VIDEO SUMMARIZATION USING HISTOGRAM ANALYSIS

Submitted in partial fulfilment of the requirements

Of the degree of

Bachelor of Engineering

By

Yash Patange- (Roll No.48)

Prashant Porwal - (Roll No.53)

Akhil Raut - (Roll No.55)

Supervisor:

PROF.SMITA JAWALE

DEPARTMENT OF COMPUTER ENGINEERING

VIDYAVARDHINI'S COLLEGE OF ENGINEERING AND TECHNOLOGY

K.T.MARG, VASAI ROAD (W), DIST-PALGHAR. PIN:401202.

(Affiliated to University of Mumbai)

(2018-2019)

CERTIFICATE

This is to certify that the project entitled "**Video Summarization using Histogram Analysis**" is a bonafide work of "**Yash Patange (48), Prashant Porwal (53) and Akhil Raut (55)**" submitted to the University of Mumbai in partial fulfilment of the requirement for the award of the degree of **"Undergraduate" in "Computer Engineering."**

_____

Prof. Smita Jawale

Guide

_____ _____

Dr. Megha Trivedi **Dr. Harish Vankudre**

Head of Department Principal

Project Report Approval for B.E.

This project report entitled "**Video Summarization using Histogram Analysis**" by "**Yash Patange (48), Prashant Porwal (53) and Akhil Raut (55)**" is approved for the degree of **Bachelor of Engineering (Computer Engineering)**.

Examiners

1. _____

2. _____

Date:

Place:

Declaration

I declare that this written submission represents my ideas in my own words and where other's ideas or words have been included. I have adequately cited and referenced the original sources. I also declare that I have adhered principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

_____

Yash Patange (48)

_____

Prashant Porwal (53)

_____

Akhil Raut (55)

Date:

Acknowledgement

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. We would like to extend our sincere thanks to all of them.

We are highly indebted to **Prof. Smita Jawale** for her guidance and constant supervision as well as for providing necessary information regarding the project and also for their support in completing the project.

We would like to thank our Principal **Dr. Harish Vankudre** for constant motivation and support to excel and having faith in our ability. We would also like to thank our professor **Dr. Megha Trivedi** (Head of Department of Computer Engineering) for providing her views of the subject.

We would like to express our gratitude towards our parents and staff of Vidyavardhini's College of Engineering & Technology for their kind co-operation and encouragement which helped us in completion of this project.

Our thanks and appreciation also goes to our colleagues in developing the project and people who have willingly helped us out with their abilities.

Abstract

Everybody in our world is so busy with their lives that taking the time out for some activities is nigh difficult. Not everybody can sit through an entire match of their favourite team and support them. Neither all can watch an entire live stream of an important event occurring in history. Through this project, we aim to condense the entire full length video into a 'highlight stream'.

Our project, Video Summarization cuts and summarizes an entire video into a small video clip of relevant information. Thus allowing all those people who couldn't get time to watch the stream, watch the highlight stream and still enjoy it. We will provide an interface through which the user can specify which video he/she wants to summarize. Once the video is selected summarization process begins and then after its end user has the summarized version of the original clip.

Table of Contents

Chapter 1

INTRODUCTION

In recent years, because of the rapid growth in multimedia information, the advance in internet communication and digital video technologies, multimedia information indexing and retrieval has become more and more important and lots of research efforts have been devoted to the video retrieval and video analysis based on audio or visual features. This analysis shows that, when developing retrieval applications and video indexing, we first have to consider the issue of structuring the huge and rich amount of heterogeneous information related to video content. In addition, to retrieve information from the audio or visual content is a very challenging since it requires the extraction of high-level semantic information from low level audio or visual data.

Video summarization is an important process that facilitates faster browsing of large video collections and also more efficient content indexing and access. There are two main video summarization techniques in the literature:

- Static video summarization (video summary)
- Dynamic video summarization (video skimming).

In order to summarize a video, most of the methods have consists on visual features computed from video frames. Also, there are methods that consider the semantic meaning implied in the video to produce a more informative summary.

1.1 Introduction

Due to the increasing volume of video content on the Web, and the human effort taken to process it, new technologies need to be researched in order to develop efficient indexing and search techniques to manage effectively and efficiently the huge amount of video data. One of the most evolving research areas is Video summarization. As the name implies, video summarization is a mechanism to produce a short summary of a video to give to the user a synthetic and useful visual abstract of video sequence, it can either be a images (keyframes) or moving images (video skims).

1.2 Aim and Objective

The main aim of Video summarization is to provide clear analysis of video by removing duplications and extracting key frames from the video. Massive growth in video content poses problem of information overload and management of content. In order to manage the growing videos on the web and also to extract an efficient and valid information from the videos, more attention has to be paid towards video and image processing technologies. Video summarization is a mechanism to produce a short summary of a video to give to the user a synthetic and useful visual abstract of video sequence; it can either be an image (key frames) or moving images. Video summarization is a vital process that facilitates well-organized storage, quick browsing, and retrieval of large collection of video data without losing important aspects. In terms of browsing and navigation, a good video abstract will enable the user to get maximum information about the target video sequence in a specified time limitation or adequate information in the minimum time.

1.3 Scope

A Video Summarization is a summary representing an abstract view of the original video sequence and can be used for video browsing and retrieval systems. Different methods used to select key frames. These methods are mainly based on low level features such as colour Histogram, edge Histogram, frame correlation and it does not consider the colours in the image. Techniques in video summarization touch various domains, such as movies, sports, news, home videos, e-learning, etc. Different techniques used for video summarization such as object base summaries, event based summaries, content based summaries, feature based summaries etc.

Chapter 2

REVIEW OF LITERATURE

2.1 Domain Explanation

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps:

- Importing the image via image acquisition tools;
- Analyzing and manipulating the image;
- Output in which result can be altered image or report that is based on image analysis.

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction.

Different sources of digital images will be discussed and examples for each source will be provided. The continuum from image processing to computer vision will be covered in this lecture. Finally we will talk about image acquisition and different types of image sensors.

Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

2.2 Existing Solution

A video summarization is a summary which represents abstract view of original video sequence and can be used as video browsing and retrieval systems. It can be a highlight of original sequence which is the concatenation of a user defined number of selected video segments or can be a collection of key frames. Different methods can be used to select key frames.

By using triangle model of perceived motion energy (PME) motion patterns are modeled in video. The frames at the turning point of the motion acceleration and motion deceleration are selected as key frames. The key-frame selection process is threshold free and fast and the extracted key frames are representative.

In Visual frame Descriptors algorithm three visual features: color histogram, wavelet statistics and edge direction histogram are used for selection of key frames. Fidelity, Shot Reconstruction Degree, Compression Ratio qualities are used to evaluate the video summarization.

Motion focusing method focuses on one constant-speed motion and aligns the video frames by fixing focused motion into a static situation. A summary is generated containing all moving objects and embedded with spatial and motion information. Background subtraction and min cut are mainly used in motion focusing.

In Camera Motion and Object Motion, the video is segmented using camera motion-based classes: pan, zoom in, zoom out and fixed. Final key frame selections from each of these segments are extracted based on confidence value formulated for the zoom, pan and steady segments.

2.3 Hardware and Software requirements

2.3.1HARDWARE REQUIREMENTS:

- Processor: Intel i3/ AMD FX 6300 or higher.
- RAM: 4 GB or higher.
- Graphics: Integrated Graphics or better.

2.3.2 SOFTWARE REQUIREMENTS:

- Operating System: Windows/Linux.
- Language Used: Python
- Software Tools: IDE, Text Editor, Jupyter Notebook.

Chapter 3

REQUIREMENT ANALYSIS

3.1 Functional requirements

- **VIDEO FORMAT**:
  User can upload any format of video to convert it into highlight stream.
- **VIDEO LENGTH**:
  User can upload the video with length up to 10-12 minutes.
- **PROCESSING TIME**:
  The video will be processed and converted into its highlight stream within 10-20 minutes.
- **VIDEO QUALITY**:
  The quality of the video is maintained in the shortened clip.

3.2 Non-functional requirements

- **PERFORMANCE**:
  The system performs efficiently for any format or any size of video to be uploaded and creates the desired highlight stream with constant response time and throughput.
- **USABILITY**:
  The system is easy to use and learn.
- **ACCURACY AND PRECISION**:
  The video converted will be accurate and précised up to 93%.
- **PLATFORM CONSTRAINTS**:
  The system has to be operated on WINDOWS and LINUX OPERATING SYSTEMS.
- **RELIABILITY**:
  The System is reliable for the videos with time duration below 10-12 minutes.

3.3 Proposed System

Our Model explores a method of **KEY-FRAME EXTRACTION ALGORITHM** based on a **K-Means Clustering Algorithm.**

Keyframes are those frames that define a start and end point in a smooth transition. They are called "frames" because they are measured in frames on a strip of film. Keyframes define the timing as well as what the viewer will see in the video. Since only 2 or 3 keyframes are present in a span of a second "inbetweener" frames are used to fill out these gaps between the keyframes.

Frame Differencing

In frame differencing method there is no need to identify the shot boundary. Whenever drastic change is identified, the next frame is selected as key frame. The identification of change is depended on threshold value. If threshold value is large, then only drastic change can identified and if it is less, then minor change can identified. The more the threshold value, the less the number of key frames, the less the threshold value, the more the number f key frames. So the number of key frames can be controlled by the threshold value.

Clustering

Clustering is a process of grouping data or objects that are more similar to each other than the other data or objects present in a given set. It is a common technique of statistical data analysis and is used widely in machine learning

We want to group all the similar histograms into the k clusters. Each histogram is representative of the corresponding video segment. Our version of the K-means algorithm is defined below:

1. Select k random centroid points on our multi-dimensional space.
2. Compute each histogram against all the cluster centroids.
3. Each histogram is assigned to the cluster that minimizes the error function.
4. Recompute cluster centroids
5. On every iteration, check to see if the centroids converged. If not, we go to step 2.

We use Euclidean distance as our error function. This is the general approach when directly comparing histograms.

Chapter 4

DESIGN

4.1 Design Consideration

Extensibility

New capabilities can be added to the software without any major change to the existing architecture.

Modularity

The software is modularized, so it is easy to maintain. New features can be added and tested in isolation without causing any problem to the main feature.

Maintainability

The software is highly maintainable and bus can be fixed easily

Usability

The user interface provided is very easy to use. It is designed keeping in mind the layman users.

Portability

The software can work under different conditions and environment.

Scalability

The software adapts well to increasing data or number of users.

4.2 Design Details

4.2.1 Entity Representation Diagram

E-R Diagram

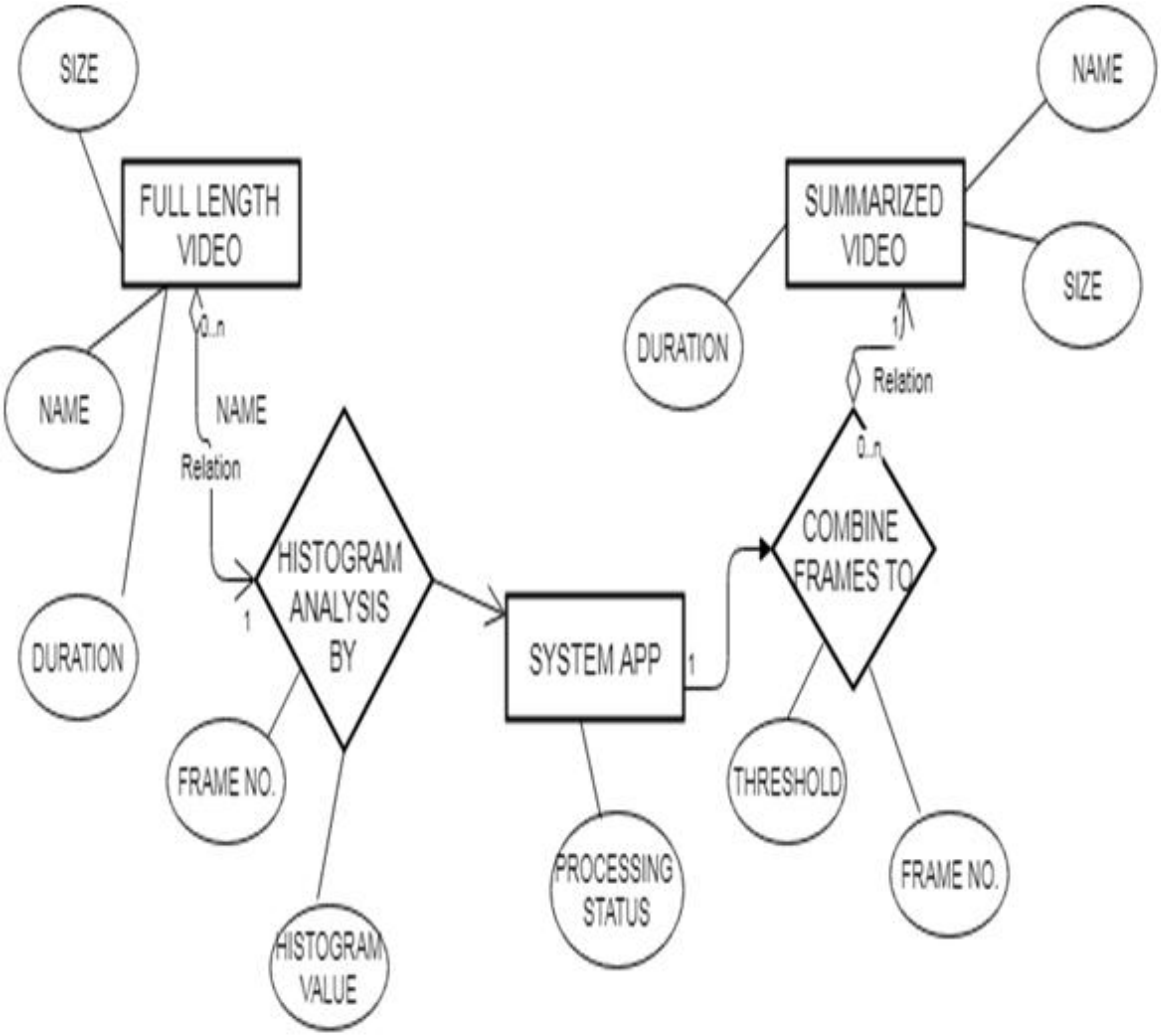
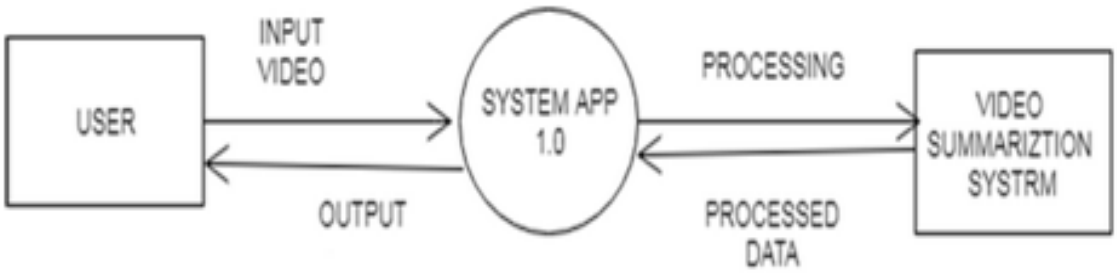


Figure 4.2.1 : Entity Representation diagram

4.2.2 Data Flow Diagrams:

Level 0:



Level 1:

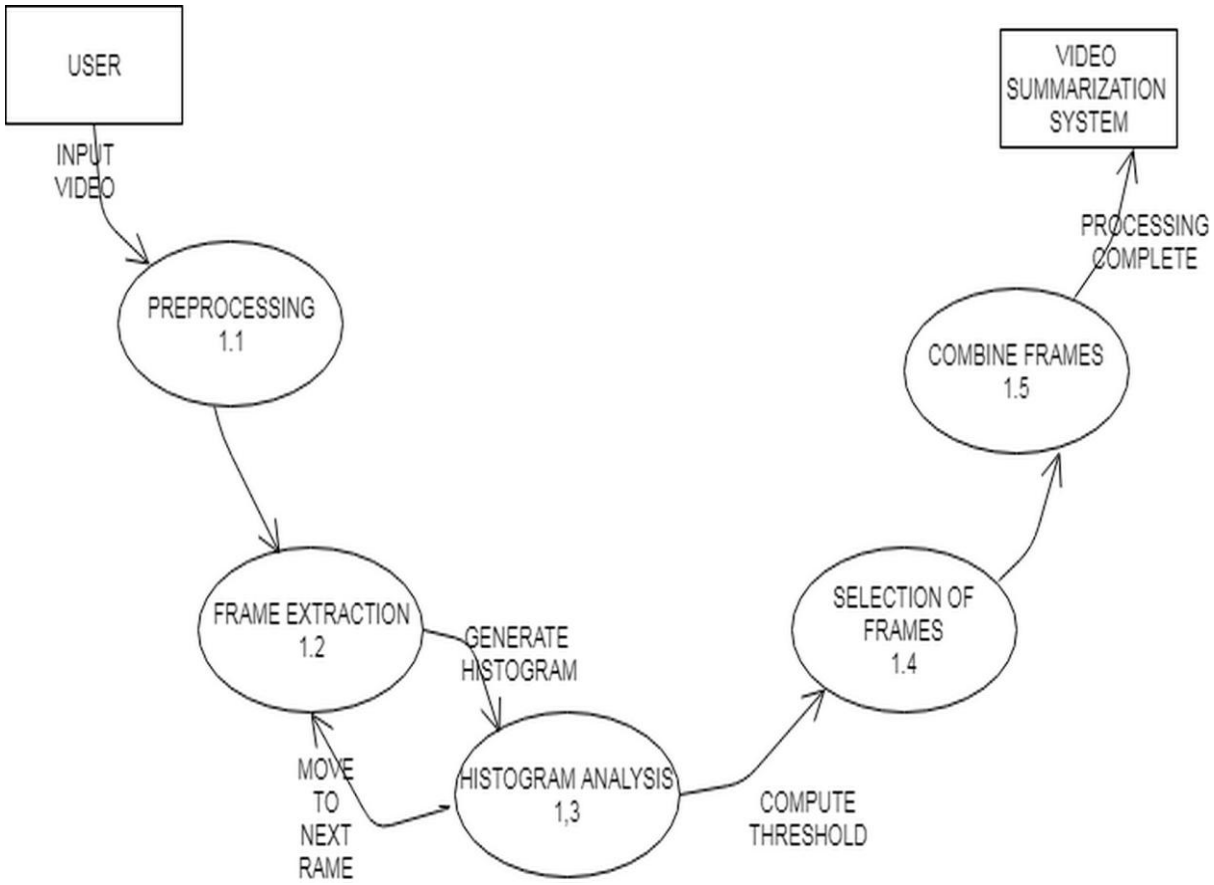


Figure 4.2.2:DFD

4.2.3 Activity Diagram

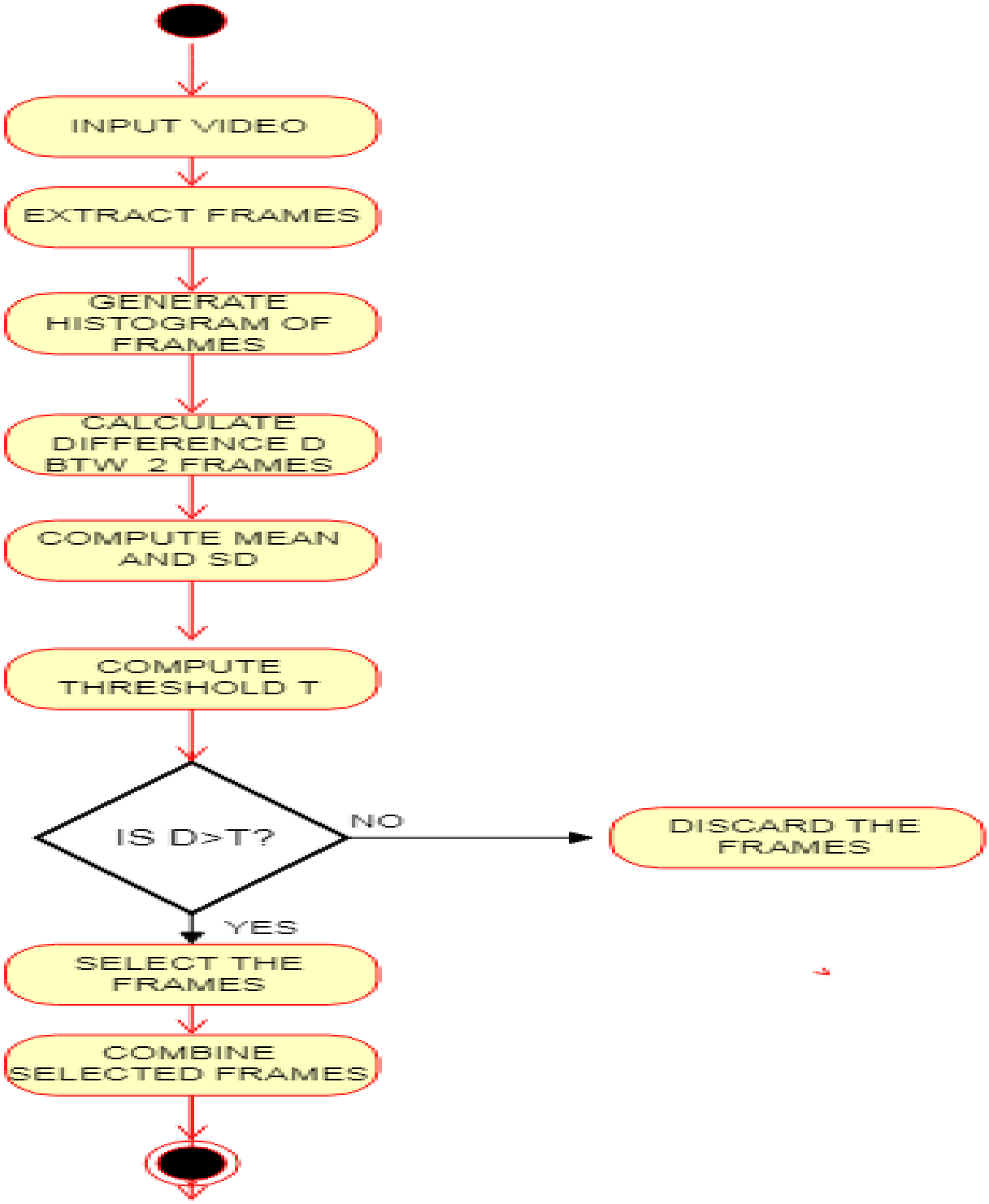
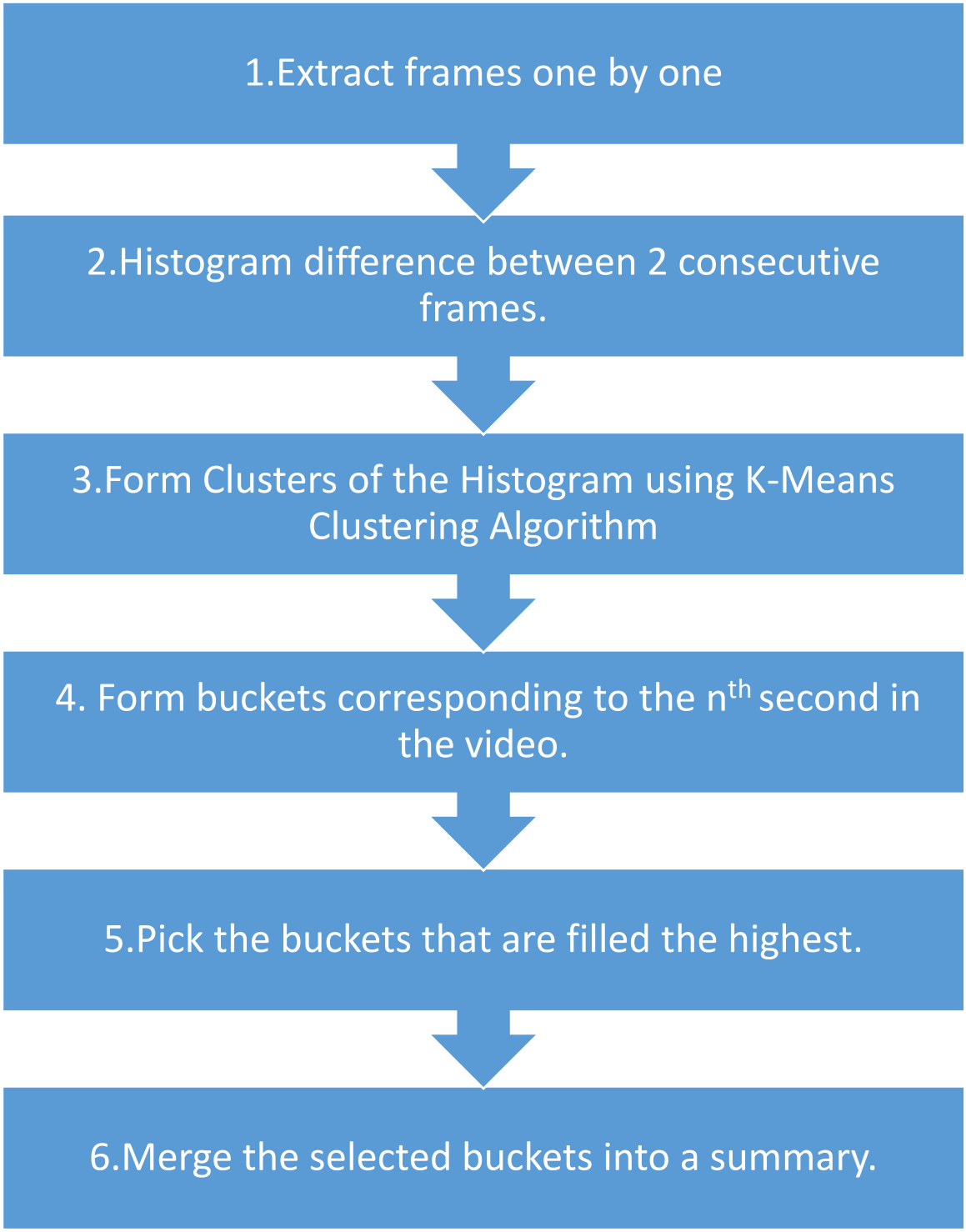


Figure 4.2.3: Activity diagram

Chapter 5

IMPLEMENTATION METHODOLOGY



5.1 Modules

5.1.1 Frame Generator

Frame Generator as the name suggests is used to generate frames from the video. The video is first selected through the desktop app and the processing begins. The first step of processing is to generate the frames of the video. The frames are generated based on the frame rate of the video to ensure video quality is maintained.

5.1.2 Histogram Analyzer

Once the frames are generated histograms are calculated. Then histogram differencing is performed to find the key frames. Once key frames are found, the corresponding video clips are extracted from the original video.

5.1.3 Clips Merger

This module is used to concatenate the key video clips extracted in the previous video. The output of this module is the final summarized video.

5.2 Sample Code

Frame Generation

```
vidcap = cv2.VideoCapture(path)
    print path
    success,image = vidcap.read()
    total_count = 0
    success = True
    while success:
        success,image = vidcap.read()
        cv2.imwrite("test%d.jpg" % total_count, image)
        print(total_count)
              if cv2.waitKey(10) == 27:
                  break
              total_count += 1
```

Histogram Analysis

```
counter=0
    index = {}
    images = {}
    results =[]
    while counter<total_count:
        name="test%d.jpg" % counter
        image = cv2.imread(name)
        images[name] = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
        hist = cv2.calcHist(image, [0, 1, 2], None, [8, 8, 8],[0, 256
  , 0, 256, 0, 256])
              hist = cv2.normalize(hist, hist).flatten()
              index[name] = hist
              counter+=25

          counter=0
          i=0
          for (k, hist) in index.items():
              results.append([])
              name1="test%d.jpg" % counter
              name2="test%d.jpg" % (counter+25)
              if(counter<total_count-25):
```

```
                    d = cv2.compareHist(index[name1], index[name2], cv2
.HISTCMP_CORREL)
                    results[i].append(d)
                    results[i].append(1)
                    counter+=25
                    i+=1
            print results
```

K-Means Clustering

```
dataset = pd.read_csv('test.csv')
    X = dataset.iloc[:,:].values

    wcss = []
    for i in range(1,11):
        kmeans = KMeans(n_clusters = i, init= 'k-
means++', max_iter=300, n_init=10, random_state=0)
        kmeans.fit(X)
        wcss.append(kmeans.inertia_)
    plt.plot(range(1,11),wcss)
        #plt.show()
        kmeans = KMeans(n_clusters =3, init= 'k-
means++', max_iter=300, n_init=10, random_state=0)
        y_kmeans = kmeans.fit_predict(X)
        plt.scatter(X[y_kmeans==0,0],X[y_kmeans==0,1], s=100, c='re
d', label='One')
        plt.scatter(X[y_kmeans==1,0],X[y_kmeans==1,1], s=100, c='gr
een', label='Two')
        plt.scatter(X[y_kmeans==2,0],X[y_kmeans==2,1], s=100, c='cy
an', label='Three')
        plt.scatter(kmeans.cluster_centers_[:,0],kmeans.cluster_cen
ters_[:,1],s=300, c='yellow',label='centroids')
```

Desktop App

```
def exitapp():
    tkMessageBox.showinfo("Video Summarization","Exiting the app...")
    answer = tkMessageBox.askquestion("Video Summarization","Do you
really want to exit?")
    if answer == 'yes':
        root.destroy()

def browse_button():
    file = tkFileDialog.askopenfilename(initialdir = "./",title = "Sel
ect file",filetypes = (("mp4 files","*.mp4"),("all files","*.*")))
```

```
    generate_frames(file)

        def about_app():


            quote = """The main aim of Video summarization is to
provide clear analysis of video by
        removing duplications and extracting key frames from the video.
Massive
        growth in video content poses problem of information overload
and
        management of content. In order to manage the growing videos on
the web and
        also to extract an efficient and valid information from the
videos, more attention
        has to be paid towards video and image processing technologies.
Video
        summarization is a mechanism to produce a short summary of a
video to give to
        the user a synthetic and useful visual abstract of video
sequence; it can either be
        an image (key frames) or moving images. Video summarization is
a vital
        process that facilitates well-organized storage, quick
browsing, and retrieval of
        large collection of video data without losing important
aspects. In terms of
        browsing and navigation, a good video abstract will enable the
user to get
        maximum information about the target video sequence in a
specified time
        limitation or adequate information in the minimum time."""
            label = Label(root, text=quote, pady=10,padx=0, justify=LEF
T)
            label.config(font=("Arial",20))
            label.pack(side=TOP)

            label2 = Label(root,pady=30, text="Developed By\n AKHIL
RAUT\n PRASHANT PORWAL\n YASH PATANGE")
            label2.config(font=("Arial",30))
            label2.pack(side=TOP)
            root.mainloop()



        #desktop application
        root = Tk()

        root.title("Video Summarization")

        myMenu = Menu(root,bg="blue")
        myMenu.add_cascade(label="Select File",command=browse_button)
```

```
46.       myMenu.add_cascade(label="About",command=about_app)
47.       myMenu.add_cascade(label="Exit",command=exitapp)
48.
49.       root.configure(menu=myMenu)
50.       toolbar =Frame(root)
51.       Try1 = Button(toolbar, text="Summarize
   video", bg="#acd9e2",command="")
52.       Try1.pack(side=LEFT,padx=5,pady=5)
53.       toolbar.pack(side=BOTTOM,fill=X)
54.       root.mainloop()
```

Chapter 6

RESULT

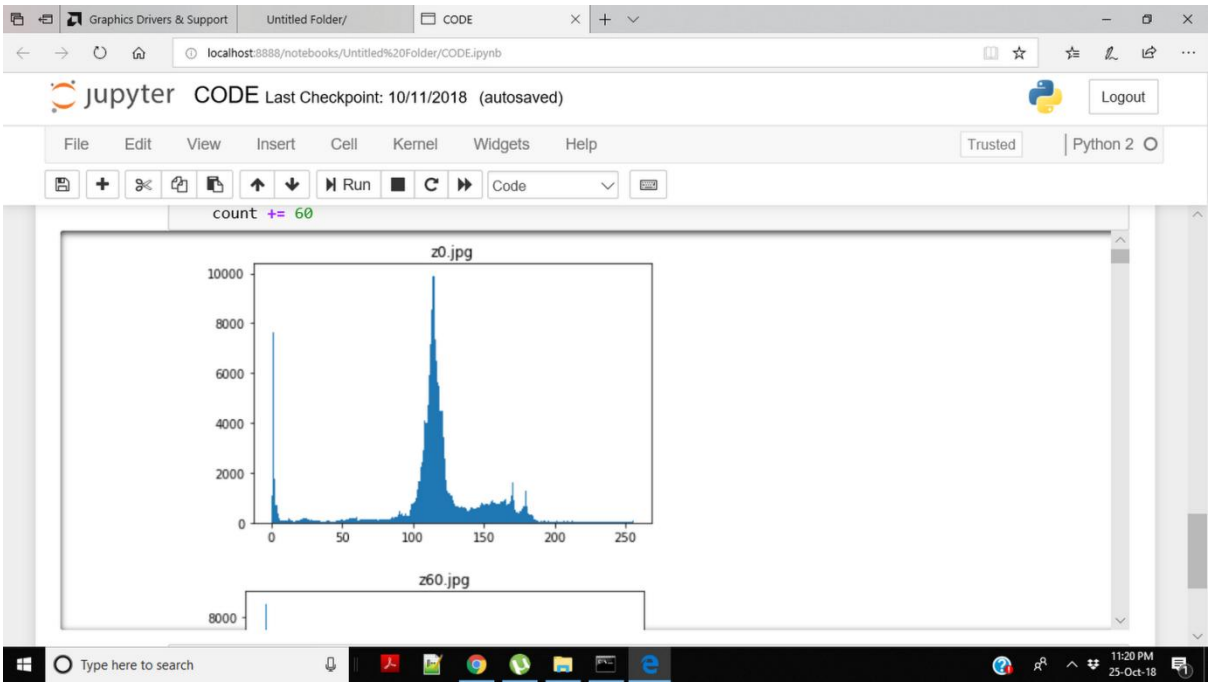
6.1 Experimented Results

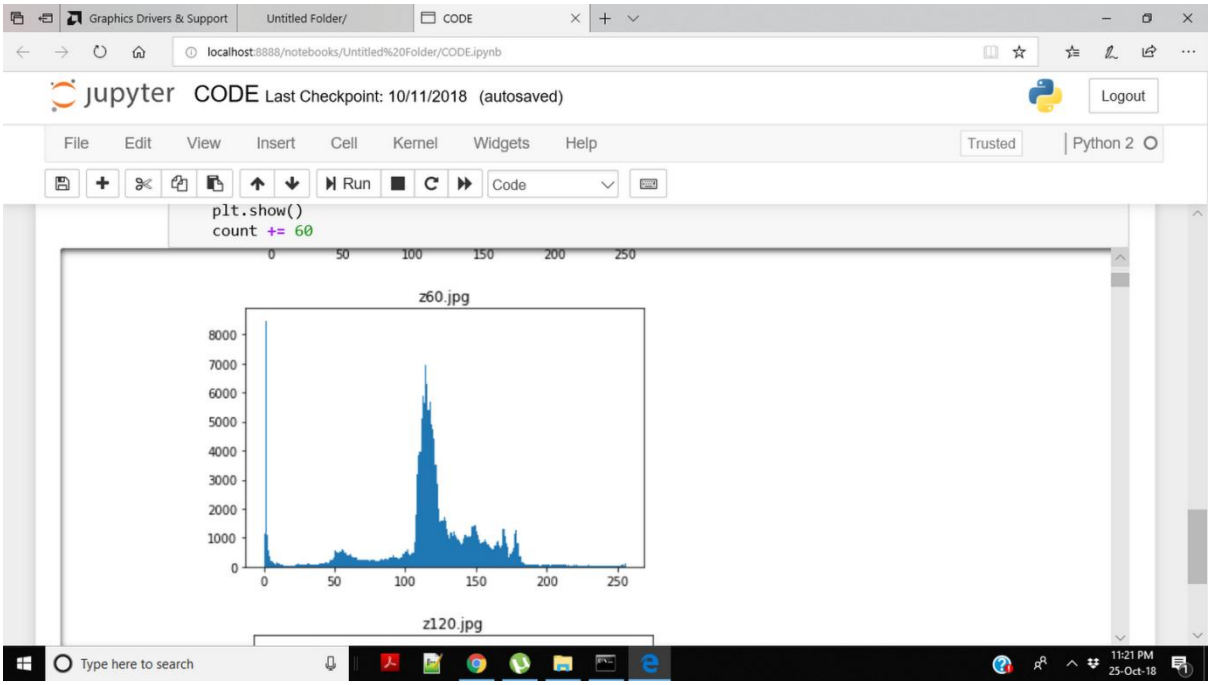
- Extensive testing was performed on a number of videos.
- Each video depending upon its frame rate, video quality and length differed in total time taken.
- The algorithm applied by us ensures no video will take an exceedingly long time. Each video will take a reasonable time to generate summary.

Sr. No.	Video Size (in MB)	Video length (in Minutes)	Summary Size (in MB)	Summary length (in Minutes)
1	56.3	13.04	2.56	1.51
2	122	12.10	2.50	3.09
3	11	4.05	5.10	2.31
4	139	17.41	6.46	2.48
5	107	12.54	8.1	1.55

Result Screenshots

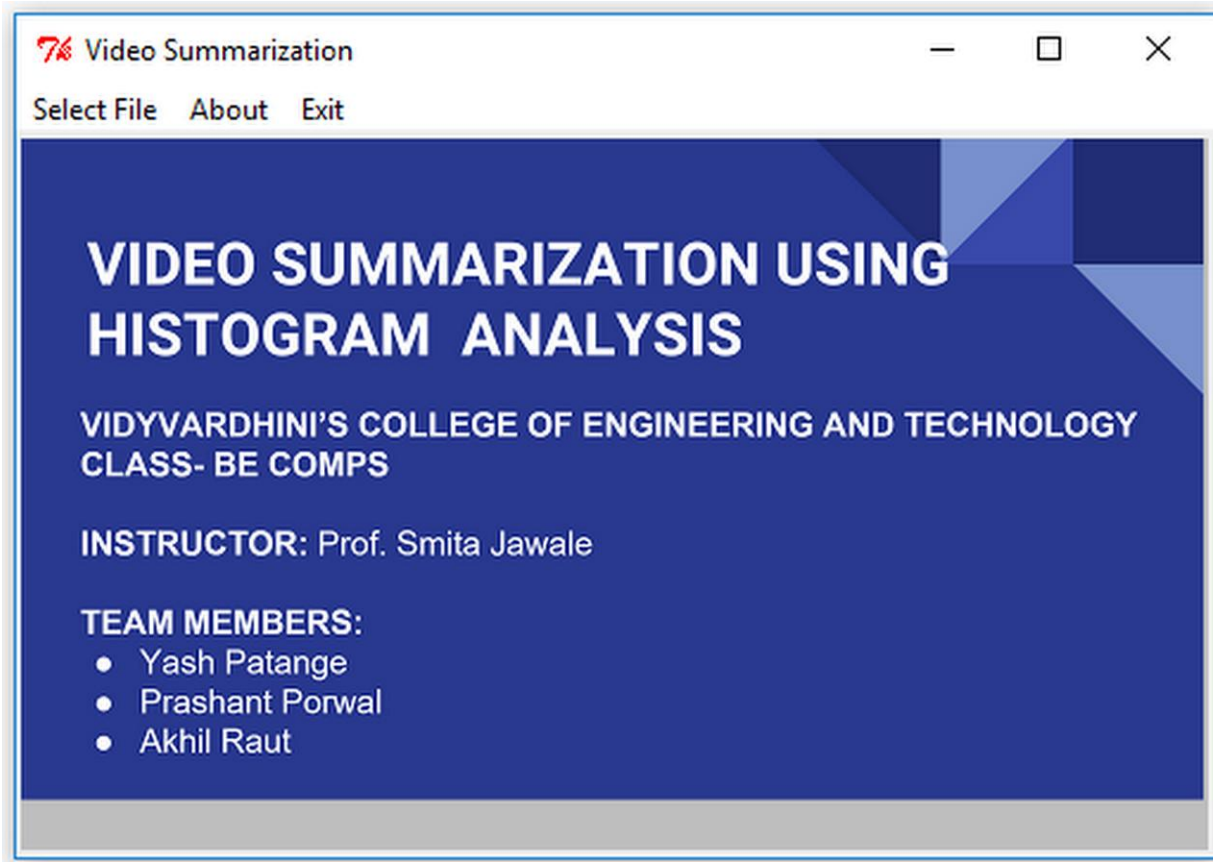
Histograms



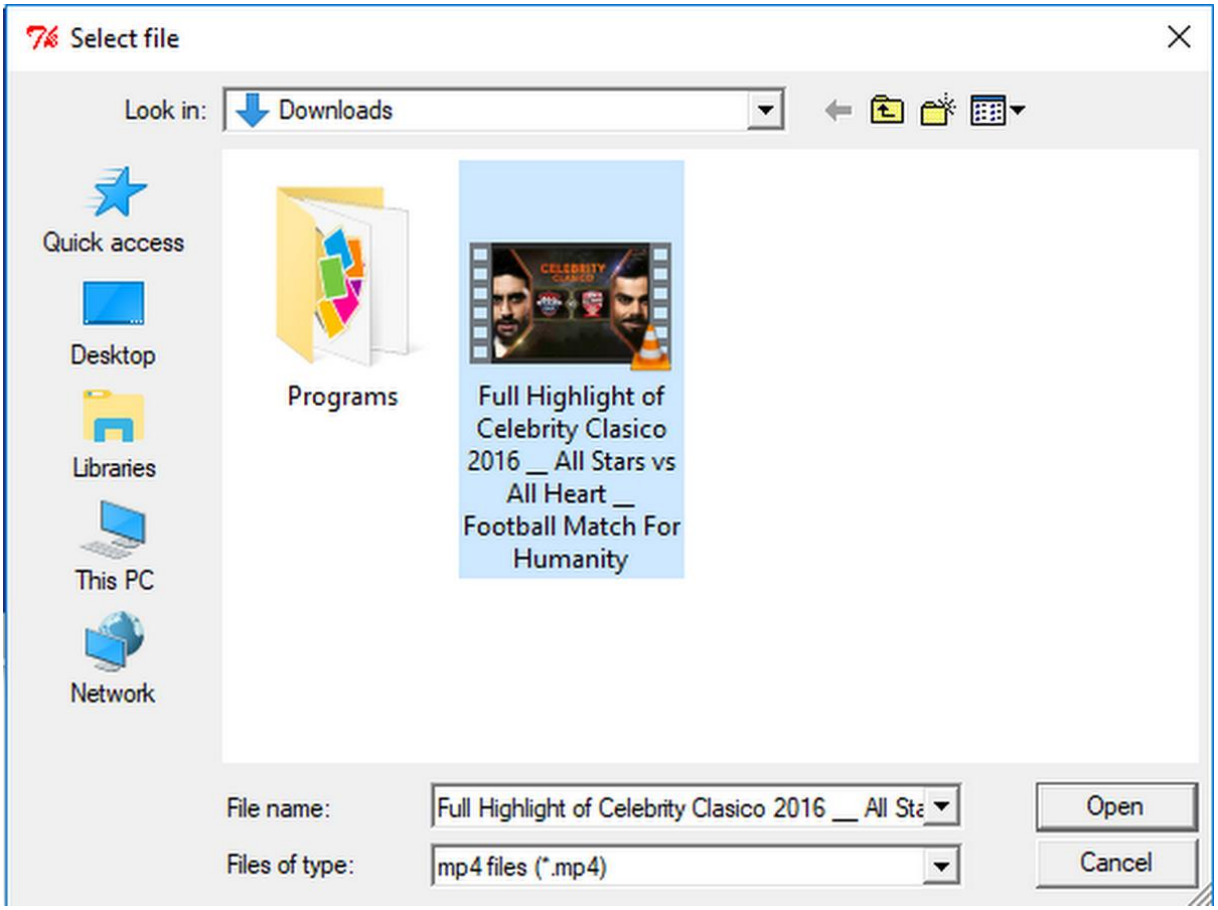


App Screenshots

Home Screen



File Selection Screen



Chapter 7

CONCLUSION

7.1 Conclusion

The methodology proposed here will certainly help laymen who are not computer savvy enough to watch highlight streams of their own choice of video. It can also be used to obtain key scenes or sequences of any type of video.

7.2 Future Scope

Our project has an incredible amount of scope in the field of video processing. This can be used in video hosting websites to show us what the video is about in its thumbnail.

GoPro cameras are lot popular in the market right now. When people go on an adventure they tend to record the whole trip in this camera. After returning from the journey it gets tedious for people to see the hours and hours of footage and edit it down to their fancy. By using video summarization, they can input the video and the software will do the work. Output is a highlight reel of their vacation.

People love to watch sports. They will sometimes spend lot of time in watching their favourite teams play the match. Sometimes it may not be possible for people to spare such large amount of time. By applying machine learning and event detection, video summarization can perfectly form the highlight stream of the match for the viewer's pleasure.

Another great use of this project is for security purposes. CCTV cameras record footage 24/7. This is a lot of data to view in case of any mishap. With the help of artificial intelligence and object detection and tracking, data to be viewed can be reduced drastically and help catch any incident easily in the summarized video

Chapter 8

REFERENCES

1. Nada Jasim Habeeb, Rana Saad Mohammed, Muntaha Khudair Abbass, "Surveillance Video Summarization Based on Histogram Differencing and Sum Conditional Variance", International Journal of Computer and Information Engineering Vol:10, No:9, 2016.
2. Gao, Yue and Qiong-Hai Dai, "Shot-based similarity measure for content-based video summarization." In Image Processing, 2008, ICIP 2008. 15th IEEE International Conference on, pp 2512-2515. IEEE, 2008.
3. Lei, Shoashuai, Gang Xie, and Gaowei Yen "A Novel Key-Frame Extraction Approach for both Video Summarization and Video Index". The Scientific World Journal 2014(2014).
4. Parminder Kaur, Dr. Rakesh Kumar "Analysis of Video Summarization Techniques". International Journal for Reasearch in Applied Science & Engineering Technology(IJRASET) Volume 6 Issue 1, January 2018.
5. Sheen C. V, N. K. Narayanan "Key-frame Extraction by Analysis of Histograms of Video Frames using Statistical Methods". 4th International Conference on Eco-friendly Computing and Communication Systems.

Chapter 9

Plagiarism Check Report

International Journal of Emerging Technologies and Innovative Research(JETIR) (An International Open Access Journal) | Impact Factor: 5.87 | UGC and ISSN Approved Approval link

JETIR206475 Review Notification | JETIR (ISSN:2349-5162) | www.jetir.org | editor@jetir.org

Dear Yash S. Patange,

Your manuscript with Registration ID: JETIR206475 has been **Accepted** for publication in the Journal of Emerging Technologies and Innovative Research (www.jetir.org). Your Review Report is as follows:

UGC and ISSN Approved | 5.87 Impact factor Review Results	
Registration ID	JETIR206475
Email ID	yash.frost1997@gmail.com
Paper Title	Video Summarization using Histogram Analysis
Review Status	Accepted
Impact Factor & Licence:	UGC,ISSN,DOI and Creative Common Approved & 5.87 Calculated by Google Scholar
Unique Contents	76 %
Comments	Paper accepted

Phase 2: Pay Publication Fees and Open Access Processing Charges
For Publishing paper at JETIR, Author needs to pay publication charges and Open Access Processing Charges as follows.

Indian Authors (online Payment)Â https://instamojo.com/@jetir	2100.00/- (Click Here for Paying Money) Payment Link 1: https://instamojo.com/@jetir  Payment Link 2:Â â‚¹2100.00/- (Click Here for Paying Money) 
Bank Details for Offline Payment Like NEFT/IMPS/RTGS/UPI	Bank Name: State Bank of India Account Name: Planet Publication (Current Account) Account No: 34375140165 Branch Name: Visnagar IFSC Code: SBIN0060246 MICR Code: 384002122 ----------Â Pay UPI payment through UPI app like BHIM app, google pay, paytm UPI, BANK UPI etc. UPI id : 9898157864@upi UPI QR Code :Â 9898157864@upiÂ