International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

VIDEO Summarization

[1]ProfSangamaesh S K [2]Prof Pavitra M Gadhar[3]Prof S T Halakatti
Department of Computer Science & Engineering
RTE Society's Rural Engineering College Hulkoti, Karnataka 582205

ABSTRACT:- The raw video is an unstructured data stream, physically consisting of a sequence of video shots. A video shot is composed of a number of frames and its visual content can be represented by key-frames. Video summarization defines as a collection of key-frames extracted from a video. In general, content-based video summarization is therefore a two-step process. The first step is partitioning a video into physical shots, called video segmentation or video shot boundary detection. The second step is to find these representative frames. Thus, video can be organized as video, shot, and key-frames hierarchy. Video summarization can provide a simple and effective way to abstract a long video sequence. They can be a generated as storyboards and video abstractions. Key frames can act as the most representatives of video shot for video indexing, browsing, and retrieval. Video summarization is indispensable processing for video management. After video is structural organized hierarchically, thus, video can be stored and transmitted by shots as minimum components and indexed by sequential key-frames, and be reassembled in receive end. When transmission errors happen, only relative shots therefore need to be resent. By using key-frames, in addition, complex video retrieval task is transformed into simple image comparison exercises among the corresponding key-frames. For a user query, the server only needs to compare key- frames and to issue a file I/O operation to retrieve the relative video segment for transmission to the client. Consequently, video summarization can prompt broadband used effectively, the amount of manipulation data-stream reduced, and the time of computation and I/O access saved.

Keywords:- Threshold, Frames, Polynomial, Shot, Scene.

1.INTRODUCTION

Recently, digital video technology is growing at a rapid rate. Due to advancement in technology, it becomes very easy to record huge volume of videos. A huge bulk of digital contents such as news, movies, sports, and documentaries etc. is available. Moreover, the need for surveillance has increased significantly due to increase in the demand of security especially after 9/11. Thousands of video cameras can be found at public places, public transport, banks, airports, etc. resulting in large amount of information which is difficult to process in real time. Furthermore, storage of huge amount of video data is not that easy. It is very important to quickly retrieve and browse huge volume of data efficiently because end user want to get all important aspects of data. Also, the techniques for automatic video content summarization have attracted numerous attentions due to its commercial potential especially for home video applications. A concise video summary, intuitively, should highlight the video content and

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

contain little redundancy while preserving the balance coverage of the original video. A video summary, nevertheless, should be different from video trailers where certain contents are intentionally hidden so as to magnify the attraction of a video.

A video shot is composed of a number of frames and its visual content can be represented by key-frames. Video summarization defines as a collection of key-frames extracted from a video. In general, content-based video summarization is therefore a two-step process. The first step is partitioning a video into physical shots, called video segmentation or video shot boundary detection. The second step is to find these representative frames. Thus, video can be organized as video, shot, and key-frames hierarchy. Video summarization can provide a simple and effective way to abstract a long video sequence. They can be a generated as storyboards and video abstractions. Key frames can act as the most representatives of video shot for video indexing, browsing, and retrieval. Video summarization is indispensable processing for video management. After video is structural organized hierarchically, thus, video can be stored and transmitted by shots as minimum components and indexed by sequential key-frames, and be reassembled in receive end. When transmission errors happen, only relative shots therefore need to be resent. By using key-frames, in addition, complex video retrieval task is transformed into simple image comparison exercises among the corresponding key-frames. For a user query, the server only needs to compare key- frames and to issue a file I/O operation to retrieve the relative video segment for transmission to the client. Consequently, video summarization can prompt broadband used effectively, the amount of manipulation data-stream reduced, and the time of computation and I/O access saved.

Objective of this project is to shorten or summarize the given video in real time

To study and develop Python programming language with concept of machine learning and thinker file dialogue.

- To develop a video summarization system using Jupiter notebook in Anaconda navigator platform.
- To make an application that shortens some specific functionalities of videos using machine learning
- To analysis the video and result obtained is used in many application. Thus, saves the time of human.
- A Multiple Visual Models Based Perceptive Analysis Framework for Multilevel Video Summarization: In this paper, we propose a generic framework to human perception analysis in video understanding based on multiple visual cues. Video features that prominently influence human perception, such as motion, contrast, special scenes, and statistical rhythm, are first extracted and modeled. A perception curve that corresponds to human perception change is then constructed from these individual models using linear or priority based fusion approach. As an important application of the perceptive analysis framework, a feasible scheme for video summarization is implemented in order to demonstrate the validity, robustness, and generality of the proposed framework. The frames that correspond to the peak points in these individual models and the fusion curve are extracted as multilevel summarizations that include video keywords, key frames, and dynamic segments. The subjective evaluations from a supplementary volunteer study on video summarizations indicate that the analysis framework is effective and offer a promising approach to semantic video management, access, and understanding. As an

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

important application of the proposed perceptive analysis framework, we have presented a feasible solution for multilevel video summarization. According to the four perceptive models and the fusion scheme applied here, we obtained a set of keywords, static key frames, and dynamic segments that accurately represent the original video contents according to the viewpoint of human perception. The experimental results indicate a very promising performance for this proposed summarization method.

II EXPERIMENTAL DESCRIPTION

Design a real time system to summarizes a surveillance video to generate a shortened video by preserving the maximum extent of salient actives of input video for a user specified time. In this paper, we propose to solve the problem of video summarization by using the temporal relationship between the frames of the video as the most relevant features for video summarization. The detailed visualization of the system is shown in figure below.

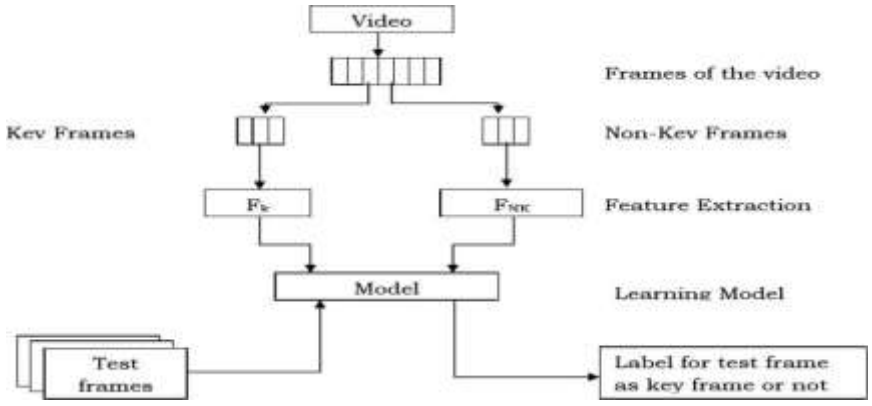

Figure 1: Proposed System for Video Summarization

We extract the optical flow features for every image and the features are then labed using a single vector. We then supply these vectors to a learning module and decide upon the hypothesis. This hypothesis is now used on a test frame. We extract same feature for a test frame and supply it to the model. We then obtain the decision of the frame being key frame or not. This task is repeated for all the frames and we obtain the faster and better video summarization.

Features of the proposed algorithm

1) It maintains the temporal relationship of the frames.
2) It completes the task at a much faster speed
3) The algorithmic complexity is just polynomial in nature, not exponential
4) The memory usage is also of the polynomial order
5) The results obtained are efficient.

Objectives

Objective of this project is to shorten or summarize the given video in real time.

The objectives of the Video summarization are as follows:

1.Convert the video into the frames: To do this, we use the code "video_to_frames.py"

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

2.Rename the frames in a particular format: To do this, we use the code "rename_images_folder.py"

3.Choose the key frames for the summarized video: To do this, we use the code "generate_selected_frames.py"

In this code, the variable "Thresholder" decides the number of images in the summarized video Range for Thresholder (30000 to 150000)

If a smaller number is given for the thresholder, we obtain more frames in the summarized video.

III METHODOLOGY

Design of a system is essential. It is a blue print or a plan for a solution for the system. Here we consider a system to be a set of modules with clearly defined behavior which interact with each other in a defined manner to produce some behavior or services for its environment. Design tells how the system is implemented.

The purpose of the design phase is to plan a solution of the problem specified by the requirements document. This phase is the first step in moving from the problem domain to the solution domain. The design of a system is perhaps the most critical factor affecting the quality of the software; it has a major impact on the later phases, particularly testing and maintenance.

The design process of a software system has two levels. At the first level the focus is on deciding which modules are needed for the system, the specification of these modules and how the modules should be interconnected. This is called as system design or top level design. In the second level, the internal design of the modules or how the specifications of the module can be satisfied is decided. This design level is often called detailed design or logic design.

Once the design is complete, most of the major decisions about the system have been made. However, many of the details about implementing the designs, which often depend on the programming language chosen, are not specified during -design. The goal of the implementation phase is to translate the design of the system into code in a given programming language. For a given design, the aim in this phase is to implement the design in the best possible manner.

The implementation phase affects both testing and maintenance profoundly. Well-written code can reduce the testing and maintenance effort. Because the testing and maintenance costs of the software are much higher than the implementation cost, the goal of coding should be to reduce the testing and maintenance effort.

Typically, video summarization can be divided into two basic modules.

1) Feature extraction

2)Key-frame extraction

1)Feature extraction**:**

In this module, we extract features from each of the frames in the video. The features extracted determine the quality of the summarization. We generally extract different shape based and content based features for every frame.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

2)Key-frame extraction**:**

In this module, we propose to decide if a particular frame in a video is fit to be a key-frame or not. This is generally done by setting up a hypothesis that measures the difference between the features of the present frame and the previous frame.

Anatomy of a Video:

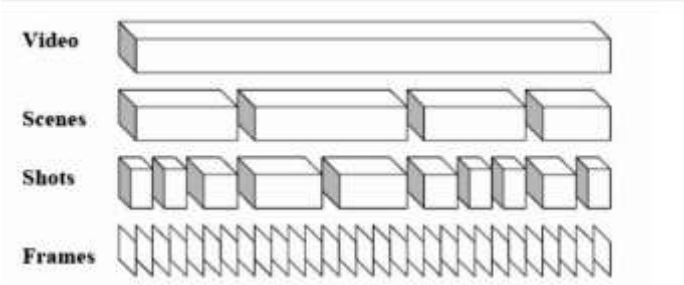

FIGURE2 : Anatomy of a video

- **frame**: a single still image from a video • 24 to 30frames/second
- **shot**: sequence of frames recorded in a single camera operation
- **scene:** collection of shots forming a semantic unity conceptually, a single time and place

There are broadly two approaches towards video summarization: static and dynamic. Static summarization techniques try to find the important frames (images) from different parts of thevideo and splice them together in a kind of story-board. Dynamic summarization techniques divide the video into small video segments/chunks and try to select and combine the important segments/chunks to create a fixed duration summary.

We chose the static approach for reasons of efficiency and utility. We had data indicating that over 80% of the viewers hovered on the thumbnail for less than 10 seconds (i.e. users don't have the patience to watch long previews). We therefore thought, it would be useful to provide a set of four diverse thumbnails that could summarize the video at a single glance. There were UX constraints that kept us from adding too many thumbnails. In this way, our problem became selecting the most relevant thumbnail (hereinafter referred to as 'primary thumbnail') and selecting the four-thumbnail set to summarize a video.

Step One: Selecting the primary thumbnail

Here's how we created a machine-learning pipeline for selecting the primary thumbnail of any video. First and foremost, you need labeled data, and lots of it. To teach our machines some examples of good and bad thumbnails, we randomly sample 30 frames (frame = still image) from the video and show it to our judges. The judges evaluate these frames using a subjective evaluation that considers attributes such as image quality, representativeness, attractiveness, etc. and assign each frame a label based on its quality as Good, Neutral, Bad Point to note – our training data is not query specific, i.e. the judges are evaluating the thumbnail

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

in isolation, and not in the context of the query. This training data, along with a host of features from these images (more on that in a bit) are used to train a boosted trees regression model that tries to predict the label on an unseen frame based on its features. The boosted trees model outputs a score between 0 and 1 that helps us decide the best frame that can be used as a primary thumbnail for the video. What were the features that turned out to be useful in selecting a good thumbnail? As it turned out, core image quality features turned out to be very useful (i.e. features like the level of contrast, the blurriness, the level of noise, etc.). We also used sophisticated features powered by face-detection (# of faces detected, face size and position relative to the frame, etc). Also used were motion detection features and frame difference/frame similarity features. Visually similar and temporally co-located frames are grouped together into video sequences called scenes, and the scene length of the corresponding frame is also used as a feature – this turns out to be helpful indeciding whether the selected thumbnail is a good one. Finally, we also use deep neural networks (DNN) to train high-dimensional image vectors on the image quality labels and these vectors are used to capture the quality of the frame layout (factors like the zoom level [absence of extreme close ups and extreme zoom outs etc.]). The frame with highest predicted frame score is selected as the primary thumbnail to be shown to the user.

Here is a visual schematic:

Step Two: Selecting the remaining thumbnails for the video summary

The next step is to create a four-thumbnail set that provides a good representative summary of the video. A key requirement is comprehensiveness and it brings in many technical challenges. For instance, we could have simply taken the four frames with the highest scores from previous step and created a summary. But that won't work in most cases because there's a high chance that the four top-scored frames are from the exact same scene, and they don't do a good job of summarizing the whole video. There are other problems too - from a computational cost point of view, it is impractical to evaluate all possible sets of four-frame candidates. Thirdly, it's hard to collect training data from users about the four frames that best summarize a video, because, it is hard for users to select the 4 best frames from a video having thousands of frames. Here's how we handle each of these problems.

To deal with the comprehensiveness, we introduce a similarity factor in the objective function. The new objective function for the expanded thumbnail set not only tries to maximize the total image quality score, but also adds an additional tuning parameter for similarity. The weight for this parameter is trained from user's labeled data (more on that below). The similarity factor currently has a negative weight (i.e. a set of 4 high quality frames in which the frames are mutually diverse, will generally be considered a better summary than a corresponding set where the frames are similar).

We deal with computational complexity by formulating the problem as a **greedy optimizationproblem**. As stated before,it'snot possible to evaluate every possible

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

select the 4 best frames from a video having thousands of frames. Here's how we handle each of these problems.

To deal with the comprehensiveness, we introduce a similarity factor in the objective function. The new objective function for the expanded thumbnail set not only tries to maximize the total image quality score, but also adds an additional tuning parameter for similarity. The weight for this parameter is trained from user's labelled data (more on that below). The similarity factor currently has a negative weight (i.e. a set of 4 high quality frames in which the frames are mutually diverse, will generally be considered a better summary than a corresponding set where the frames are similar).

$$f(I;w)(x,y,t) = f(I;w)(x + \Delta x, y + \Delta y, t + \Delta t), \qquad (2)$$

We deal with computational complexity by formulating the problem as a **greedy optimizationproblem**. As stated before,it'snot possible to evaluate every possible combination of 4-framesummaries. Moreover, the best combination of 4 frames need not contain the primary thumbnail (it's possible that the best combination excludes the primary thumbnail). But since we've already taken great pains to select the primary thumbnail, it can greatly simplify our task if we use this as a starting point to select just three more thumbnails that help maximize the total score. That's greedy optimization. Here's how we generate training data for learning the weights for similarity and other features. We show judges a set of 4 frames on LHS and RHS (these frames are randomly selected from the video) and ask them to do a side-by-side judgment (label as **"left better", "right better", or "equal"**). This training data is then used to derive the thumbnail-set model bytraining the new objective function (total image quality score and similarity) for the 4-frame set. As it turned out, based on the training data, the weight for similarity is negative (i.e. in general, more visually diverse frame-sets lead to better summaries). That's how we select the 4-thumbnail set.

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

Optical Flow Guided Feature: It is inspired by the famous brightness constant constraintdefinedby traditional optical flow. It is formulated as follows:

$$I(x,y,t) = I(x + \Delta x, y + \Delta y, t + \Delta t), \text{-------(1)}$$

where I(x,y,t) denotes the pixel at the location (x,y) of a frame at time t. For frames t and (t + Δt),Δx and Δy is the spatial pixel displacement in x and y axes respectively.It assumes that for any point that moves from(x,y) at frame t to (x+Δx,y +Δy) at frame t+Δt, its brightness keeps unchanged over time. When we apply this constraint at the feature level, we havewhere f is a mapping function for extracting features from the image I. w denotes the parameters in the mapping function. The mapping function f can be any differentiable function. In this paper, we employ trainable CNNs consisted of stacks of convolution, ReLU, and pooling operations. According to the definition of optical flow, we assume that p = (x,y,t) and obtain the equation as follows:

$$\partial f(I;w)(p)\ \partial x\ \Delta x + \partial f(I;w)(p)\ \partial y\ \Delta y + \partial f(I;w)(p)\ \partial t\ \Delta t = 0. \quad (3)$$

By dividing Δt in both sides of Equation 3, we obtain

$$\partial f(I;w)(p)\ \partial x\ vx + \partial f(I;w)(p)\ \partial y\ vy + \partial f(I;w)(p)\ \partial t = 0, \quad (4)$$

where p = (x,y,t), and (vx,vy) denotes the two dimensional velocity of feature point at p. $\partial f(I;w)(p)$ ∂x and $\partial f(I;w)(p)$ ∂y are the spatial gradients of $\partial f(I;w)(p)$ in x and y axes respectively. $\partial f(I;w)$ ∂t is the temporal gradient along time axis. As a special case, when f(I;w)(p) = I(p), then f(I;w)(p)simply represents pixel at p. In this special case, (vx,vy) are called optical flow. Optical flow is obtained by solving an optimization problem with the constraint in Equation 4 for each p [1, 4, 2]. Here in this case, the term $\partial f(I;w)(p)$ ∂t represents the difference between RGB frames. Previous works have shown that the temporal difference between frames is useful in video related tasks , however, there is no theoretical evidence to help explain why this simple idea.

Here, we can find its correlation to spatial features and optical flow. We generalize the representation of optical flow from pixel I(p) to feature f(I;w)(p). In this general case, [vx,vy] are called the feature flow. We can see from Equation4that ~ F(I;w)(p) = [$\partial f(I;w)(p)$ ∂x , $\partial f(I;w)(p)$ $\partial y \partial f(I;w)(p)$ ∂t] is orthogonal to the vector [vx,vy,1] containing feature level optical flow. ~ F(I;w)(p) changes as the feature-level optical flow changes. Therefore, ~ F(I;w)(p) is guided by the feature-level optical flow. We call ~ F(I;w)(p)as Optical Flow guided Feature (OFF).

RESULTS

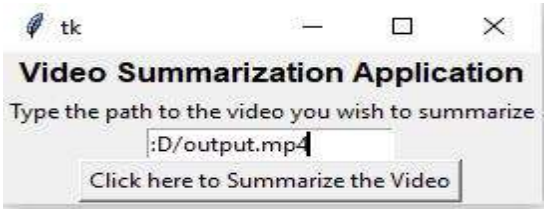

Figure 3

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

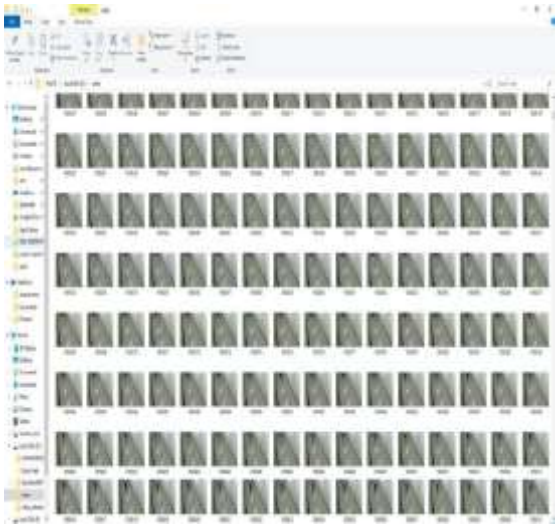

Figure 4

The output of code is divided into two types here is the first type where all the frames in the video are displayed.

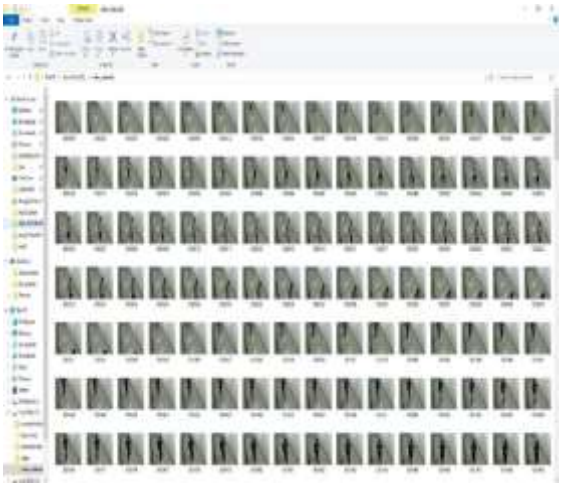

Figure 5 Here is the second type of execution where only the selected frames of video are displayed.

CONCLUSION

The recent advancements in the field of video analytics have driven the need for automatic video summarization. There are many techniques for video summarization, which were studied and categorized. It is observed that not all the summarization techniques fit well in each and every situation. Some of the techniques (Low level feature based) are good for real time applications as they are computationally simple and fast; where as some techniques (High level feature based, User attention model based) are particularly suitable for applications that require precise and accurate results regardless of the time taken for producing the summary (e.g. surveillance

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 6, Issue 10 - October 2018 - Pages 86-95

applications). Each technique has its own merits and demerits but the need for a technique which is independent of the application is realized. Secondly there is lack of standard evaluation techniques, earlier user generated summaries were used to evaluate the automatic generated summary later shorter construction degree, fidelity were proposed and used.

REFERENCES

[1]. H. S. Chang, S. S. Sull, and S. U. Lee, "Efficient video indexing scheme for content-based retrieval," IEEE Trans. Circuits Syst. Video Technol., vol. 9, no. 8, pp. 1269–1279, Dec.1999.

[2] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, Introduction to Algorithms. New York/Cambridge, MA: McGraw-Hill/MIT Press, 1990.

[3].D. DeMenthon, V. Kobla, and D. Doermann, "Video summarization by cur simplification," in Proc. 6th ACM Int. Conf. Multimedia, 1998, pp. 211–218.

[4].Y. H. Gong and X. Liu, "Video summarization using singular value decomposition," in Proc. Int. Conf. Comput. Vis. Pattern Recognit., vol. 2, 2000, pp. 174–180.

[5].A. Hanjalic and H. J. Zhang, "An integrated scheme for automated video abstraction based on unsupervised cluster-validity analysis," IEEE Trans. Circuits Syst. Video Technol., vol. 9, no. 8, pp. 1280–1289, Dec. 1999

[6]]L. Itti and C. Koch, "Computational modeling of visual attention," Nature Rev.Neurosci., vol. 2, no. 3, pp. 194–203, Mar. 2001.

[7].L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Machine Intell., vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[8]. Azra Nasreen, Dr Shobha G, (2013). `Key Frame Extraction using Edge Change Ratio for Shot Segmentation', International Journal of Advanced Research in Computer and Communication Engineering, Vol 2.

[9] Proceedings, published by Springer Berlin Heidelberg, (2008). `Video Summarization: Techniques and Classification ', International Conference, ICCVG 2012, Warsaw,Poland,pp.24-26

[10] The Open-Video Project: http://www.open-video.org

[11] Karim M. Mohamed, Mohamed A. Ismail, and Nagia M. Ghanem (2014) VSCAN: An Enhanced Video Summarization using Density-based Spatial Clustering. Computer and Systems Engineering Department Faculty of Engineering, Alexandria University Alexandria, Egypt.