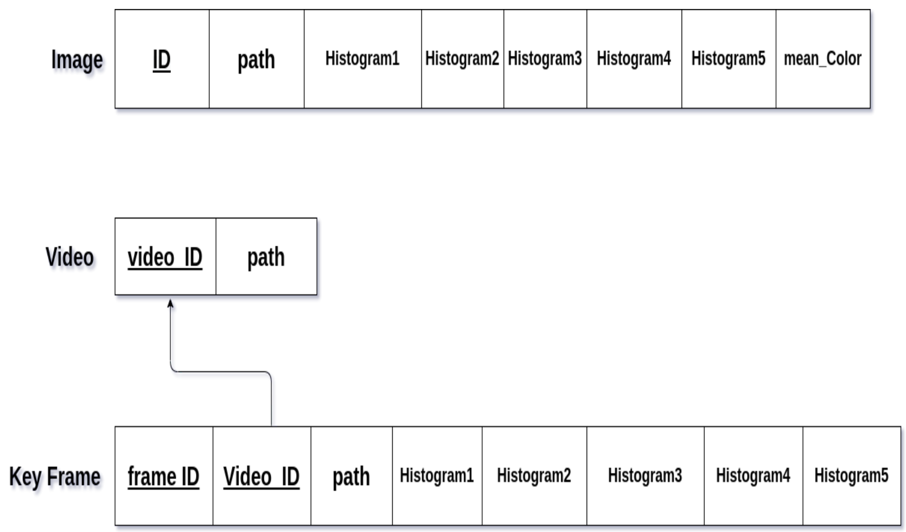

Figure 4.2: Data base schema

4.2.4 Primary keys and foriegn keys

- Primary keys of each table are:
  1. Image table Primary key is the Id of each image.
  2. Video table Primary key is the Id of each video.
  3. Key frame table Primary key is a composite key from the id of each key frame and the id of the refered video.
- Foriegn keys used are:
  - Key frame table has the referenced video id as a foriegn key which relates each key frame to its original video.

4.3 System design

We have designed and implemented some classes to simplify our target project.

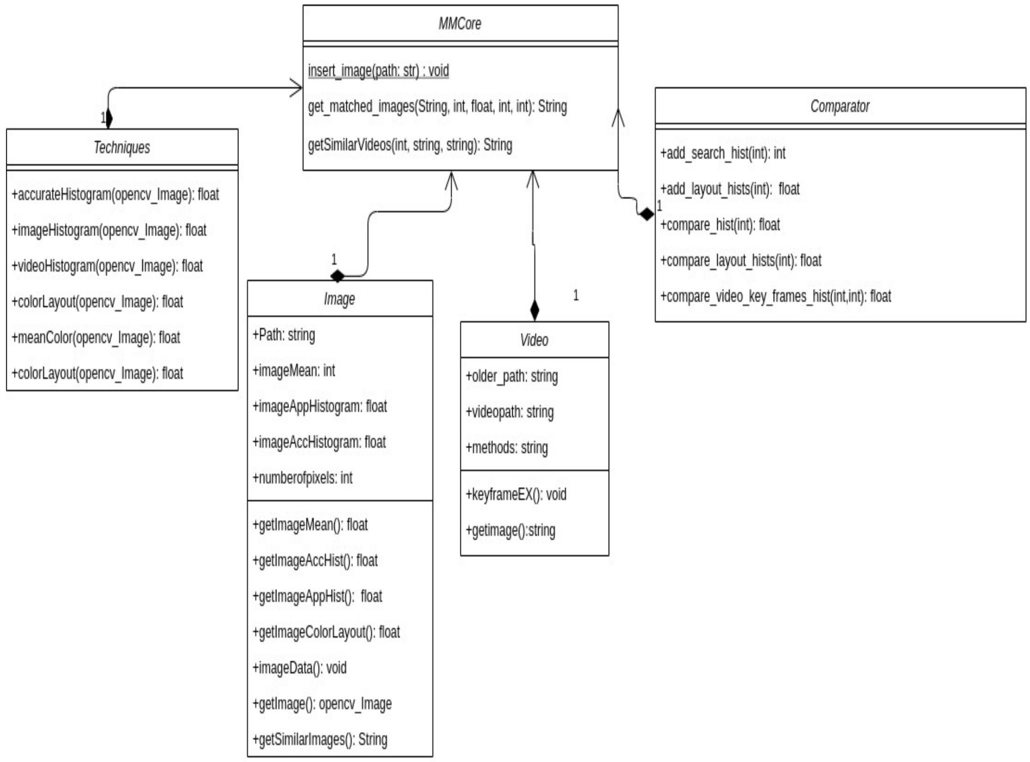

Figure 4.3: Class UML

4.4 Testing scenarios and results

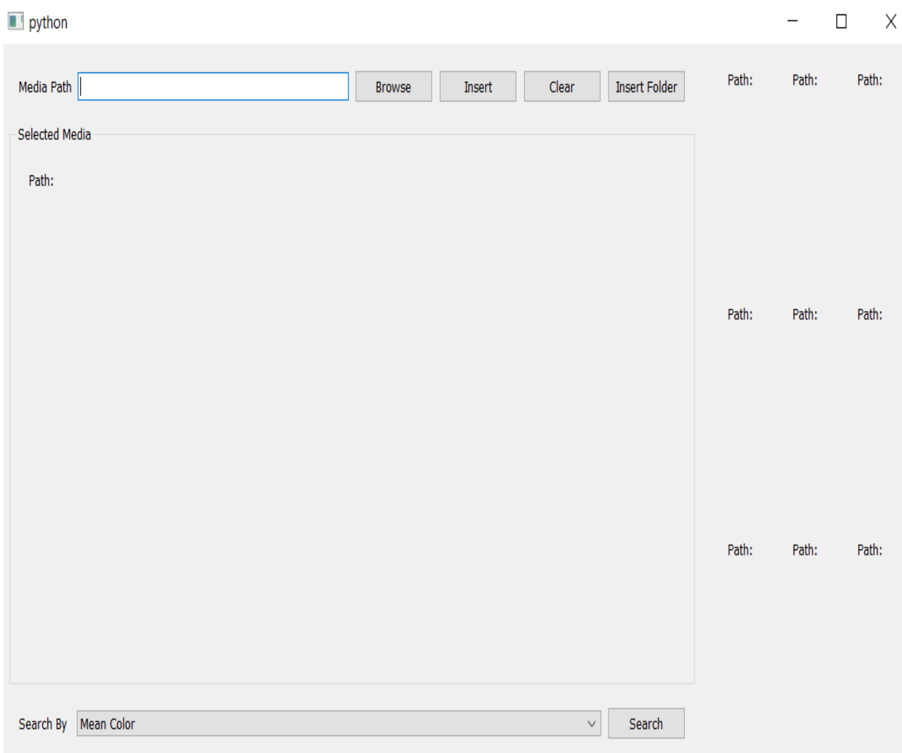

Figure 4.4: Testing user interface

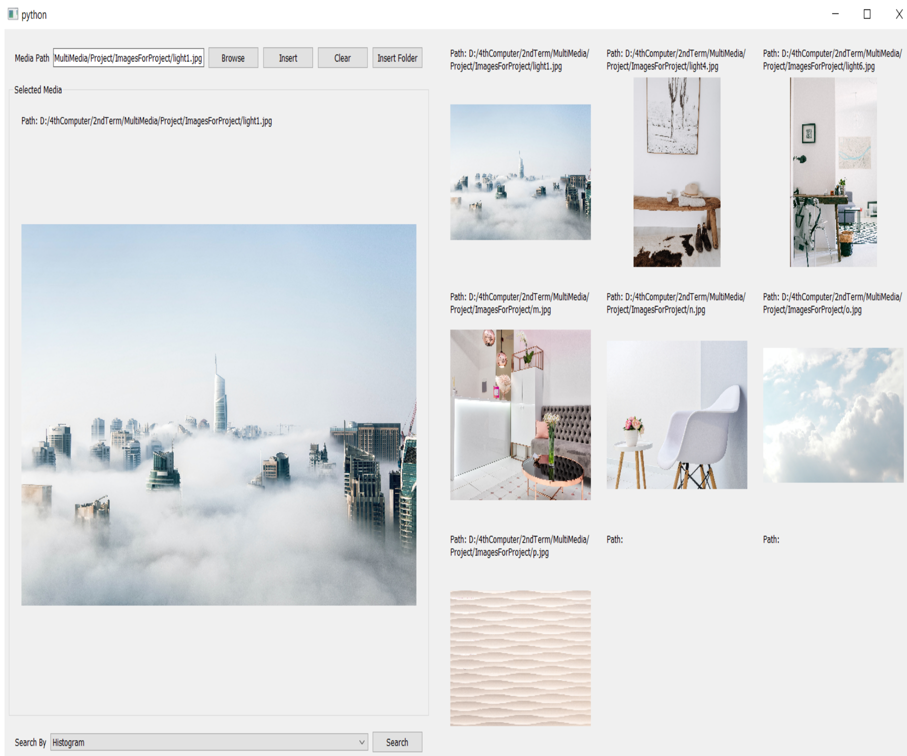

Figure 4.5: Testing Images with histogram

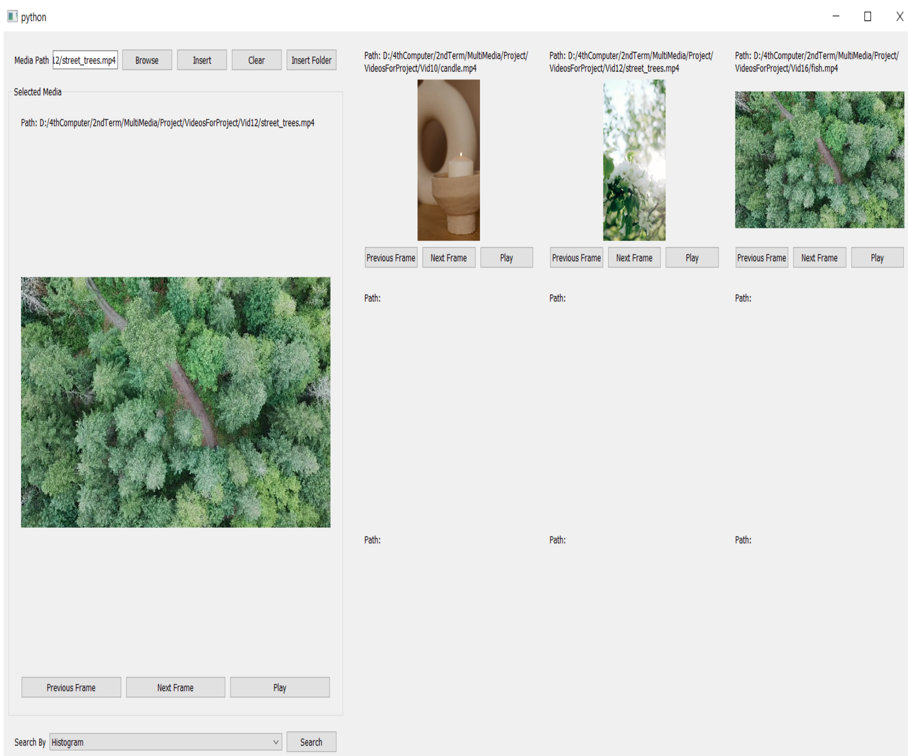

Figure 4.6: Testing Videos with histogram

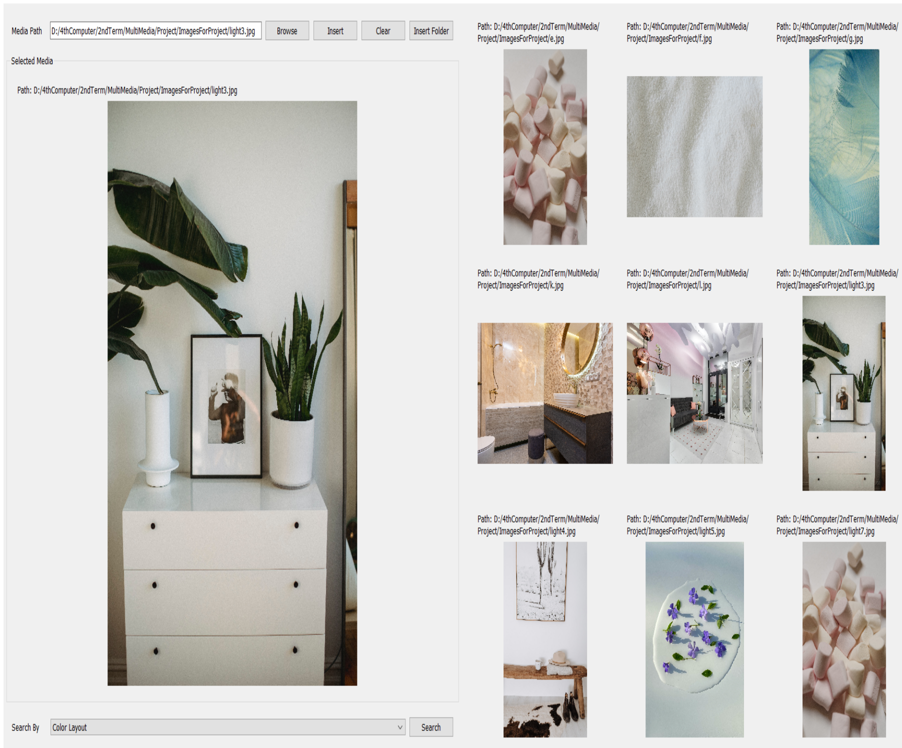

Figure 4.7: Testing with color layout

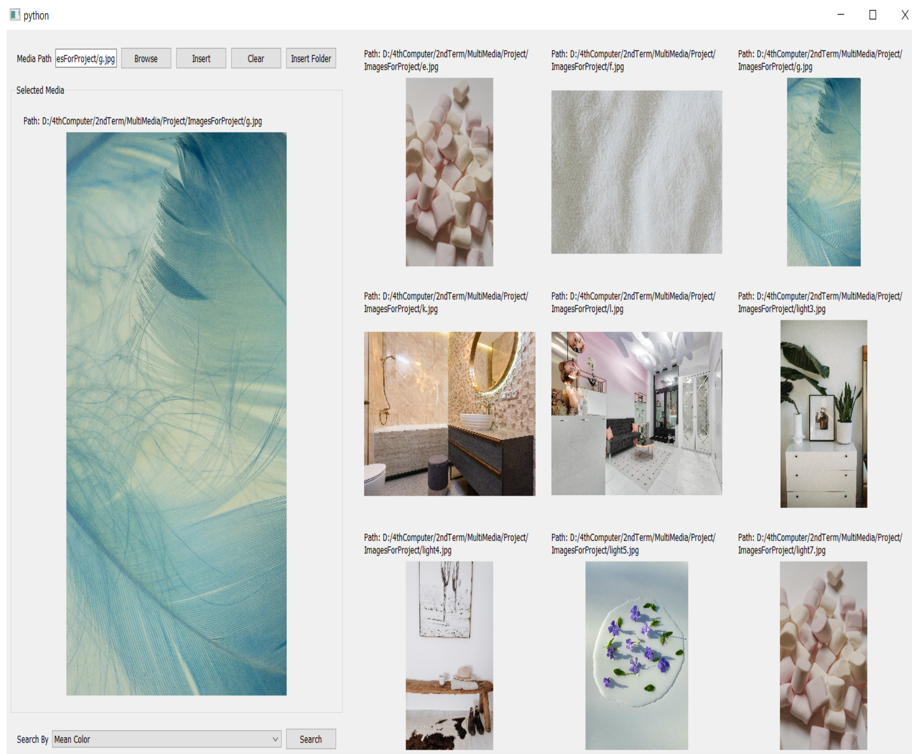

Figure 4.8: Testing mean color

Chapter 5

End User Guide

5.1 Steps To use the project

1. Clone the repository and then run main.py.
2. Install the dependencies mentioned:

openCV2 - Reading input images and videos.
pyside2 - open source python gui framework.
sqlalchemy - connects python with database.
video-kf - helps extracting video key frames.

5.2 Features

1. Retrive input video with Naive Video Similarity (by key frames Histogram) Technique.
2. Retrive input images with Mean Color and Color Layout Technique.
3. Support a well designed graphical user interface.

And of course our project itself is open source with a [https://github.com/mohammedBadawi/Multimedia-Project-Document.git][dill] on GitHub.

Chapter 6

Conclusion

CBIR is a fast-developing technology with considerable potential in digital libraries, architectural and engineering design, crime prevention, historical research and medicine. Nevertheless, the effectiveness of current CBIR systems is inherently limited because they most effective operate at the primitive feature level. Furthermore, the technology still lacks maturity, and is not widely used on a significant scale. Consequently, examine examines exclusive techniques used in CBIR systems. The examine reviewed numerous literatures that are related to CBIR. The examine located out that there are 3 basic capabilities that can be extracted in CBIR. These consist of shade, texture and shape. The examine additionally discovered that each of those capabilities has exclusive extraction techniques. For instance, shade can be extracted in images using shade histogram, geometric moments, shade area Olaleke et al.; AJRCOS, 3(2): 1-15, 2019; Article no.AJRCOS.48235 14 and shade moments. The examine discovered the strengths and weaknesses of each of those techniques. For instance, the shade area technique is straightforward to put in force however it isn't uniform while the shade histogram is faster and greater green than other shade extraction techniques. It can however be identical for two images with exclusive colours. The examine additionally famous that the GLCM, Tamura, Fourier transform, Ranklet transform and discrete wavelets are ordinary examples of textural extraction techniques. Similarly, the edge technique, Fourier descriptors and Zernike technique were the shape extraction techniques discovered on this examine. Furthermore, the examine investigated the techniques for computing the similarity between a query image and the images in the database. The result of the examine showed that examples of similarity measures utilized in CBIR consist of sum of absolute difference, sum of the squared differences of absolute values and metropolis block distance. In latest times, there's no popular leap forward in CBIR regardless of the numerous techniques and equipment developed to formulate and execute queries in large databases based on their visual contents. Hence, future works should be tailored closer to the development of CBIR systems that will solve the trouble of semantic hole in CBIR.

Chapter 7

References

- Aigrain, P et al (1996) "Content-based representation and retrieval of visual media - a state-of-the-art review" Multimedia Tools and Applications 3(3), 179-202.
- Ashley, W (1996) "What shoe was that? The use of computerised image database to assist in identification." Forensic Science International, 82, 7-20
- IEE Proc.-Vis. Image Signal Process., Vol. 152, No. 6, December 2005 851
- Image and Video Indexing and Retrieval Michael A. Smith and Tsuhan Chen Carnegie Mellon University.
- Alsuth, P et al (1998) "On video retrieval: content analysis by ImageMiner" in Storage and Retrieval for Image and Video Databases VI, Proc SPIE 3312, 236-247
- Mandal, Mrinal Kr. Multimedia Signals and Systems. Kluwer Academic Publishers, 2003.
- Chen, Tsuhan, and Atul Puri. Multimedia Systems, Standards, and Networks. Dekker, 2000.