

VIDEO SEGMENTATION & SUMMARIZATION USING MODIFIED GENETIC ALGORITHM

H S Prashantha

Professor, Department of Electronics & Communication Engineering, Nitte Meenakshi
Institute of Technology, Bangalore, India

ABSTRACT

Video summarization of the segmented video is an essential process for video thumbnails, video surveillance and video downloading. Summarization deals with extracting few frames from each scene and creating a summary video which explains all course of action of full video with in short duration of time. The proposed research work discusses about the segmentation and summarization of the frames. A genetic algorithm (GA) for segmentation and summarization is required to view the highlight of an event by selecting few important frames required. The GA is modified to select only key frames for summarization and the comparison of modified GA is done with the GA.

KEYWORDS

Video segmentation, video summarization, Genetic Algorithm, video streams

1.INTRODUCTION

Segmenting multimedia data streams is a necessary requirement for some of the applications. Segmentation is the process of partitioning a piece of information into elementary parts termed as segments. Video segmentation is used to describe a range of different processes for partitioning the video into meaningful parts of different granularities. Properly summarized video streams can be better organized and reused after segmentation for some applications. The purpose of segmentation is to partition the video sequence into shots based on the contents and key frames are selected. Since the adjacent frames in the video contain similar information and hence distinct frames need to be considered to summarize the mass of material.

There are two types of scene changes such as abrupt scene change (sudden change in the scene, i.e. the next scene starts immediately after completion of the previous scene) and Gradual scene change (change in the scene with a delay). The scene change can be viewed in three ways such as Fade in, Fade out and Dissolve. Fade in is the change of scene from bright or dark to normal color. Fade out is the change of scene from normal color to bright (or) dark. Dissolve is the change in the scene which is the weighted average of both the previous and the next scene.

The classification of the video segmentation algorithms includes pixel comparison, block based comparison, histogram comparison, feature based comparison, clustering based temporal segmentation and model driven video segmentation. The different algorithms for segmentation of the video are studied. Temporal video segmentation for real-time key frame extraction is discussed (Calic, J 1993). A temporal video segmentation method proposed is based on the detection of shot abrupt transition and gradual transition, and then takes into account the conditions of user terminals, which could generate different video summarization for each user (Chen Yinzi 2010). Implementation of image segmentation using GA involves defining a fitness evaluation function, designing a 'Population' (set of chromosomes), defining genetically inspired operators such as crossover and mutation to evolve new population and deciding the termination

of the evolutionary search for the optimal solution (Aravind.I 2002). The genetic segmentation algorithm for video is defined and the evolutionary nature of genetic algorithms offers an advantage by enabling incremental segmentation (Patric Chiu 2000). In the computational process, the improved GA adjusts crossover probability and mutation probability automatically according to the variance between the target and background, thus overcoming the problems of Simple Genetic Algorithm (Lei Hui 2008) (Bir Bhanu 1995). A new approach based on GA is proposed for selection of threshold from the histogram of images (P. Kanungo 2006) (Chi-Chun Lo 2003). Literature documents the ability of GA for video segmentation. GA imitates the principles of biological evolution and it is more suitable method for segmentation and summarization compared to the other traditional methods. The algorithm makes use of global search and optimization method of video segmentation and summarization GA is a search algorithm based on the mechanics of natural selection and natural genetics. The algorithm is probabilistic iterative. GA transforms a set (called a population) of mathematical objects (typically fixed-length binary character strings), each with an associated fitness value, into a new population of offspring objects using the Darwinian principle of natural selection and using operations that are patterned after naturally occurring genetic operations, such as crossover and mutation. The genetic algorithm works by randomly selecting pairs of individual chromosomes to reproduce for the next generation. The selection of chromosome is mainly based on the fitness function value to the other chromosomes in the same segmentation. Usually chromosomes are randomly split and merged with the consequence that some genes of a child come from one parent while others come from the other parents. This mechanism is called crossover.

A simple GA consists of following five steps.

- 1.Start with a randomly generated population of N chromosomes, where N is the size of population and l is the length of chromosome x .
- 2.Calculate the fitness value of function $\varphi(x)$ of each chromosome x in the population.
- 3.Repeat until N off-springs are created
 - 3.1 Probabilistically select a pair of chromosomes from current population using value of fitness function.Produce an offspring y^i
 - 3.2 using crossover and mutation operators, where $i = 1,2,3,\dots$.
- 4.Replace current population with newly created one.
- 5.Go to step 2.

In case of simple GA, the whole population is formed of strings having the same length. These strings contain encoded information.

2.METHODOLOGY USED FOR VIDEO SEGMENTATION

The reason for using GA is its ability to deal with large complex search spaces in situations where only minimum knowledge is available about the objective functions. The corresponding search space in many situations is quite large and there are complex interactions among parameters. Most of the video segmentation algorithms have many parameters that need to be adjusted for effective segmentation. One of the parameter can be considered is the fitness function which indicates the quality of the individual. GA involves defining a fitness evaluation function, designing a population (set of chromosomes) defining genetically inspired operators such as cross over and mutation to evolve new population and deciding the termination of evolutionary search for the optimal solution.

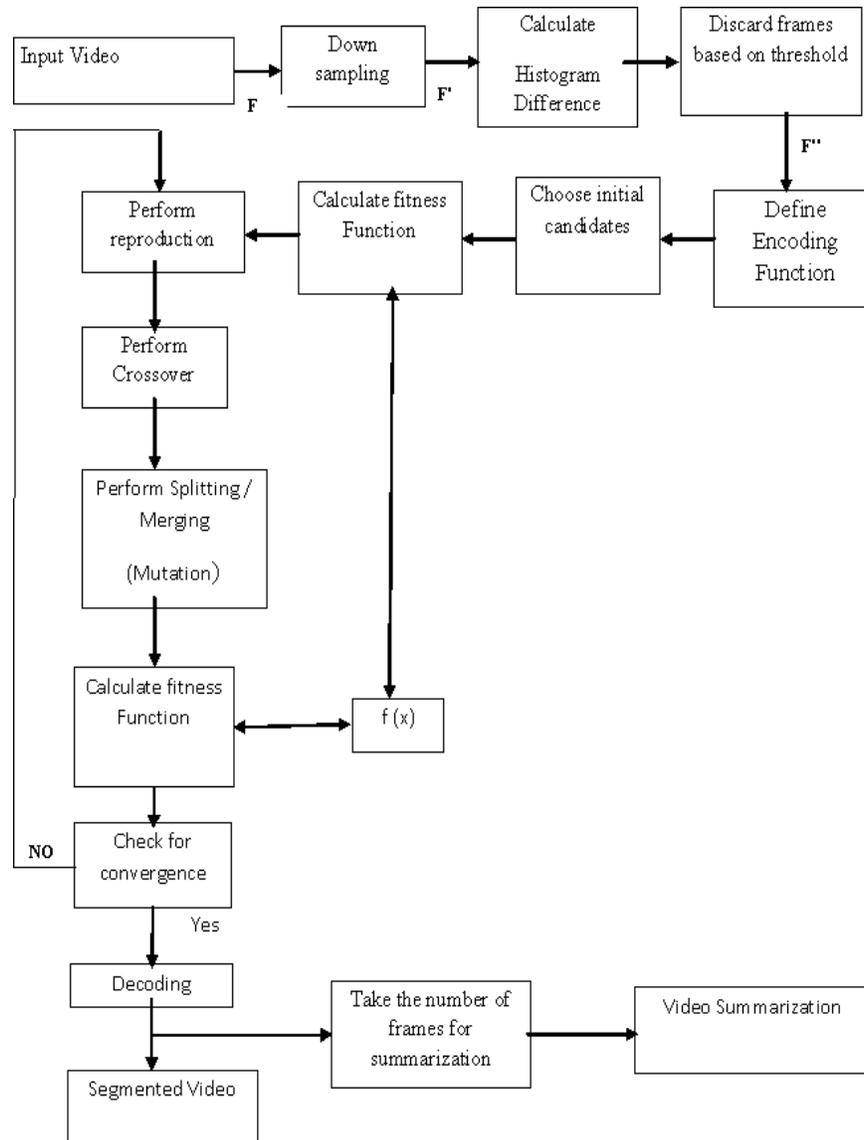


Figure 1: Block diagram of Video Segmentation and Summarization using proposed approach

For the segmentation of video, an efficient and optimal idea is to reduce the number of frames. A video can have thousands of frames, and many adjacent frames are generally similar to provide the illusion of continuity. Since adjacent frames are similar, sub sampling is done at a lower rate to obtain few frames out of many thousands of frames. The selection of the sampling factor depends on the type of scene changes. For gradual scene changes, choose higher value of sampling factor and for abrupt scene changes choose lower value of sampling factor. Obtain the least similar images by measuring their difference with standard techniques of color histograms. The number of frames depends on the input video to be segmented. The video to be segmented can have either abrupt scene change or gradual scene changes such as dissolve, fade-in and fade out. The user can choose the number of segments required for segmentation of the video. Summarization is carried out for the length mentioned by the user. Selection function selects the excellent individual so that cross over and mutation operations become more effective. The

International Journal on Computational Science & Applications (IJCSA) Vol.8, No.4/5, October 2018
 research work proposes to modify the given genetic algorithm shown in figure 1 to make it
 suitable for the given application of video segmentation and summarization.

Let F be the input video to be segmented for the given application. The input video is down
 sampled to obtain few frames and F^1 represents the down sampled video. By finding the
 histogram difference procedure, obtain the video frames which are dissimilar and neglect the
 frames which are similar and F^{11} represents the video obtained after discarding frames. Apply GA
 by defining the encoding function and choosing initial candidates to calculate the fitness function.

The fitness function of GA is given by

$$f(S_k) = \sum_{i,j \in S_k} \alpha(i,j)h(i,j)$$

Where $\alpha(i,j)$ is a function for weighting the histogram differences $h(i,j)$.

The GA is modified by changing the fitness function which is given by

$$f(S_k) = \sum_{i,j \in S_k} \alpha(i,j)h(i,j) + k' \sum_{i \in S_k} h(i, i-1)$$

Where $h(i,i-1)$ or $dh(i)$ is the histogram difference between i^{th} and $(i-1)^{th}$ frame and K^1
 is $\frac{\text{constant}}{\text{scaling factor}}$ and user can choose the value. The experiments are conducted for different

values of K^1 and the result is discussed by selecting $K^1 = 5$. The proposed modified genetic
 algorithm works as follows.

1.If there is a scene change, the segmentation is done encoding it as '1'.

If there is a scene change at i^{th} frame, obviously the $(i-1)^{th}$

2.frame will be different.

3.Hence, the histogram difference between them may be high.

Perform cross over function to complete the GA.

3.IMPLEMENTATION DETAILS

Table 1: Input video considered for video segmentation & summarization (YUV video)

Video name	No of frames	Video name	No of frames
Akiyo	300 Frames	Carphone	382 Frames
Flower	250 Frames	Claire	494 Frames
Bridge (Close)	2000 Frames	Stefan	90 Frames
Bridge (Far)	210 Frames	Suzie	150 Frames
Coastguard	300 Frames	Tempete	260 Frames
Hall Monitor	300 Frames	Grandma	870 Frames
Bus	150 Frames	Mobile	300 Frames
Container	300 Frames	Silent	300 Frames
Foreman	300 Frames	Waterfall	260 Frames
Highway	2000 Frames	News	300 Frames
Miss America	150 Frames	Paris	1065 Frames
Mother and Daughter	300 Frames	Salesman	449 Frames

The experiments are conducted for different video clips with different types of scene changes such as abrupt and gradual scene changes. The different videos considered for the experimentation are shown in table 1. The different sampling factors and number of segments are considered for experimentation and the sample result is displayed for the discussion.

4.EXPERIMENTAL RESULTS AND DISCUSSIONS

Experiments are conducted for the different videos shown in the table 1 for different value of sampling factors such as 2, 3, 4, 5, 6, 7, 8, 9 and 10 for sub-sampling. The event considered for the experimentation consists of different scene changes such as movie scenes, television news, sports etc where scene changes may be abrupt or gradual. Also the experiments are conducted for different values of number of segments such as 2, 3, 4, 5, 6, 7, 8, 9 and 10. The summarization is carried out by considering different number of frames for summarization.

Experiments are conducted for various values of summarization such as 2, 3, 4, 5, 6, 7, 8, 9 and 10.

Input video considered: **Hall monitor**

Number of frames: 300

Frame Data set $F = \{F1, F2, \dots, F300\}$

Sampling factor = 10

Number of segments=8

Preprocessing step for segmentation is done using sub-sampling to obtain Frame data set F^1 given by

$$F^1 = \{F1, F11, F21, F31, F41, F51, F61, F71, F81, F91, F101, F111, F121, F131, F141, F151, F161, F171, F181, F191, F201, F211, F231, F241, F251, F261, F271, F281, F291\}$$

Least similar images are considered by measuring their difference with the standard technique of color histograms. On the reduced data set F^1 , define the length of an element i given by $\delta(i)$ which is the number of frames in F from 1 to the next element in F^1 . On F^1 , defining $dH(i)$ which is the histogram difference between the $(i-1)^{th}$ and i^{th} element of F^1 . Obtain the data set of frames by considering only those frames which are not similar by defining histogram difference. The data set contains

$$F^{11} = \{F1, F11, F31, F71, F81, F101, F141, F151, F211, F221, F231, F251, F261, F271\}$$

Histogram gives an estimate of the difference between the two frames. Two unlike frames with similar histogram gives rise to falls estimate of frames. The data set F^{11} may not be properly segmented the output data set if all the combinations of the histogram difference of the two frames of F^{11} are considered. Since most salient images may be similar to each other and too much repetition may occur in the new data set obtained due to histogram difference. In order to avoid the computational complexity involved in formulating the histogram data base by considering each frame distinctly, GA may be applied to achieve the reduction in the data set. GA is more methodological evaluation for segmentation and summarization of the frames.

GA starts by using the initialization procedure to generate the first population. The members of the population are usually strings of symbols (chromosomes) that represent possible solutions to the problem to be solved. To take into consideration the relative differences among all the selected images, we define similarity adjacency function. Genetic

segmentation algorithm is applied to obtain key frames. The genetic algorithm can be described by specifying the encoding, fitness function, and crossover operations. For the encoding of genetic algorithm, we take a string of 0^s and 1^s called a chromosome. The bit position of a chromosome string is an index for an element of the image data stream. Using 1^s to denote the boundaries, the desired segments can be obtained. The genetic mechanism works by randomly selecting pairs of individual chromosomes to reproduce for the next generation. The probabilities of a chromosome being selected are proportional to its fitness function value relative to the other in the same generation. The fitness function depends on the histogram difference and the chromosome structure depends on the maximum histogram difference. The maximum histogram difference is considered for summarization and denoted as ‘1’ where as ‘0’ indicates minimum histogram differences. Since the number of segments is 8, the chromosomes which indicate the segment boundaries contain 8 ones and remaining are filled with zeros. The structure of chromosomes follows the order F^{11} and is given by

{1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0}

Defining similarity adjacency function to obtain the better result of segmentation and the genetic algorithm is applied to obtain new data set of frames. To reproduce a crossover procedure is defined. The proposed application of segmentation does not uses mutation operation since it makes the segments unstable. The data set F^{111} is obtained by using chromosome patterns and F^{11} . Eliminate all 0’s and retain 1’s as 0’s correspond to similar frame information. Grouping is done to obtain the data set given by F^{111} and the result obtained for the input video hall monitor is given by

$F^{111}=\{F_1, F_{11}, F_{31}, F_{81}, F_{151}, F_{211}, F_{251}, F_{261}\}$

The modification of the genetic algorithm is done by modifying the fitness function which gives better results than the genetic algorithm. The chromosome of the modified GA is given by

{1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0}

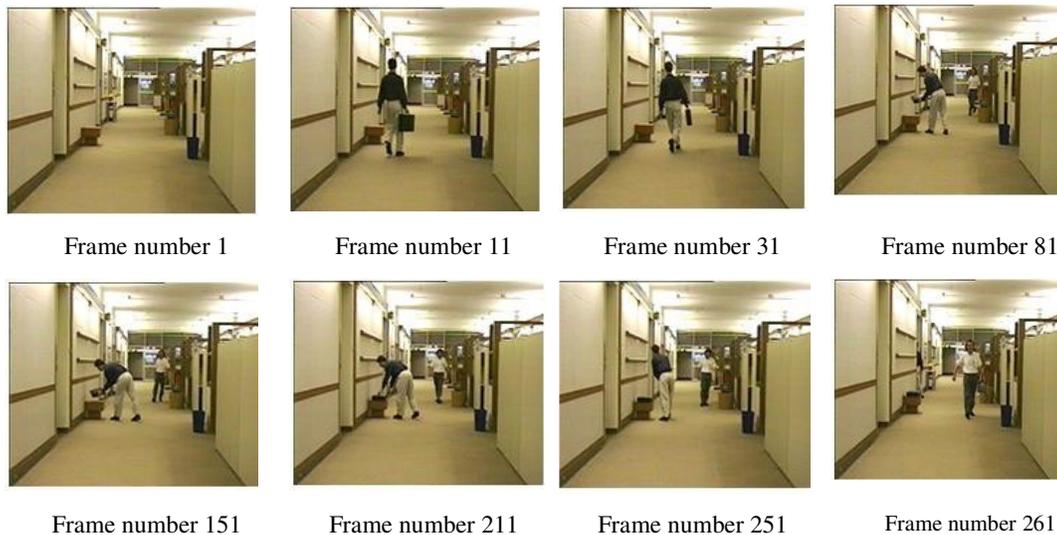


Figure 2: Summary frames of segmented video with number of segments=8 using GA

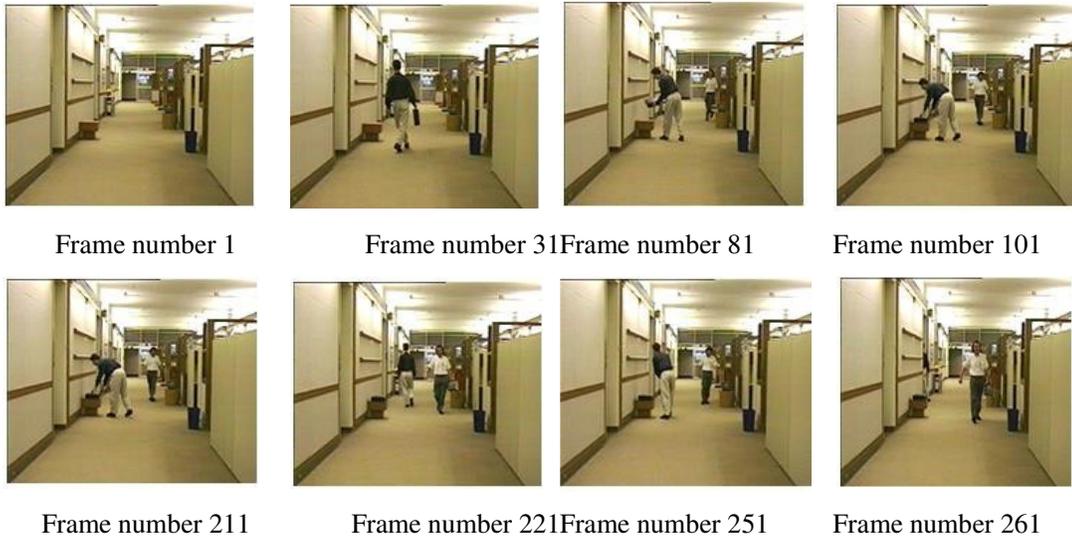


Figure 3: Summary frames of segmented video with number of segments=8 using proposed approach (modified GA)

Grouping is done to obtain the data set given by F^{111}

$$F^{111} = \{F_1^{11}, F_3^{11}, F_5^{11}, F_6^{11}, F_9^{11}, F_{10}^{11}, F_{12}^{11}, F_{13}^{11}\}$$

$$F^{111} = \{F_1, F_{31}, F_{81}, F_{101}, F_{211}, F_{221}, F_{251}, F_{261}\}$$

Table 2: Modified GA comparison with GA for different sampling factors

Sampling factor	Algorithm without modification selected frames	Algorithm with modification Selected frames
2	$\{F_1, F_{47}, F_{93}, F_{97}, F_{109}, F_{139}, F_{141}, F_{183}, F_{219}, F_{243}\}$	$\{F_1, F_5, F_{17}, F_{39}, F_{61}, F_{93}, F_{141}, F_{151}, F_{223}, F_{229}\}$
3	$\{F_1, F_{10}, F_{43}, F_{49}, F_{106}, F_{157}, F_{211}, F_{223}, F_{229}, F_{232}\}$	$\{F_1, F_{49}, F_{85}, F_{106}, F_{121}, F_{211}, F_{229}, F_{244}, F_{247}, F_{250}\}$
4	$\{F_1, F_{53}, F_{101}, F_{125}, F_{129}, F_{189}, F_{209}, F_{237}, F_{241}, F_{277}\}$	$\{F_1, F_{45}, F_{97}, F_{101}, F_{129}, F_{149}, F_{161}, F_{185}, F_{189}, F_{241}\}$
5	$\{F_1, F_{31}, F_{111}, F_{121}, F_{141}, F_{151}, F_{211}, F_{221}, F_{256}, F_{281}\}$	$\{F_1, F_{31}, F_{51}, F_{101}, F_{111}, F_{141}, F_{216}, F_{231}, F_{246}, F_{256}\}$
6	$\{F_1, F_7, F_{55}, F_{61}, F_{103}, F_{133}, F_{157}, F_{181}, F_{187}, F_{271}\}$	$\{F_1, F_{61}, F_{91}, F_{133}, F_{139}, F_{181}, F_{187}, F_{193}, F_{199}, F_{229}\}$
7	$\{F_1, F_8, F_{22}, F_{36}, F_{78}, F_{99}, F_{148}, F_{197}, F_{232}, F_{253}\}$	$\{F_1, F_8, F_{22}, F_{78}, F_{85}, F_{99}, F_{148}, F_{169}, F_{232}, F_{253}\}$
8	$\{F_1, F_{25}, F_{41}, F_{73}, F_{89}, F_{97}, F_{153}, F_{177}, F_{193}, F_{273}\}$	$\{F_1, F_{33}, F_{41}, F_{89}, F_{97}, F_{153}, F_{177}, F_{193}, F_{217}, F_{273}\}$
9	$\{F_1, F_{37}, F_{64}, F_{73}, F_{82}, F_{118}, F_{136}, F_{163}, F_{181}, F_{253}\}$	$\{F_1, F_{37}, F_{64}, F_{82}, F_{100}, F_{118}, F_{136}, F_{163}, F_{181}, F_{253}\}$
10	$\{F_1, F_{11}, F_{31}, F_{81}, F_{151}, F_{211}, F_{221}, F_{231}, F_{251}, F_{271}\}$	$\{F_1, F_{11}, F_{31}, F_{71}, F_{101}, F_{141}, F_{151}, F_{211}, F_{261}, F_{271}\}$

The modified algorithm removes the frames F_{11} and F_{151} which contain the information which is almost same as the information present in frames present in the F_1 and F_{81} . But frames F_{101}

International Journal on Computational Science & Applications (IJCSA) Vol.8, No.4/5, October 2018
 and F_{221} are summarized which are more significant for the given application shown in figure 3. The results show that modification in the genetic algorithm has helped to improve the efficiency of the GA for the input video clip. Some of the frames with almost similar information content are repeated in the GA and modified algorithm removes those frames to a greater extent.

Similar procedure is conducted for different value of scaling factors and the table 2 shows the result obtained for various sampling factors.

The experimental results in the table 2 shows that the result obtained using modified GA is better than the genetic algorithm in summarizing some of the key frames. The experiments are conducted for different target segments for the video hall monitor using GA and modified GA and sampling factor of 10. The table 3 shows the result obtained for the given input video hall monitor.

Table 3: Modified GA comparison with GA for different target segments

Target segments	Algorithm without modification	Algorithm with modification
5	{ $F_1, F_{11}, F_{81}, F_{151}, F_{271}$ }	{ $F_1, F_{31}, F_{71}, F_{81}, F_{271}$ }
6	{ $F_1, F_{11}, F_{31}, F_{71}, F_{221}, F_{231}$ }	{ $F_1, F_{81}, F_{101}, F_{141}, F_{221}, F_{271}$ }
7	{ $F_1, F_{31}, F_{141}, F_{151}, F_{221}, F_{251}, F_{271}$ }	{ $F_1, F_{11}, F_{71}, F_{81}, F_{151}, F_{221}, F_{231}$ }
8	{ $F_1, F_{11}, F_{31}, F_{81}, F_{151}, F_{211}, F_{251}, F_{261}$ }	{ $F_1, F_{31}, F_{81}, F_{101}, F_{211}, F_{221}, F_{251}, F_{261}$ }
9	{ $F_1, F_{31}, F_{81}, F_{101}, F_{141}, F_{151}, F_{231}, F_{251}, F_{271}$ }	{ $F_1, F_{31}, F_{71}, F_{101}, F_{141}, F_{151}, F_{231}, F_{251}, F_{271}$ }
10	{ $F_1, F_{11}, F_{31}, F_{81}, F_{101}, F_{151}, F_{221}, F_{231}, F_{251}, F_{271}$ }	{ $F_1, F_{11}, F_{31}, F_{71}, F_{81}, F_{101}, F_{141}, F_{211}, F_{251}, F_{271}$ }

It is seen that if the number of target segments is more, the modified GA result is almost same as GA. Since most of the applications need is to segment only few frames, modified GA performs better than GA.

Table 4: Profiling result for scaling factor of 10

Function Name	Algorithm without modification			Algorithm with modification		
	Calls	Clock cycles 1 (MCPS)	Clock cycles 2 (MCPS)	Calls	Clock cycles 1 (MCPS)	Clock cycles 2 (MCPS)
calculate fitness function	1344	43.4	43.4	1680	43.4	43.4
calculate fitness function Value	12	326.2	239.4	15	634	460.6
histogram difference of two frames	3	21	21	3	0	0
histogram matrix of the frames)	360	1047	740.4	360	457.2	347.2

The experiments are conducted for the different video clips and the profiling result obtained is shown in the table 4. The profiling result gives an estimation of clock cycles consumed and the number of times the function is called. The clock cycle 2 is the number of clock cycles spent in a function excluding the clock cycle spent in its child functions. The clock cycle 2 includes overhead resulting from the process of profiling. The profiling result obtained indicate that the number of clock cycles required to segment and summarize the frames using modified GA is

International Journal on Computational Science & Applications (IJCSA) Vol.8, No.4/5, October 2018
almost same as GA. Hence, the profiling result in table 4 indicates that the modified GA doesn't increase the computational complexity. The modified GA performs better than GA in segmentation and summarizing the frames without increasing the computational complexity.

5.SUMMARY & CONCLUSIONS

The number of segments can be defined by the user and for every time when the user wants to change the numbers of segments, the user no need to change the program just the user need to change the number of segments. For summarization also, the number of frames can be defined by the user and can be changed whenever the user needs to change. The genetic algorithm which is used for video segmentation and summarization is useful because it is the random search method and is an optimization algorithm which estimates the solution and the algorithm approaches the global value rather than the optimal value. From the results obtained, it is noted that the modified GA works better than GA in summarizing the important frames which are non-repetitive.

6.REFERENCES

- [1] Calic, J, et.al. (1993), "Temporal video segmentation for real-time key frame extraction", IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-93. Volume 4.
- [2] Chen Yinzi, et. al. (2010), "A Temporal Video Segmentation and Summary Generation Method Based on Shots' Abrupt and Gradual Transition Boundary Detecting", Second International Conference on Communication Software and Networks.
- [3] Aravind. I, et. al. (2002), "Implementation of Image Segmentation and Reconstruction Using Genetic algorithms", IEEE, ICIT'02, Bangkok, Thailand
- [4] Patrick Chiu, et.al. (2000), "A genetic algorithm for video segmentation summarization", IEEE International conference on multimedia and expo ICME2000, Proceedings, Vol.3, pp.1329-133.
- [5] Lei Hui, et. al.(2008), "Application of an Improved Genetic Algorithm in image segmentation", International Conference on computer Science and Software Engineering".
- [6] Bir Bhanu, et. al. (1995), "Adaptive Image segmentation using a genetic algorithm", IEEE transaction on systems, man, and cybernetics, vol 25, No.12, pp.1543-1567.
- [7] Chi-Chun Lo, et. al. (2001),"Video segmentation using a histogram-based fuzzy c-means clustering algorithm", Computer Standards & Interfaces 23, pp 429-438.
- [8] P.Kanungo, et. al.(2006),"Parallel Genetic Algorithm Based Thresholding for Image Segmentation", Proceedings of the National Seminar on IT and Soft computing", ITSC06, India.
- [9] Chi-Chun Lo, et. al. (2003), "A histogram-based moment-preserving clustering algorithm for video segmentation", Pattern Recognition Letters 24, pp. 2209-2218.

Author

Dr. Prasantha. H S received Bachelor degree from Bangalore University, Master Degree from V.T.U, Belgaum, and Ph.D from Anna University, Chennai, in the area of Multimedia and Image Processing. He has 19+ years of teaching and research experience. His research interest includes Multimedia and SignalProcessing and published over 30 papers in conferences and Journals. He is currently guiding students for their research program under VTU and other university. Currently, he is working as a Professor in the department of Electronics and Communication Engineering, Nitte Meenakshi Institute of Technology, Bangalore.

