COMPARATIVE STUDY OF KEY FRAME EXTRACTION ALGORITHMS

PROJECT COMPLETION CERTIFICATES

This is to certify that **Mr Vaibhav Singh** bearing **Registration No. 20120423**, **Mr Tanmay Disoriya** bearing **Registration No. 20120000** and **Mr Suryadeep Chatterjee** bearing **Registration No. 20120407** of **Computer Science & Engineering** Department of Sikkim Manipal Institute of Technology has worked under my supervision and guidance from **01 January 2015 to 01 June 2015** and has successfully completed the project entitled **"Comparative Study of Key Frame Extraction Algorithms"** in partial fulfilment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering.

Mr Sanjoy Ghatak

Assistant Professor I

Computer Science & Engineering Department

Sikkim Manipal Institute of Technology

DECLARATION

We hereby declare that the work recorded in this project report entitled **"Comparative Study of The Key Frame Extraction Algorithms"** in partial fulfilment for the requirements of award of B-Tech in Computer Science & Engineering from Sikkim Manipal Institute of Technology (a constituent college of Sikkim Manipal University) is a faithful and bonafide project work carried out at **SMIT, Majitar** under the supervision and guidance of **Mr Sanjoy Ghatak** Assistant Professor I, Computer Science & Engineering Department, Sikkim Manipal Institute of Technology. The results of this investigation reported in this project have so far not been reported for any other Degree / Diploma or any other Technical forum.

The assistance and help received during the course of the investigation have been duly acknowledged.

Mr Vaibhav Singh (Reg No. 20120423) ……………………………

Mr Tabmay Disoriya (Reg No. 20120042) …………………………

Mr Suryadeep Chatterjee (Reg No. 20120407)…………………….

ACKNOWLEDGEMENT

In performing our project, we had to take the help and guideline of some respected persons, who deserve our greatest gratitude. The completion of our project gives us much pleasure. We would like to show our gratitude Mr Sanjoy Ghatak, Assistant Professor I, Sikkim Manipal Institute of Technology for giving us good guideline throughout numerous consultations. We would also like to expand our deepest gratitude to:

- Mr Sanjoy Ghatak, Assistant Professor, Guide, SMIT.
- Prof. (Dr.) Kalpana Sharma, Computer Science and Engineering, HOD, SMIT.
- Computer Science and Engineering Department, SMIT.

Mr Vaibhav Singh (20120423) ………………….

Mr Tanmay Disoriya (20120042) ……………….

Mr Suryadeep Chatterjee (20120407) ..…………

DOCUMENT CONTROL SHEET

1.	Report Number	CSE/Mini Project/B.Tech/Mr Sanjoy_ Ghatak/6/2015
2.	Title of the Report	Comparison of key frame extraction algorithms.
3.	Type of Report	Technical
4.	Author(s)	Mr Vaibhav Singh Mr Tanmay Disoriya Mr Suryadeep Chatterjee
5.	Organizing Units	SMIT
6.	Language of the Document	English
7.	Abstract	The key frame extraction algorithms aim at reducing the amount of data in video processing
8.	Security Classification	General
9.	Distribution Statement	General

LIST OF CONTENT

0. ABSTRACT--10

1. INTRODUCTION--11
   1.1 GENERAL OVERVIEW OF THE PROBLEM-----------------------------12
   1.2 LITERATURE SURVEY---12
   1.3 PROBLEM DEFINITION---13
   1.4 ANALYSIS OF PROBLEM AND SRS---------------------------------------14
   1.5 SOLUTION STRATEGY--.15
   1.6 PROJECT SCHEDULE--.-19

2. DESIGN--.20
   2.1. ARCHITECTURAL DIAGRAM--.21
   2.2. BLOCK DIAGRAM---.22
   2.3. FLOW CHART--24

3. IMPLEMENTATION DETAILS---25
   3.1. COLOUR HISTOGRAM--26
   3.2. EUCLIDEAN DISTANCE--27
   3.3. MAHALANOBIS DISTANCE--29

4. RESULTS AND DISCUSSIONS-------------------------------------.------31
   4.1. RESULTS--32
   4.2. DISCUSSION---34

5. CONCLUSION---35
   5.1 SUMMARY--36
   5.2 CONCLUSION---37
   5.3 FUTURE SCOPE--37

6. REFERENCES--38

LIST OF FIGURES

Sl. No.	FIGURE No.	FIGURE NAME	PAGE No.
1.	Figure-1	Various Parts of A Video	15
2.	Figure-2	Gantt chart	19
4.	Figure-3	Block Diagram	23
5.	Figure-4	Flowchart	24
6.	Figure-5	Euclidean Distance	26
7.	Figure-6	Bar Graph	31
8.	Figure-7(a)	Mahalanobis Formula Output	33
9.	Figure-7(b)	Mahalanobis Formula Output	34
10.	Figure-8	Euclidean Distance Formula Output	34
10.	Figure-9	Histogram Difference Output	35

ABSTRACT

In order to extract valid information from video, process video data efficiently, and reduce the transfer stress of network, more and more attention is being paid to the video processing technology. The amount of data in video processing is significantly reduced by using video segmentation and key-frame extraction. So, these two technologies have gradually become the focus of research. With the features of MPEG compressed video stream, a new method is presented for extracting key frames.

Firstly, an improved histogram matching method is used for video segmentation. Secondly, the key frames are extracted utilizing the features of I-frame, P-frame and B-frame for each sub-lens.

Fidelity and compression ratio are used to measure the validity of the method. Experimental results show that the extracted key frames can summarize the salient content of the video and the method is of good feasibility, high efficiency, and high robustness.

1. INTRODUCTION

1. INTRODUCTION

Video contains huge amount of information at different levels in terms of scenes, shots and frames. To discover knowledge from videos the issue that needs to be addressed is the elimination of redundant information. The Objective is to remove the redundant data which will significantly reduce the amount of information that needs to be processed. So, key frame extraction is the fundamental step in any of the video retrieval applications. It is necessary to discard the frames with repetitive or redundant information during the extraction. In recent years, many algorithms of key frame extraction focused on original video stream have been proposed. This paper provides an extensive survey in this area to bring out the advantages, drawbacks, suitability to an application, and precision of each method for video retrieval systems. Key frame is the frame which can represent the salient content of the shot.

The key frames extracted must summarize the characteristics of the video, all the key frames on the time sequence gives visual summary of the video to the user. There are great redundancies among the frames in the same shot, so only those frames that best reflect the shot contents are selected as key frames to represent the shot. The extracted key frames should contain as much salient content of the shot as possible and avoid as much redundancy as possible.

The features used for key frame extraction can include colours (particularly the colour histogram), edges, shapes, optical flow, MPEG motion descriptors, MPEG discrete cosine coefficient, motion vectors, camera activity etc. The key frames can be extracted utilizing the features of I-frame, P-frame and B-frame for each sub-lens. Fidelity and compression ratio are used to measure the validity of the method. Experimental results show that the by this method extracted key frames can summarize the salient content of the video and is of good feasibility, high efficiency, and high robustness. A framework has been provided to assess the quality of the video against a given reference summary using both subjective and objective measures. A framework for automatic evaluation is needed based on both subjective and objective measures without the reference summary.

1.1. GENERAL OVERVIEW OF THE PROBLEM

Motion is the more salient feature in presenting actions or events in video and, thus, should be the feature to determine key frames. A triangle model of perceived motion energy (PME) is proposed to model motion patterns in video and a scheme to extract key frames based on this model. 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. By focusing the analysis on the compressed video features, paper introduces a real-time algorithm for scene change detection and key frame extraction that generates the frame difference metrics by analysing statistics of the macro-block features extracted from the MPEG compressed stream. The key-frame extraction method is implemented using difference metrics curve simplification by discrete contour evolution algorithm. This approach resulted in a fast and robust algorithm.

Key frames are extracted utilizing the features of I-frame, P-frame and B-frame for each sub-lens. Key frames can also be extracted based on macro-block statistical characteristics of MPEG video stream. The frame difference metrics Azra are generated by analysing statistics of the macro-block features extracted from the MPEG compressed stream. The key-frame extraction method is implemented using difference metrics curve simplification by discrete contour evolution algorithm.

1.2. LITERATURE SURVEY

The MPEG video compression algorithm has two main advantages- (1) Macro block-based motion compensation for the reduction of the temporal redundancy, (2) Transform domain based compression for the reduction of spatial redundancy. In the compression of the video stream, frames can be grouped into sequences called a group of pictures (GOP). The types of frames can be classified into I-frames, P-frames and B-frame.

They are regularly arranged in the video stream and compose the GOPs. Within a GOP, and I-frame is the first frame. I-frames and P frames act as reference frames. I-frames are interceded. The frames are processed with discrete cosine transform (DCT) using 8*8 blocks, and DC coefficients contain the main information. P-frames are inter-frame coded. P frames refer to the preceding I frame or P frame, and are predicatively coded with only forward motion compensation based on macro blocks.

The forward motion vectors for forward motion predication and DCT coefficients of residual error after motion compensation are obtained. B frames are inter-frame coded for forward motion prediction, backward motion prediction and bi-directional motion prediction. Each Macro-block of 16*16 pixels in P-frames and B-frames search for the optimal matching macro block in corresponding reference frames, then reduce predictive error of motion compensation with DCT coding.

Key frames are extracted using the characteristics of I frames, P frames and B frames in the MPEG video stream after shot segmentation. If a scene cut occurs, the first I frame is chosen as a key frame. P frames are coded with forward motion compensation.

1.3. PROBLEM DEFINITION

When a shot transition occurs at a P frame, great change can take place in the P frame corresponding to the previous reference frames. The advantages of this method are: It compensates for the shortcomings of other algorithm and improves the techniques of key frame extraction based on MPEG video stream As shown in figure, the Input video is segmented into shots using the shot change detection techniques and then once the shot is identified, the key frames can be extracted from the candidate frames to represent each shot. All the key frames can be combined together to create a video summary which represent the video as a whole.

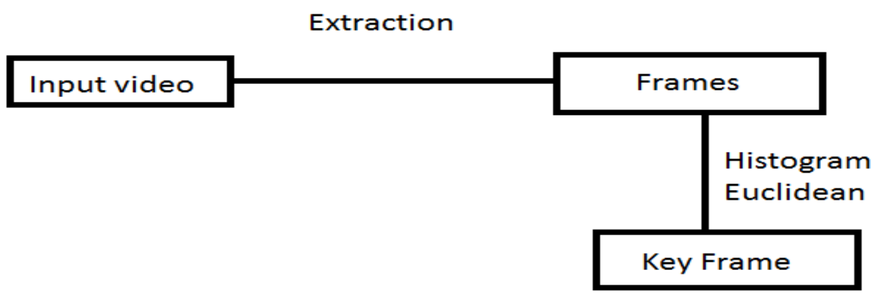

Fig: Design of key frame extraction project

1.4. ANALYSIS OF THE PROBLEM AND THE SRS

This is a digital era. Anything can be done with the help of internet and multimedia. Digital video has been included in every aspect of human life. We use various applications of videos in our daily life. In this digital era, education system is also getting incorporated with digital videos. The main trend of organizing smart class is all about showing digital videos. These videos contain lectures of professors from various colleges. With the help of these digital videos, students from any corner of the world can have access to lectures of any professor. This is achieved with the help of digital videos.

Digital videos have got a huge recognition in this digital era. With the increase use of videos we need to focus on the effective ways of indexing, storing and transferring these videos. Video indexing is the process of providing watchers a way to access and navigate contents easily; similar to book indexing. The basic multimedia information is required for dynamic video indexing and retrieval. There are necessary components for storing, sorting and accessing multimedia contents. They also help in finding the desired components to form a multimedia repository with an ease. Video segmentation is the process of partitioning pieces of information into meaningful elementary parts termed as video segments.

The extraction will help us reduce the network stress while transferring the video by removing unwanted frames from it. Here we use few key frame extraction algorithms and come up with a comparative study of the algorithms mentioning the ways in which the algorithms work.

Each shot is made up of many frames. A frame is the smallest unit of a video. A video is an integrated form of numerous frames. These frames can be further analysed to extract key frames from them.

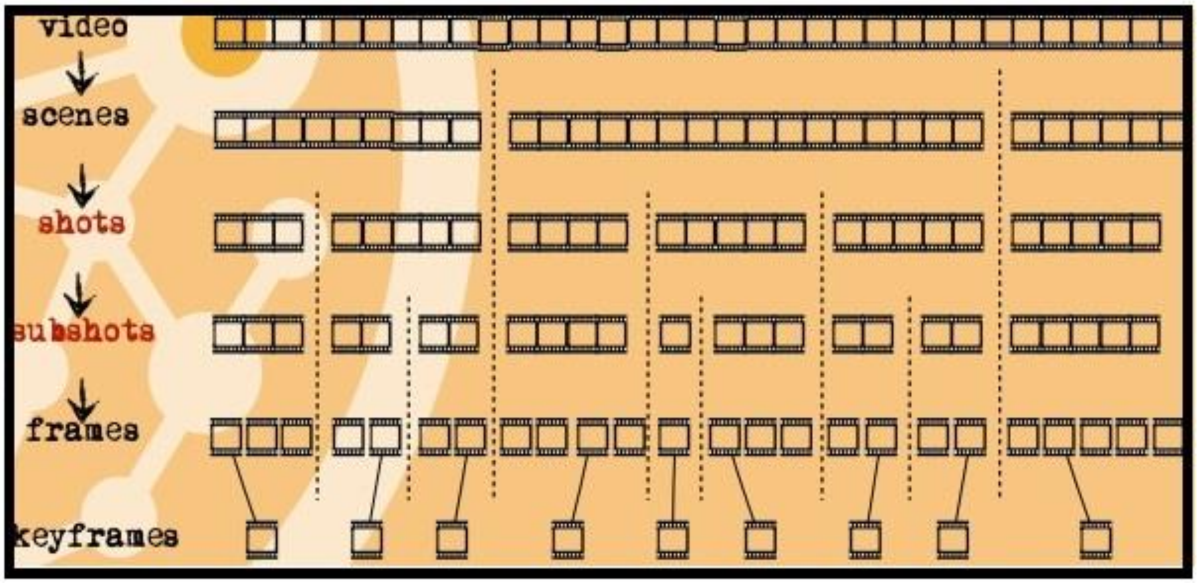

Figure 1: Various parts of a video. (Source: Google images)

1.5. SYSTEM REQUIREMENT AND SPECIFICATON

Hardware Requirements

- Minimum RAM:-2GB
- Hard Disk:-360 GB
- Processor:-Intel Core i3

Software Requirements

- Language: - MATLAB 2015
- Operating System: - Windows 10

Performance Requirements

- The performance the product depends on the hardware.
- Response of the MATLAB software should be fast.
- The dataset video should be clear and it should be a standard dataset.

1.6. PROPOSED SOLUTION STRATEGY

This topic is one of the most trending research topics in today’s time. A lot of researches have already been done in this topic. Key-frame extraction techniques have been used extensively in video retrieval. This is used video browsing. Of most of the works that has been done till now, the basic idea has been measurement and comparison between the consecutive frames. Few of the available techniques have been discussed below:

A. Pairwise-pixel comparison

Among the available techniques, this comparison technique is a straightforward technique. It is probably the simplest way of performing key frame extraction.

In this technique the number of pixels changed between two consecutive frames is counted. In short, pairwise comparison generally is any process of comparing entities in pairs to judge which of each entity is preferred, or has a greater amount of some quantitative property, or whether or not the two entities are identical.

B. Region based technique

Likelihood ratio is a region based technique. Often while capturing a video we come across small camera motions. These small camera motions result into false face detection. With the help of region based techniques we can remove the problem of false face detection caused by small camera motions.

The comparison explained above, compared individual pixels. This technique, instead of comparing individual pixels emphasizes on statistical characteristics. This statistical characteristic is known as likelihood ratio (LR).

LR is the likelihood that a given test result would be expected in a patient with the target disorder compared to the likelihood that the same result would be expected in a patient without the target disorder. LR of the corresponding regions in two successive frames is calculated.

We need to decide a threshold value (th). If the LR value is greater than the threshold value, the region is regarded as being changed. If more number of blocks have changed it is regarded as a shot boundary. It is less sensitive to noise and camera motion than pairwise comparison.

C. <u>Histogram differences</u>

It is the most widely used shot boundary detection technique. Using this technique we find the frame whose histogram value varies from another frame by a considerable amount.

We take this considerable amount because this enables us to detect shot boundaries. While capturing video a common phenomenon is always encountered, this phenomenon is motion. Motion may be of the camera or the object. This technique is less sensitive to motion.

Even in this approach we use the concept of assuming a threshold value. If the change in histogram value is more than the threshold value we find shot boundary. A colour histogram is basically the representation of the distribution of colours in an image. For digital images, a colour histogram represents the number of pixels that have colours in each of a fixed list of colour ranges that span the image's colour space, the set of all possible colours.

A limitation of this approach is that, two images can have exactly same histogram differences while their contents vary extremely. Region based histogram differences are adapted. Frames are divided into various blocks. After disintegrating them into several blocks, grey scale histogram is computed for each one of them.

Grey scale is a range of shades of grey without apparent colour. The darkest possible shade is black, which is the total absence of transmitted or reflected light. The lightest possible shade is white, the total transmission or reflection of light at all visible wavelengths.

In the algorithm explained above a fixed threshold was previously set. We could also have used one global threshold value in the above algorithm but that would limit the scope and consequence of the algorithm.
To overcome this problem we set two threshold values whose value is set adaptively. This helps us enhance the performance. In this method of setting two values of threshold, a higher threshold value is used to get hard cut detection and a lower threshold is used for special effects.

In adaptive threshold method the local thresholds consider various features and adapt the thresholds between various windows. Like in the previous case, if the feature is higher than the corresponding threshold a shot can be declared.

Unlike the difference comparison methods, the edge change ratio (ECR) attempts to compare the actual content of two consecutive frames. It transforms both frames to edge pictures, and compares them using dilatation to compute a probability that the second frame contains the same objects as the first frame. This approach is very sensitive to hard cuts.

D. Distance based approaches

The similarity of frames is also measured by the distance between two frames. Distance between frames mean the distance between corresponding pixels in the frames. We often use masking techniques to find the distance between the frames. There are various masks that are available for image processing which enables us to find the pixel concentration in the consecutive frames under question.

These masks help us to focus only on the pixels mentioned in the mask matrix. The frames are believed to be similar if the distance between the frames is extremely small. If this is the case then we eliminate the frames that have smaller distance among them. It is important to select a measurement method which makes the distance of feature match with the similarity of frames. In this paper we have used Euclidean and Mahalanobis correlation based distance formulae to implement distance based approach of key frame extraction.

1.7. PROJECT SCHEDULE

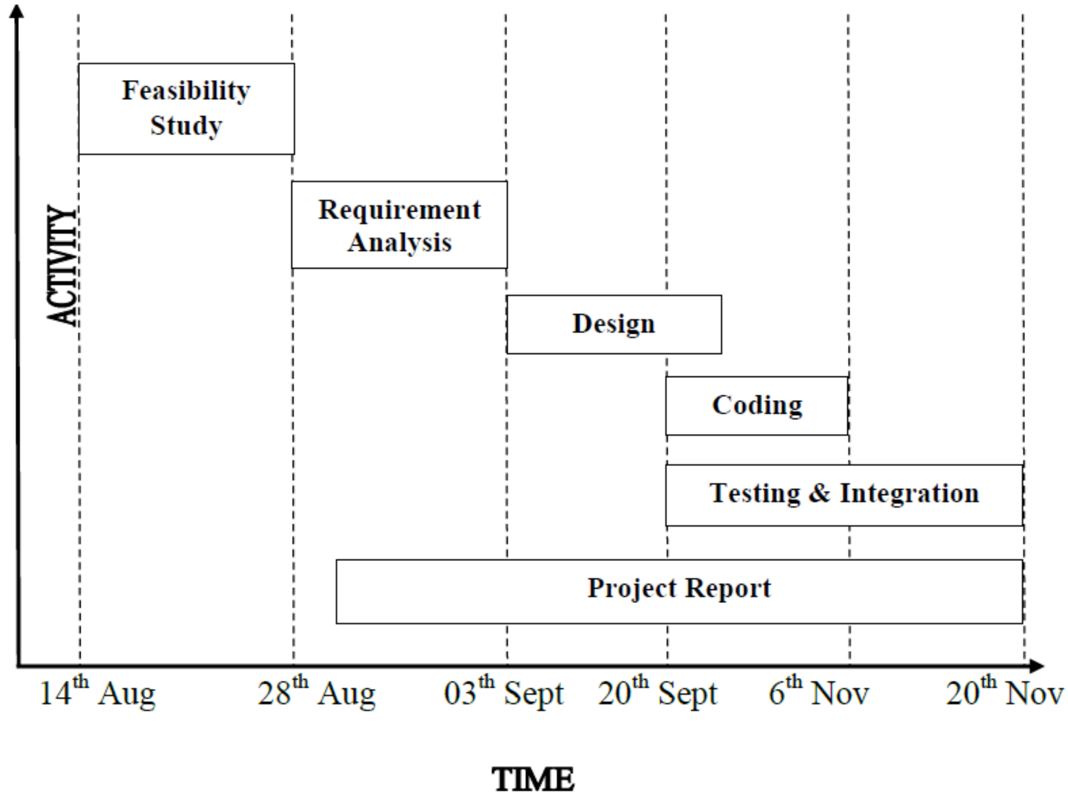

Figure2: Gantt chart

2. DESIGN

2. DESIGN STRATEGIES

Key frame extraction is one of the most important research issues in video information retrieval. To extract information from a video to reduce the transfer rate in the network, we need to process video data efficiently. The amount of data in a video is significantly reduced by video segmentation and key frame extraction. A lot of researches have been going on in the ground off key frame extraction.

There are few approaches which have higher complexity on the basis of processing time. Here, selection of key fame is done on the basis of dissimilarities between two consecutive frames. In this paper we do a comparative study of three key frame extraction algorithms which are effective and have a considerable complexity in terms of processing time. These algorithms analyse the video to find the change in visual frame by computing the frame difference values.

2.1. BLOCK DIAGRAM:

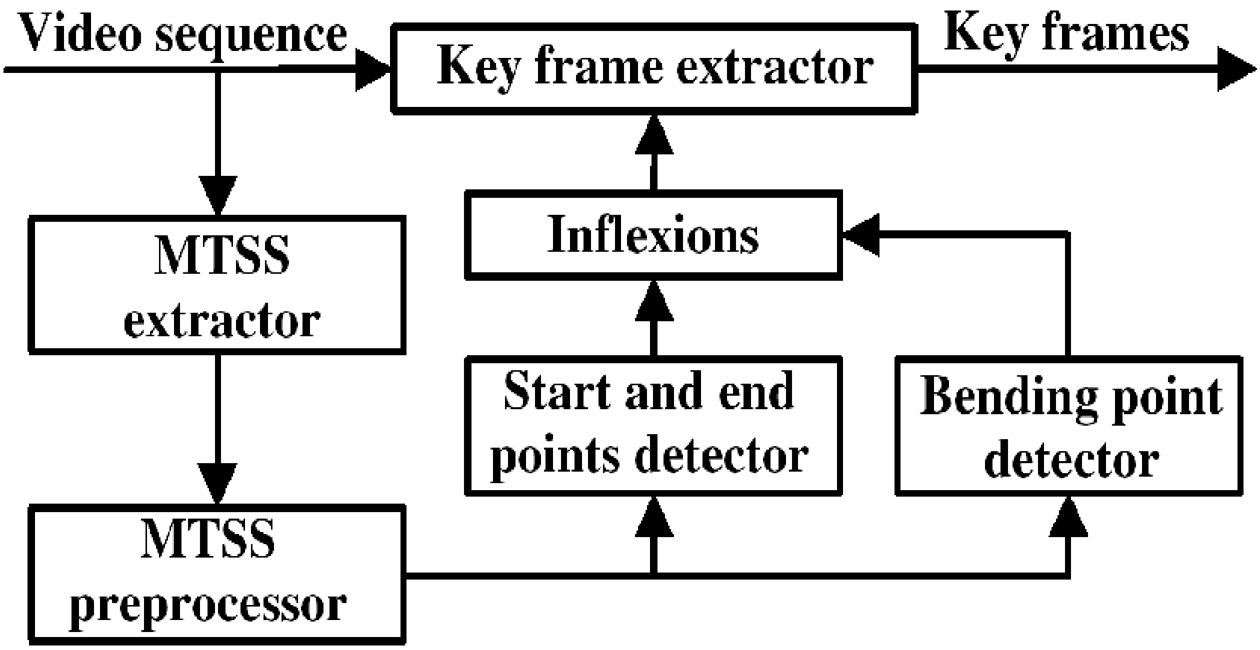

Figure 4: Block Diagram

2.2. FLOWCHART

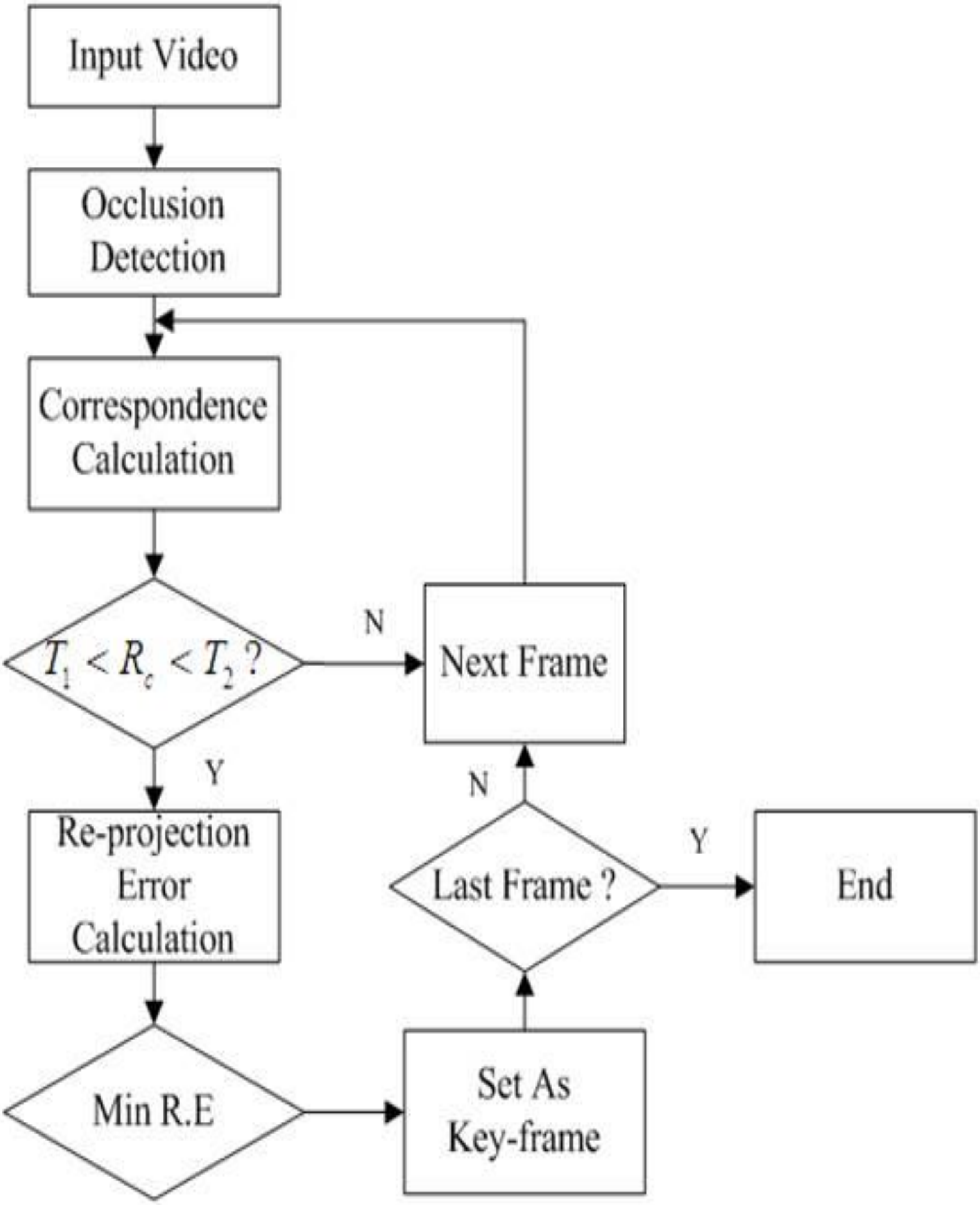

Figure 5: Flowchart

3. IMPLEMENTATION DETAILS

3. IMPLEMENTATION DETAILS

This topic is one of the most trending research topics in today's time. A lot of researches have already been done in this topic. Key-frame extraction techniques have been used extensively in video retrieval. This is used video browsing. Of most of the works that has been done till now, the basic idea has been measurement and comparison between the consecutive frames. Shot segmentation is the first step of the key frame extraction, which mainly refers to detecting the transition between successive shots. The domain of video shot segmentation falls into two categories: uncompressed and compressed.

The detection methods can be broadly classified into abrupt transition detection and gradual transition detection. Nowadays, the common used methods of shot transition detection are as follows: pixel-based comparison, template matching and histogram-based. Since key frame extraction plays an important role in video retrieval and video indexing, a lot of research has
been done on the techniques. The widely used key frame extraction techniques are as follows:

(1) Key frame extraction based on shot activity. Gresleand Huang [7] computed the intra and reference histograms and then computes an activity indicator. Based on the activity curve, the local minima are selected as the key frames.

(2) Key frame extraction based on macro-block statistical characteristics of MPEG video stream. Janko and Ebroul [8] generate the frame difference metrics by analysing statistics of the macro-block features extracted from the MPEG compressed stream. The key-frame extraction method is implemented using difference metrics curve simplification by discrete contour evolution algorithm.

(3) Key frame extraction based on motion analysis. Wolf [9] computed the optical flow for each frame and then used a simple motion metric to evaluate the changes in the optical flow along the sequence. Key frames are then found at places where the metric as a function of time has its local minima. Nowadays, most of the video

are stored in the compressed form of MPEG. The MPEG video compression algorithm has two main advantages: macro block-based motion compensation for the reduction of the temporal redundancy and transform domain based compression for the reduction of spatial redundancy [10]. In the compression of the video stream, frames can be grouped into sequences called a group of pictures (GOP).The types of frames can be classified into I-frames, P-frames and B-frame. They are regularly arranged in the video stream and compose the GOPs.

3.1. COLOUR HISTOGRAM

Using this technique we find the frame whose histogram value varies from another frame by a considerable amount. While capturing video a common phenomenon is always encountered, this phenomenon is motion. Motion may be of the camera or the object. This technique is less sensitive to motion.

A colour histogram is basically the representation of the distribution of colours in an image. Region based histogram differences are adapted. Frames are divided into various blocks. After disintegrating them into several blocks, grey scale histogram is computed for each one of them. The colour histogram is most commonly used frame extraction technique.
Colour is one of the most important features to describe an image. Colour histograms are robust in case of small camera motion and needless to say, they are easy to compute. In histogram based approaches, two frames with unchanging background and unchanging objects are believed to have little difference in their histograms.

This is the list of steps that we follow in order to implement Histogram method:

1. At first, we need to convert the RGB colour mode to HSV colour model.

2. The RGB colour model is an additive colour model in which red, green, and blue light are added together in various ways to reproduce a broad array of colours.

3. The name of the model comes from the initials of the three additive primary colours, red, green, and blue.

4. The HSV, or HSB, model describes colours in terms of hue, saturation, and value (brightness).

5. Saturation corresponds directly to the concept of tint in the Colour Basics section, except that full saturation produces no tint, while zero saturation produces white, a shade of grey, or black.

6. Find colour histogram for each frame.

7. Normalize the colour histogram of each frame.

8. The distance between two consecutive frames is calculated.

9. Find the mean value of HSV colour histograms and this is taken as dissimilarity between frames.

10. If the dissimilarity is larger than a predefined threshold, we consider the two as cut.

11. After going through the whole video, the hard cuts can be detected, the video can be broken into shots and the key frames are extracted.

3.2. EUCLIDEAN DISTANCE

This is a form of distance based approach; this formula is a simple form of distance based approach. Distance between frames mean the distance between corresponding pixels in the frames. In mathematics the Euclidean distance or Euclidean metric is the "ordinary" (i.e. straight-line) distance between two points in Euclidean space. With this distance, Euclidean space becomes a metric space.

The associated norm is called the Euclidean norm. Older literature refers to the metric as Pythagorean metric. A central problem in image recognition and computer vision is determining the distance between images. Considerable efforts have been made to define image distances that provide intuitively reasonable results.

Among others, two representative measures are the tangent distance and the generalized Hausdorff distance. Tangent distance is locally invariant with respect to some chosen transformations, and has been widely used in handwritten digit recognition.

The generalized Hausdorff distance is not only robust to noise but also allows portions of one image to be compared with another, and has become a standard tool for comparing shapes. Among all the image metrics, Euclidean distance is the most commonly used due to its simplicity.

Let x, y be two M by N images, x = (x1, x 2, • • •, x MN), y = (y1, y2, • • •, y MN), where xkN+1, ykN+1 are the grey levels at location (k, l). The Euclidean distance dE(x, y) is given by

$$dE\ 2(x, y) = \sum (xk - yk)2$$

The value of k varies from 1 to MN.

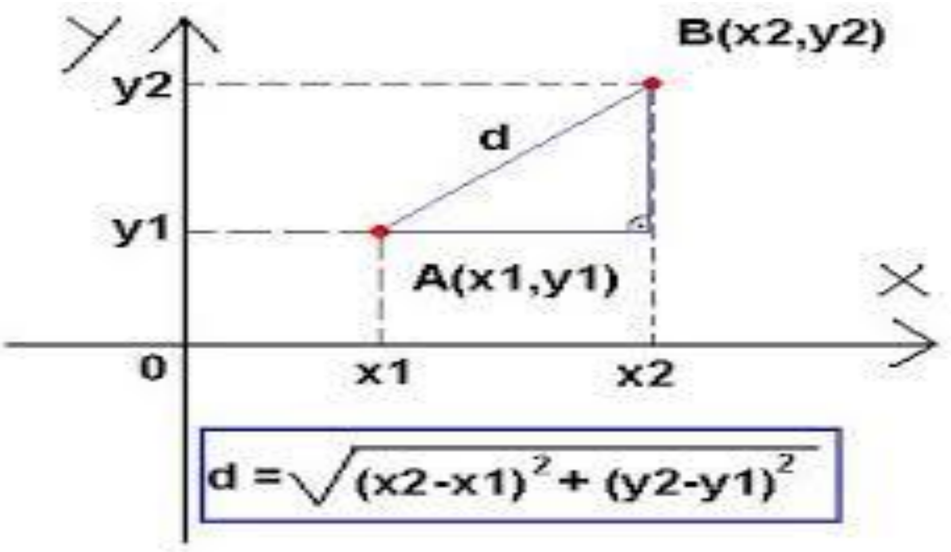

Figure 2: Euclidean distance.(*Source*: Google images)

3.2. MAHALANOBIS FORMULA

The Mahalanobis distance is a measure of the distance between a point P and a distribution D, introduced by P. C. Mahalanobis in 1936. It is a multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D.

Consider the problem of estimating the probability that a test point in N-dimensional Euclideanspace belongs to a set, where we are given sample points that definitely belong to that set. Our first step would be to find the average or centre of mass of the sample points. Intuitively, the closer the point in question is to this centre of mass, the more likely it is to belong to the set.

However, we also need to know if the set is spread out over a large range or a small range, so that we can decide whether a given distance from the centre is noteworthy or not. The simplistic approach is to estimate the standard deviation of the distances of the sample points from the centre of mass. If the distance between the test point and the centre of mass is less than one standard deviation, then we might conclude that it is highly probable that the test point belongs to the set. The further away it is, the more likely that the test point should not be classified as belonging to the set.

The Mahalanobis distance of an observation $\underline{x} = (x_1, x_2, x_3, \ldots, x_N)^T$ from a set of observations with mean

$$\underline{\mu} = (\mu_1, \mu_2, \mu_3, \ldots, \mu_N)^T$$

Covariance matrix *S* is defined as:

$$D_M(\underline{x}) = \sqrt{(\underline{x} - \underline{\mu})^T S^{-1} (\underline{x} - \underline{\mu})}.$$

If the covariance matrix is the identity matrix, the Mahalanobis distance reduces to the Euclidean distance. If the covariance matrix is diagonal, then the resulting distance measure is called a normalized Euclidean distance:

$$d(\underline{x}, \underline{y}) = \sqrt{\sum_{i=1}^{N} \frac{(x_i - y_i)^2}{s_i}},$$

4. RESULTS AND DISCUSSIONS

4.1. RESULTS

To get an overall idea about the algorithms they are implemented on a set of video samples. In a table the outcome of all the algorithms on the respective video sets are shown.

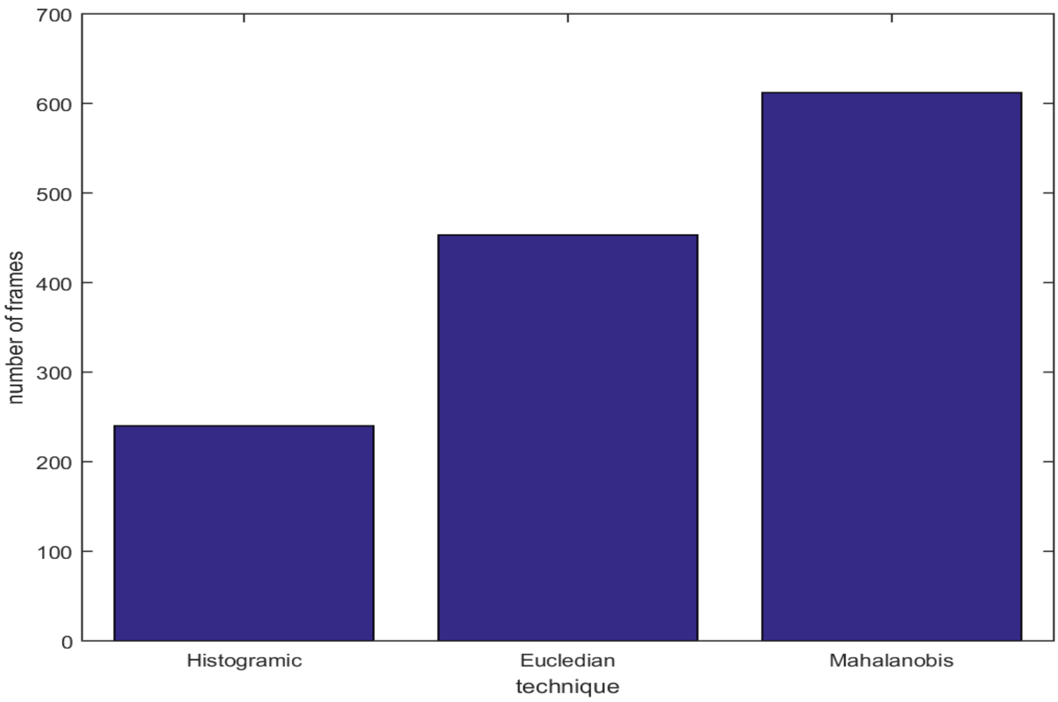

Figure 6: Bar Graph

In the above graph the number of frames after extraction is shown with respect to the algorithms.

We have applied the algorithms in standard database videos with more precession and less noise. The results we got are depicted in the graph.

X-axis: Algorithms

1 denotes COLOUR HISTOGRAM.
2 denote EUCLIDEAN DISTANCE.
3 denote MAHALANOBIS DISTANCE.

Y-axis: number of frames extracted.

4.2. DISCCUSSION

The extraction of the key-frames is done in a totally automatic way without requiring that the user specifying the number of key frames to be extracted as a parameter. Also it is flexible enough to extract the variable number of key frames. Key frames or Representative frames that are obtained are sufficient enough to represent the original shot with little or no change.

Variable numbers of key frames are generated depending on the size of the input video shot. By the inclusion of three frame frames descriptors and distance formula, the obtained results generated lesser number of key frames yet they are able to reflect the significant properties of the input video frames. The important part is that the results are entirely based on the visual details of the input video.

After converting the video into frames we applied the algorithms one by one on the extracted frames. The algorithms decreased the number of frames by deleting the redundant data from the frames.

We are considering a particular range (index no. 12 to index no. 43) of frames to show the comparison between the algorithms.

1. APPLYING MAHALANOBIS DISTANCE FORMULA:

On applying the Mahalanobis distance formula on the extracted frame the number of frames was affected as follows:

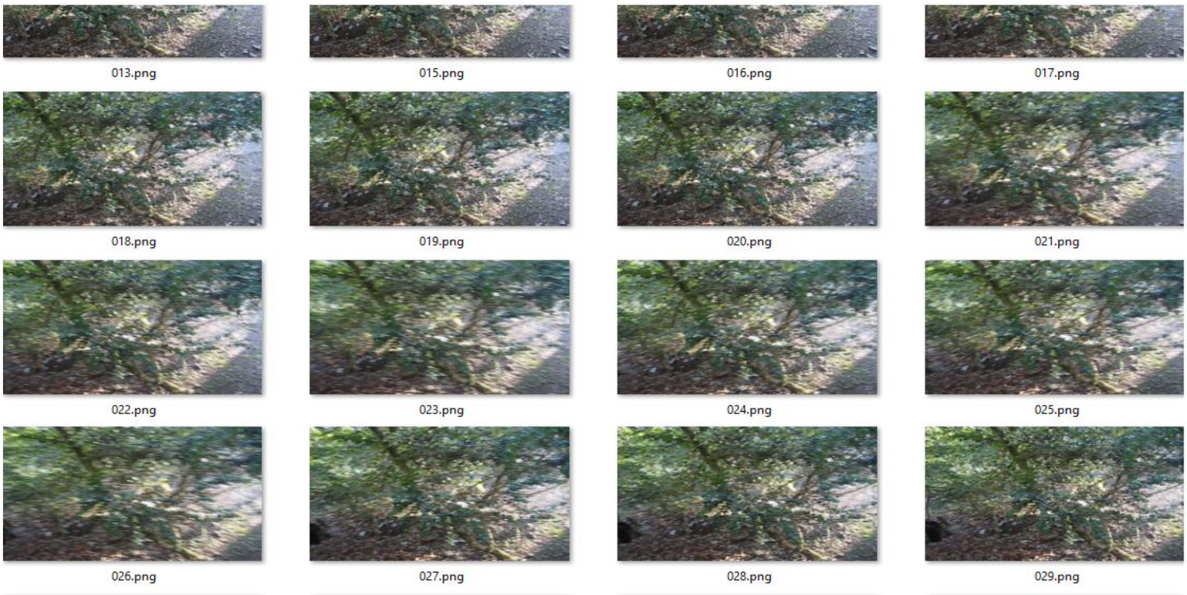

Figure – 7(a): Mahalanobis output

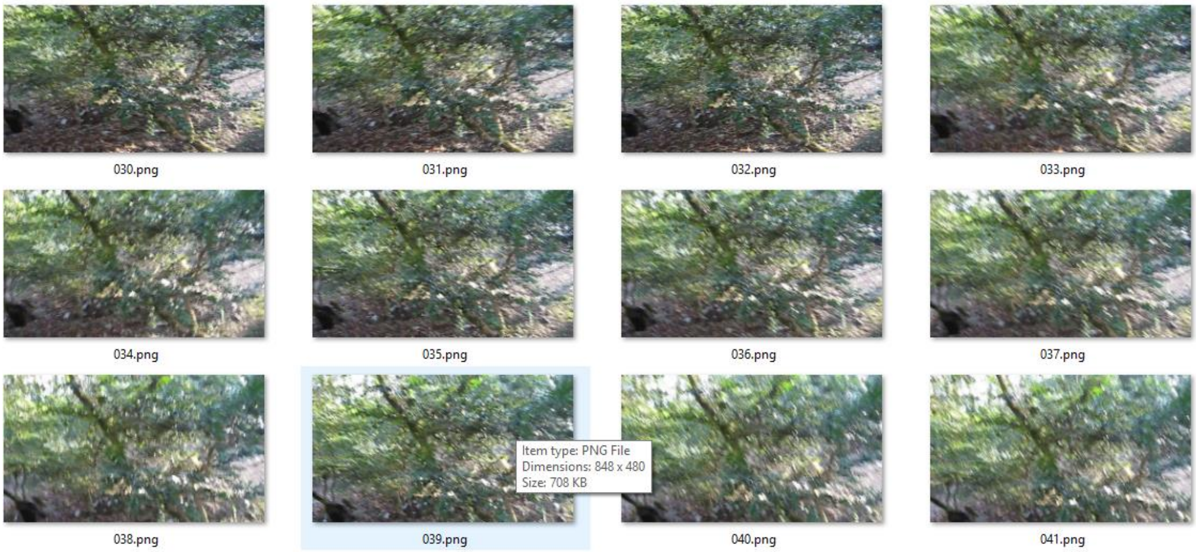

Figure – 7(b): Mahalanobis output

2. APPLYING EUCLIDEAN DISTANCE FORMULA:

On applying the Euclidean distance formula on the extracted frame the number of frames was affected as follows:

Figure 8: Euclidan formula output

3. APPLYING HISTOGRAMNIC DIFFERENCE:

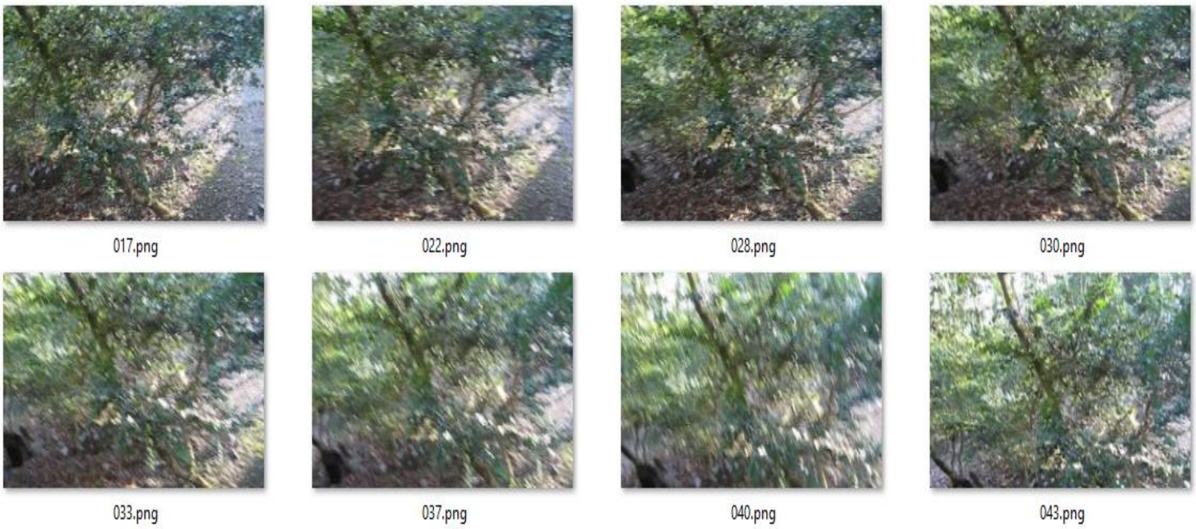

Figure 9: Histogramic difference.

This is how the results vary when we apply the various algorithms on the extracted frame. These pictures show us a good comparison.

5. CONCLUSION

5.1 SUMMARY

With the advent of social networking, sharing of multimedia has gained tremendous amount of importance and is the widely used form of communication worldwide. In the process of discovering knowledge from videos, challenge is to process huge amount of information which is resource intensive. One way to minimize the cost of computations is to reduce the amount of information that undergoes processing. The redundant information can be greatly reduced by extracting key frames that represent the entire video. It is an essential task in any video analysis and indexing applications.

To analyse the comparative study of key frame identification of faces in video with the help of different algorithms like genetic algorithm, adaptive key frame extraction using unsupervised clustering and nearest feature line algorithm for face detection.

5.2 CONCLUSION

In this paper, the basic algorithms used for key frame extraction had been analysed for different surveillance videos. And the detailed discussion is presented in the previous section. Our aim is to extract the key frames that can be used to represent the video as a whole and summarizes the salient content of the video. From this analysis it is evident that the extracted key frames depend on the input video content and the threshold values used in the above methods.

5.3 FUTURE SCOPE

This project has a very useful future aspect. In the era of multimedia advancement every data we deal with are in terms of videos. With the increasing number of videos it has become very important to extract valid information from a video because video contains several data which are not required. This extraction of key frames can be done manually by some person but since it is the era of computer application development we have found a way for it to be done automatically.

This project also helps us do indexing of videos. With the increasing amount of videos we need to be very careful about the video indexing. A proper video indexing may reduce the pain of going through a large bulk of videos and select the required one. Thus, with the help of video indexing we can easily achieve this task.

6. REFERENCES

REFERENCES

[1]T. Kikukawa, S. Kawafuchi, “Development of an automatic summary editing system for the audiovisual resources”, IEEE Trans. on Electronics and Information, J75-A, pp. 204-212, 1992.

[2]T. Liu, H. Zhang, and F. Qi, “A novel video key frame-extraction algorithm based on perceived motion energy model,” IEEE Transactions On Circuits And Systems For Video Technology, vol. 13, no. 10, pp. 1006-1013, 2003.

[3]J. S. Borecsky, L. A. Rowe, “Comparison of video shot boundary detection techniques”, In Proceedings of SPIE, vol. 2670, pp. 170-179, 1996.

[4]N. Hirzalla, “Media processing and retrieval model for multimedia documents”, PhD thesis, Ottawa University, Jan. 1997.

[5]R. Kasturi, R. Jain, Dynamic vision, in Computer Vision: Principles (R. Kasturi and R. Jain, eds.), pp. 469-480, IEEE Computer Society Press, 1991.

[6]Y.S. Avrithis,A.D. Doulamis,N.D. Doulamis, and S.D. Kollias. A stochastic framework for optimal key frame extraction from mpeg video databases. Computer Vision and Image Understanding,75(1/2):3 24,July-August 1999.

[7]Girgensohn A., Boreczky J. Time-Constrained Key frame Selection Technique. Multimedia Tools and Application, 2000; 11:347-358.

[8]Li Zhao, Wei Qi, Stan Z. Li, S.Q.Yang, H.J. Zhang. Key-frame Extraction and Shot Retrieval Using Nearest Feature Line (NFL). Proc. ACM Int. Workshops on Multimedia Information Retrieval, 2000; 217-220.