International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference

Ganesh. I. Rathod[1], Dipali. A. Nikam[2]
[1]Department of Computer Science & Engineering, Dr. J. J. Magdum college of Engineering, Jaysingpur
[2]HOD, Department of Computer Science & Engineering, Dr. J. J. Magdum college of Engineering, Jaysingpur

***Abstract*— Shot boundary detection is the complete segmentation of a video into continuously imaged temporal video segments. Condensed video representation is the extraction of video frames or short clips that are semantically representative of the corresponding video. Both tasks are very significant for the organization of video data into more manageable forms. 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 square histogram difference considered at block level for the video frames, a new method of extracting the keyframes based on shot type is presented.**

***Keywords*—Shot Boundary Detection, KeyFrame Extraction.**

I. INTRODUCTION

Video is the most effective media for capturing the world around us. Applications such as multimedia information systems, distance learning uses huge amount of video data. This has lead to an increasing demand of efficient techniques to store, retrieve, index and summarize the video content. Searching video media on the web often means to use the available search interfaces provided by online video portals. YouTube as the market leader provides keyword search based on manually generated meta information and tags. Unfortunately, metadata is limited in its ability of representing the content of a video, and tags are subjective labels that might be misleading in their semantics. This is why content-based video retrieval (CBVR) can improve video search. One key process in content-based video retrieval is the extraction of keyframes which then can be analyzed with known image processing methodologies. The increased demand for intelligent processing and analysis of multimedia information has led to the development of different approaches for intelligent video management. Among these approaches, shot transition detection is the first step of content-based video analysis and keyframe is a simple yet efficient form of video abstract.

II. PREVIOUS WORK

Shot boundary detection is an early step for most of the video applications involving the understanding, indexing, characterization, or categorization of video, temporal video segmentation has been an active topic of research in the area of content based video analysis. In this chapter an overview of different categories of solutions to primary problem of shot boundary detection is presented. The key frame extraction is primary task in producing summaries and skims of video. There are various algorithms proposed as solution to problem of key frame extraction based on various modality information of frames and special events on the temporal axis of the video. The chapter is also presenting literature review of these mechanisms. The detection of shot boundaries provides a base for nearly all video abstraction, indexing and high-level video segmentation. Therefore, solving the problem of shot-boundary detection is one of the major prerequisites for revealing higher level video content structure analysis and retrieving. Moreover, other research areas can profit considerably from successful automation of shot-boundary detection processes as well. There are various approaches proposed based on different feature parameters to detect shot boundary and keyframe extraction. The problem of high quality key frame extraction and video summarization detection algorithm by using QR decomposition [1] method some efficient shot measures are derived with the help of same. And further the keyframes are derived from QR decomposition. An efficient video summarization by keyframe extraction system, need two presumptions. First one of which is that a mathematical criterion to measure the video dynamicity for detecting the number of keyframes in each shot needed to produce a summary with a predefined length. The second presumption is an accurately method that detects the independent keyframes within shots. It has efficient properties of keyframe extraction with low redundancy.

A neural network scheme for adaptive video indexing and retrieval [2] makes use of extraction of limited but characteristic amount of frames by using the cross correlation criteria to represent efficient scene summarization.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

Low level features are extracted to indicate the frame characteristics. A two step algorithm for efficient extraction of keyframe makes use of minimal spanning tree graph where each frame forms the node in first step, then in the second step extract the keyframe based on their maximum spread [3]. Pixel- or region-based approaches use absolute or relative differences between pairs of single pixels or pixel regions. Frame-based approaches compare features extracted from consecutive frames e.g. frame size or the change of intensity of edges .Histogram based approaches compare color or grayscale histograms [4]. Edge-based approaches measure the similarity or dissimilarity as change in intensity of edges. The method of extracting the keyframes by Accumulated Histogram Intersection Measure [5] first extract key frames by matching of DC image sequence constructed from the MPEG video sequence. Then using the region segmentation-based projective histogram and its moments as database indices for video retrieval. Results clearly validate and dominate the method and that in conjunction with other indexing techniques, such as color, can provide a poweful framework for video indexing and retrieval. Efficient keyframe extraction, using local semantics in form of a region thesaurus specifically, works on certain MPEG-7 color and texture features are locally extracted from keyframe regions. Then, using a hierarchical clustering approach a local region thesaurus is constructed to facilitate the description of each frame in terms of higher semantic features [6]. These region types carry semantic information. Each keyframe is represented by a vector consisting of the degrees of confidence of the existence of all region types within this shot. A method of keyframe selection [7] is a thing where a single frame is selected as the keyframe from a video frame sequence. The technique is composed of three steps: shot boundaries detection, shot selection, and key frame extraction within the selected shot like, first the video is divided into shots, then the best shot is determined, and finally a representative frame is extracted from the selected shot. The shot and key frame are selected based on measures of motion and spatial activity and the likeliness to include people .Clearly, the technique can be extended to extract *n* key frames by selecting *n* shots and subsequently a key frame from each selected shot. By analyzing the similarity between two consecutive frames of a video sequence, the algorithm determines the complexity of the sequence in terms of changes in the visual content expressed by different frame descriptor features [8].

The keyframes are extracted by detecting curvature points within the curve of the cumulative frame differences. Method of Independent Component Analysis (ICA) [9], suggests that projecting video frames from illumination-invariant raw feature space into low dimensional ICA subspace; each video frame is represented by a two-dimensional compact feature vector. An iterative clustering algorithm based on adaptive thresholding is developed to detect cuts and gradual transitions simultaneously in ICA subspace. A simple approach stills an effective framework for features extraction from an athletic sport sequence [10].

On analysing all the research works carried the methods for Shot boundary detection can be listed as, Pixel based, Histogram based, Edge based, Motion based etc. In case of pixel based basic idea is that the intensity values of the pixels at the same locations of the sequential frames do not change significantly unless there is a shot boundary. The initial pixel based algorithms investigates the sum of absolute pixel intensity differences and if the difference is above a certain value a shot boundary is assigned. Even very small changes in the illumination or very small vibration in the camera can result in significant changes in total value of the pixel differences. Another method called as the histograms which do not change with the spatial changes within a frame, histogram differences are more robust against the object motion with a constant background. However, histogram differences are also sensitive to camera motion, such as panning, tilting or zooming. Histograms represents the global intensity of colors in a frame, sometimes two frames may have significantly same histogram. Further the edges also proved useful in shot boundary.

III. Proposed Algorithm For SBD And KFE

Shot boundary Detection task can be achieved using various approaches such as pixel intensity based, histogram-based, edge-based, and motion vectors based, are implemented and analyzed. Among all the approaches Histogram difference is the popular approach. Histogram based method is interested with the global percentage of colors that an image contains. In these histogram-based approaches, pixels, space distribution was neglected. The diagrammatic representation of proposed problem can be shown as follows:

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

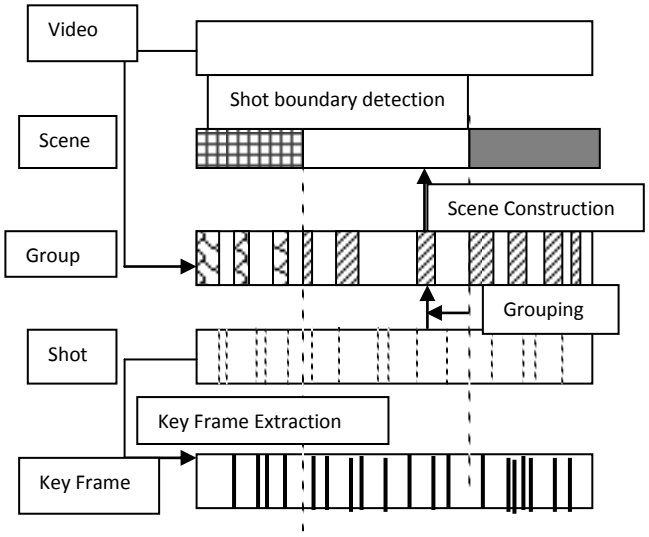

Figure 3.1. Proposed Work.

The problem is implemented by dividing each frame in to blocks and considering histogram difference of the consecutive frames between corresponding blocks by assigning weights to blocks. Then the threshold is calculated by calculating mean deviation and standard deviation. At the last if the frame difference is above the threshold then it is keyframe.

The proposed algorithm can be explained in following steps:

Algorithm:

Step 1: Read the input video and is then passed through a function called ShotBoundaryDetection .

Step 2: The function ShotBoundaryDetection converts frames into sub frames by using function divideIntoSubFrames.

Step 3: The function divideIntoSubFrames determines the block difference for each sub frame.

Step 4: Sum of block differences is the calculated using the formula

Block difference = (histogram of 1^{st} – histogram of 2^{nd}) no of gray levels

---

1^{st} Histogram .

Step 4: Mean deviation and standard deviation are also calculated. Threshold is determined by the formula

Threshold = mean deviation+ (a* standard deviation).

Step 5: If the frames block difference is greater than threshold then it is a keyframe.

Step 6: The above steps are repeated for the complete video and all the keyframes are determined and stored in a folder called keyframes.

Step 7: Finally display the Shot Type, as Static or Dynamic depend on input video.

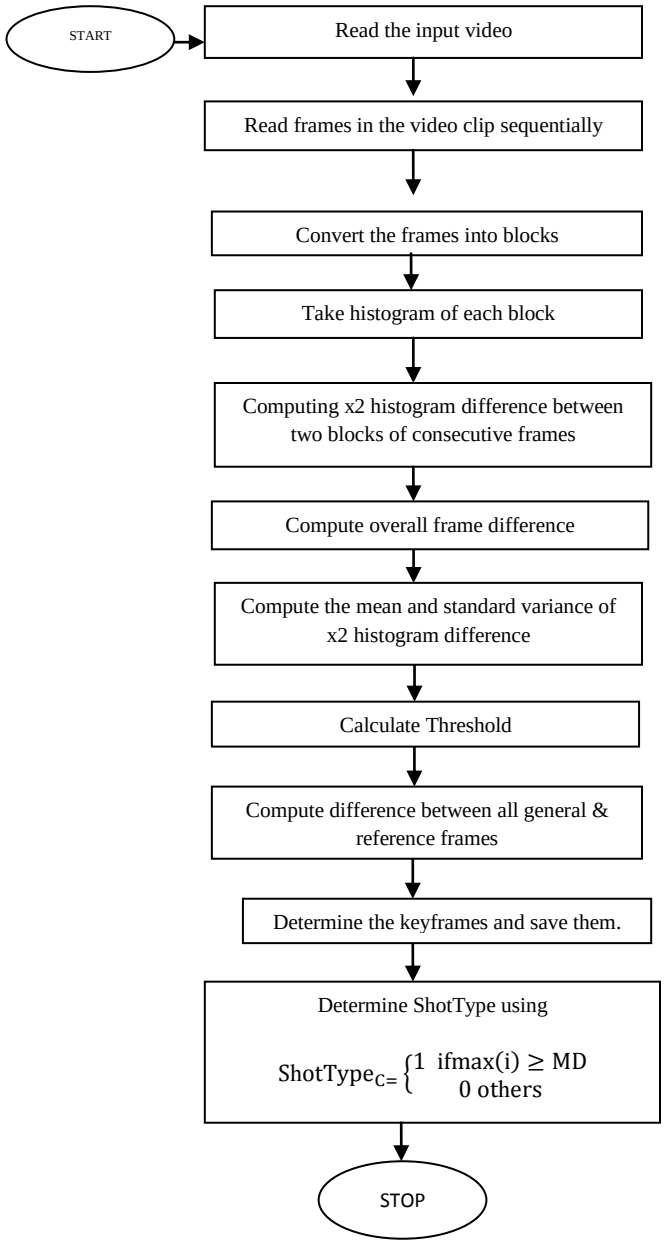

Figure 3.2. Design Flow of proposed algorithm for SBD and KFE.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

A. Shot boundary detection

Shot Boundary Detection is an early step for most of the video applications involving the understanding, indexing, characterization, or categorization of video, temporal video segmentation etc. This has been an active topic of research in the area of content based video analysis. At the implementation level the idea of Shot boundary detection is achieved by three major steps,

- *Image Segmentation:*-In the Image segmentation level, first dividing each frames obtained from the video in to blocks of m-Rows and n-columns. Then the difference of the corresponding blocks between two consecutive frames is computed. Finally, the final difference of two frames is obtained by adding up all the differences through different weights.
- *Attention Model:* - Attention, a neurobiological concept, means the concentration of mental powers upon an object by close or careful observing or listening, which is the ability or power to concentrate. Attention model means that, from the visual viewpoint, different contents are ranked based on importance correspondingly, it also reflected the importance of frames. Based on the consideration, one can think that different position's pixels have different contribution to shot boundary detection: pixels on the edge are more important than others. Thus, different weights are given to blocks of different position. Both the space distribution characteristic of pixels of different gray and the different importance of pixels of different position are considered.
- *Matching the difference:* - There are many kinds of histogram match. Color histogram was used in computing the matching difference in most literatures. However, through comparing several kinds of histogram matching methods, thus reached a conclusion that x^2 histogram outperformed others in shot boundary recognition [11]. Hence, x^2 histogram matching method is proposed.

Algorithm for Shot Boundary Detection:-

Let F(k) be the k^{th} frame in video sequence, k = 1,2,.., Fv （ Fv denotes the total number of video). The algorithm of shot boundary detection is described as follows:

Step 1: Partitioning a frame into blocks with *m* rows and *n* columns, and *B*(*i*, *j*,*k*) stands for the block at(i, j) in the k^{th} frame.

Step 2: Computing x^2 histogram matching difference between the corresponding blocks between consecutive frames in video sequence. *H*(*i*, *j*,*k*) and *H*(*i*, *j*, *k* +1) stand for the histogram of blocks at (*i*, *j*) in the k^{th} and $(k+1)^{th}$ frame respectively. Block's difference is measured by the following equation:

$$D_B(k, k+1, i, j) = \sum_{l=0}^{L-1}[(H(i,j,k) - \mathrm{H(i,j,k+1)}]^2 \ /\mathrm{H(i,j,k)}.$$

Where, L is the number of gray in an image;

Step 3: Computing x^2 histogram difference between two consecutive frames:

$$\mathrm{D}(k, k+1) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij} D_B(k, k+1, i, j)$$

where, w_{ij} stands for the weight of block at (*i*, *j*) .

Step 4: Computing Threshold Automatically: Computing the Mean and standard variance of x^2 histogram difference over the whole video sequence. Mean and standard variance are defined as follows：

$$\mathrm{MD} = \sum_{k=1}^{F_{V-1}} \mathrm{D(k,k+1)} \ /F_{V-1}$$

$$STD = \sqrt{\sum_{k=1}^{F_{V-1}} (D(k,k+1) - \mathrm{MD})^2 \ /F_{V-1}}$$

Step 5: Shot boundary detection.

Let threshold $T = MD + a \times STD$.
Where a is the constant. Say a=1.

Shot candidate detection:

If $D(i,i+1) \geq T$, the i^{th} frame is the end frame of previous shot, and the $(i+1)^{th}$ frame is the end frame of next shot.Final shot detection: Shots may be very long but not much short, because those shots with only several frames cannot be captured by people and they cannot convey a whole message. Usually, a shortest shot should last for 1 to 2.5 s. For the reason of fluency, frame rate is at least 25 fps, (it is 30 fps in most cases), or flash will appear. So, a shot contains at least a minimum number of 30 to 45 frames. In our case, video sequences are down sampled at 10 fps to improve simulation speed. On this condition, the shortest shot should contain 10 to 15 frames. 13 is selected for our experiment. Thus formulate a "shots merging principle": if a detected shot contain fewer frames than 13 frames, it will be merged into previous shot, or it will be thought as an independent one. With reference to the above discussions certain definitions are defined as,

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

- Reference Frame: It is the first frame of each shot.
- General Frames: All the frames except for reference frame;
- "Shot Dynamic Factor" max(i): The maximum x^2 histogram within shot i;
- Dynamic Shot and Static Shot: a shot will be declared as dynamic shot, if its max(i) is bigger than MD; otherwise it is static shot;
- $F_C(k)$: The kth frame within the current shot, k=1,2,3… $F_{CN}(k)$ ($F_{CN}(k)$ is the total number of the current shot)

B. Key Frame Extraction

Keyframes play an important role in video abstraction. Keyframes are a set of salient images extracted from video sequences. They provide a simple yet effective way of summarizing the content of videos for browsing and retrieval and are also widely used in video abstraction due to their compactness. Much research has been conducted in the past few years in understanding the problem of key-frame extraction and developing effective algorithms. Although simple and computationally efficient sampling-based methods may produce no keyframes for a shot, yet semantically producing too many keyframes with identical content to represent a long static segment thus failing to effectively represent the actual video content. The keyframe extraction algorithm discussed, here will focus only on techniques that take into account the underlying dynamics, to different degrees and from varying viewpoints, of the video sequence. The algorithm of keyframe extraction is described as follows:-

Step 1: Computing the difference between all the general frames and reference frame with the above algorithm:

$$D_C(1,k) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij} D_{CB}(1,k,i,j), k = 2,3,4, \ldots\ldots F_{CN}$$

Step 2: Searching for the maximum difference within a shot:

$$\max(i) = \{D_C(1,k)\}_{max}, K = 2,3,4, \ldots\ldots F_{CN}$$

Step 3: Determining "Shot Type" according to the relationship between max(i) and MD:
Static Shot(0) or Dynamic Shot:

$$ShotType_{C=}\begin{cases}1 & if max(i) \geq MD \\ 0 & others\end{cases}$$

Step 4: Determining the position of keyframe:

If $ShotType_C$=0, with respect to the odd number of a shot's frames, the frame in the middle of shot is chose as keyframe; in the case of the even number, any one frame between the two frames in the middle of shot can be chose as keyframe.

If $ShotType_C$=1, the frame with the maximum difference is declared as keyframe.

IV. Experimental Results

The proposed work is implemented in Matlab Version 2010b, image and video processing tools are employed here. The algorithm is executed on Pentium® Dual-Core processor with 3GB RAM memory. The proposed model accepts the input video in the form of ".avi" and the size of the video ranges from 3MB to 4.5MB. The experimental dataset used in this project are from the Open Video Project.

Proposed algorithm is tested with videos of different types. Results for such samples are listed as below.

- *Sample 1:*

Figure 4.1. Input clip Sample 1 with Size: 3.42 MB, Total number of frames: 623.

For input sample clip1, the algorithm gives following frames as the keyframes for this video.

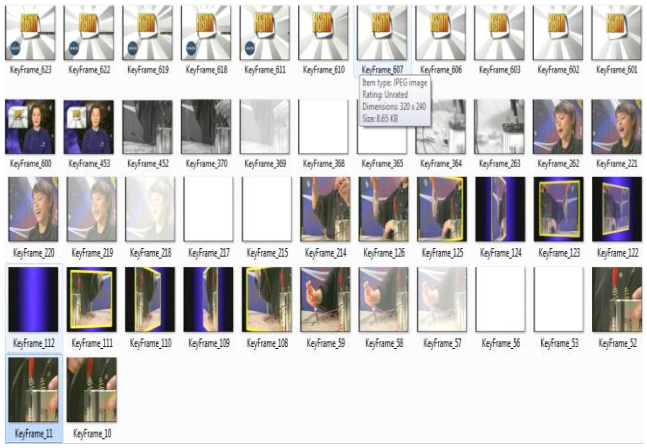

Figure 4.2. Extracted keyframes for the input clip Sample 1 Keyframes =46, Shot Type: Dynamic Shot.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

Figure 4.2 clearly demonstrate the extracted keyframes. Frame 10 and 11 shows the transition effect of hand movement. Abrupt cut can be seen in keyframe 56, and keyframe 53 differs from 56 by a small dot. Keyframes from 57 to 59 demonstrates dissolve effect. Hence the frames are continuous. Keyframes from108 to 214 indicates wipe effect hence all the frames are extracted at output. Keyframe 112 denotes cut. Further the same dissolve effect can be seen from Keyframes from 215 to 221 and 364 to 370. Further the frames from 602 to 623 demonstrate the wipe effect. Thus the algorithm has successfully created the summary of the video and reduction in frames at around less than 10%.

- *Sample 2 (Cartoon Clip) :*

Figure 4.3. Input cartoon clip Sample 2 with Size: 2.36MB Total number of frames: 407.

The output for sample 2 video is shown below:-

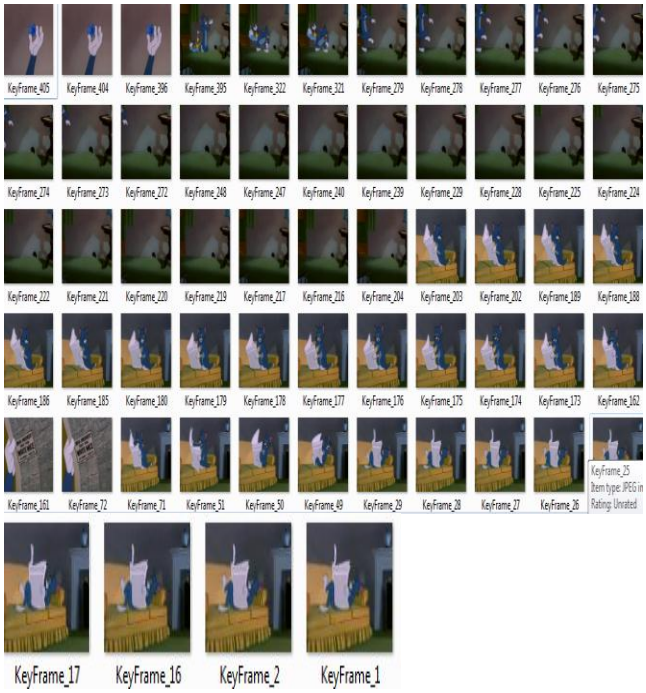

Figure 4.4. Extracted keyframes for the input cartoon clip sample2. Keyframes: 59, Shot Type: Dynamic Shot.

For the input sample 2 generated keyframes demonstrate the detection of shot boundary detection. Keyframes from 1 to 17, 26 to 71 and 162 to 203 demonstrates the considerable movement of the object thus causing considerable effect on the background of the frame. So all such frames are listed as keyframes. Abrupt cut can be seen in Keyframe 72. Keyframe from 204 to 395 clearly demonstrates the wipe effect in shot boundary detection. Thus there is considerable increase in the resultant keyframes. Thus the algorithm has successfully created the summary of the video and reduction in frames at around 10%.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

- *Sample 3 (Sports Clip)*

Figure 4.5. Input sports clip Sample 3 with Size: 3.59 MB Total number of frames: 718.

From figure 4.6 below, it can be clearly seen that the wipe effect lasts from frame 260 to 272 and 493 to 505, as the object is slowly entering and it leaves the frame. Thus all the frames are extracted. Keyframes 335 to 340 demonstrates the shot play actions by a batsman. Keyframe 341 to 349 and 419 to 434 demonstrates the movement of camera with respect to the ball, and continuous change in the background is seen. Thus all these are named as keyframes. Thus we conclude that wipe effect and camera motion rate is high in sports video. Thus the algorithm has successfully created the summary of the video and reduction in frames at around less than 10%.

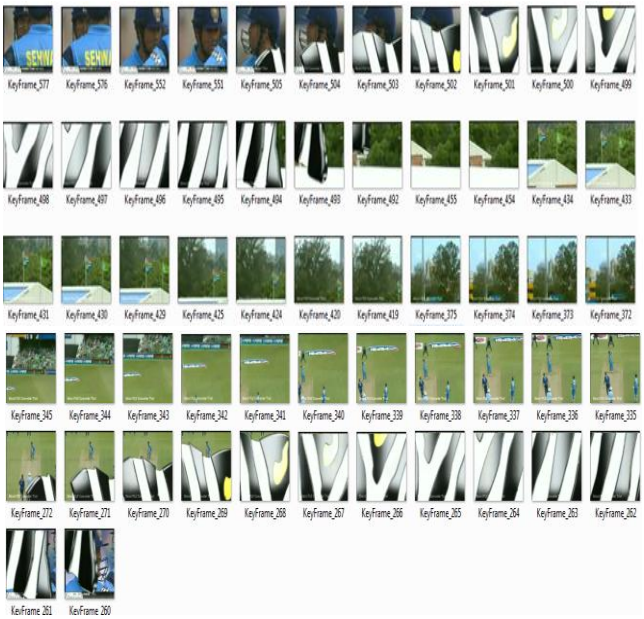

Figure 4.6.Extracted keyframes for the input sports clip. Keyframes: 70, Shot Type: Dynamic Shot.

TABLE I
COMPARISON OF NUMBER OF KEYFRAMES OBTAINED.

Video	Size of Video	Number of Keyframes
Movie	3.42MB	49
Cartoon	2.36MB	59
Sports	3.59MB	70

The table above shows the comparison of the outputs i.e., the number of keyframes detected for different types of video clips.

For the valuation purpose a test on around 40 videos from open video project source. It is observed that the keyframes extracted from the video forms the effective summary of the video and also our algorithm can identify significant shot changes.

TABLE II
SAMPLE VIDEO FROM THE OPEN VIDEO PROJECT WITH SIZE ON DISK, TOTAL FRAMES, AND EXTRACTED KEYFRAMES.

Sl.No	Size of video in MB	Total frames in video	Keyframes Obtained
Sample1	3.68	605	64
Sample2	3.42	623	46
Sample3	3.15	575	47
Sample4	3.80	700	81
Sample5	3.06	600	68
Sample6	2.05	533	68
Sample7	2.49	638	85
Sample8	2.43	545	99
Sample9	2.13	598	45
Sample10	3.58	582	53

In order to know the efficiency of the designed algorithm a simple method is adopted, i.e, analyzing each video then listing out the manual shot change. Then the same video is applied as the input to the algorithm, finally the obtained keyframes are checked for those manually identified shots changes. Surprisingly the algorithm had given almost 100% results.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 8, August 2013)

Further usually more number of keyframes are observed in most of the scenes because of the wipe effect, dissolve action and fade in/ out effects of video shot transition. The result of the same is shown in the table 3.One more effort is made to check the efficiency of the algorithm by conducting the following experiment. Twenty people are divided into two groups: group A and group B. Each group consists of ten people. Group A first watch video and then evaluate the keyframes [Summarization]. Contrarily, Group B first watches keyframes [Summarization] and then watch video, and finally evaluate the summarization. Zero score means that the keyframe cannot represent video's content at all. Six scores mean that it can represent the main content of a video basically. Ten scores mean that it can represent the content completely. Table 4. shows the final result of evaluation. The table shows that the algorithm is very efficient. The summarization can not only compress the video and but also conceive its content. The reason for Scores of group B being slightly higher than that of group A is that under condition of viewers without knowing the content they are satisfied with the summarization.

TABLE III
PERFORMANCE EVALUATION OF KEYFRAME EXTRACTION VIA MANUAL SHOT BOUNDARIES SEEN.

Sample Clips [Size]	Total Number of Frames in clip	Keyframes detected by algorithm	Manual SBD Seen	Detected SBD by Algorithm
Sample1[3.68MB]	605	64	09	09
Sample2[3.42MB]	623	46	08	08
Sample3[3.15MB]	575	47	06	06
Sample4[3.80MB]	700	81	07	07

TABLE IV
PERFORMANCE EVALUATION OF KEYFRAME EXTRACTION

Video	Score A	Score B
Movie	8.5	8.9
Sport	8.0	8.8
Cartoon	8.6	8.5
Education	8.4	8.7

V. CONCLUSION

Video abstraction is an integral part of many video applications, including video indexing, browsing, and retrieval. Shot boundary detection is the process of identifying the significant content change in the video. In this work a Square histogram based model is developed using frame segmentation and automatic threshold calculation. In this paper the keyframe is extracted by using a reference frame approach per shot. A total of around 40 videos of different types are tested on this model and the model is able to detect all shot boundaries and is storing the suitable frames as keyframes to represent the video summary. An efficiency of almost 95% to 98% is observed using this algorithm.

A key point to note that significance increase in the number of keyframes occurs whenever the special effects such as wipe, dissolve and fading are observed in the video. The limitation of the present work is on tool used, as it has certain constraint related to memory.

REFERENCES

[1] Ali Amiri, Mahmood Fathy, Atusa Naseri , Computer Engineering Department, Iran University of Science and Technology, Tehran, Iran " **Keyframe extraction and video summarization using QR-Decomposition**", IEEE 2007.

[2] Anastasios D. Doulamis, Nikolaos D. Doulamis and Stefanos D. Kollias , National Technical University of Athens, Department of Electrical and Computer Engineering , Heroon Polytechniou, 157 73 Zografou, Greece "**Relevance feedback for content-based retrieval in video databases: a neural network approach**".

[3] D.Besiris, F. Fotopoulou, N. Laskaris, G. Economou, Department of Physics, Electronics Laboratory, University of Patras "**Key frame extraction in video sequences: a vantage points approach**", IEEE 2007.

[4] EMRAH AŞAN, M.S, Department of Electrical and Electronics Engineering, The Graduate School Of Natural And Applied Sciences of Middle East Technical University "**Video Shot Boundary Detection by Graph Theoretic Approaches**", September 2008.

[5] Eung Kwan Kang, Sung Joo Kim, and Joon So0 Choi, Dept.of Physics Chung-Ang University, 221 Huksuk-Dong, "**VIDEO RETRIEVAL BASED ON SCENE CHANGE DETECTION IN COMPRESSED STREAMS",** IIT assessment.

[6] Evaggelos Spyrou and Yannis Avrithis, National Technical University of Athens Image, Video and Multimedia Laboratory Zographou, 15773 Athens, Greece. "**Keyframe Extraction using Local Visual Semantics in the form of a Region Thesaurus**", IEEE 2007.

[7] Frkdkric Dufaux, Compaq Computer Corp., Cambridge Research Lab.," **Key frame selection to represent a video",** 0-7803-6297-7/00/ IEEE 2000.

[8] G. Ciocca1, R. Schettini Dipartimento di Informatica Sistemistica e Comunicazione (DISCo) Università degli studi di Milano-Bicocca "**An innovative algorithm for key frame extraction in video summarization**".

[9] Gentao Liu Xiangming Wen Wei Zheng Peizhou He School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, **"Shot Boundary Detection and Keyframe Extraction based on Scale Invariant Feature Transform",** Eigth IEEE/ACIS International Conference on Computer and Information Science, 2009

[10] Giuseppe Caccia, Rosa Lancini, CEFRIEL - Politecnico di Milano Milan, Italy**, " Algorithm for Summarization and Ketframes extraction in Athletic video" 0-**7803-?713-3/02 IEEE 2002

[11] Lijie Liu, Student Member, IEEE and Guoliang Fan, Member, IEEE, "**Combined Key-Frame Extraction and Object-Based Video segmentation**", IEEE Transactions on circuits and systems for video technology, vol. 15, no. 7, July 2005.