An Effective Method of Video Segmentation and Summarization for Surveillance

Pratyush Pranjal, Rohit Lakhdive, Aditya Vasave, AtulkumarVarma and AmitBarve
Vishwakarma Institute of Information Technology
Pune, India

ABSTRACT
In today's world, video surveillance has become an integral part of investigation system. Video Segmentation and Summarization plays a crucial role in the field of Video analysis. The paper presents an approach that helps in extracting the movable segments from a large stable video. The system takes the video as input and extracts the movements from the video and identifies the time-intervals. Initially frames (images) are extracted from the video and compared consecutively. If some movements are detected then the movement signatures (start time – end time) are stored. Based on signaturesthe segmented and summarized video is generated. The result shows substantial improvement in video summarization process.

Keywords
Key frame; Video Segmentation and Summarization, P-Hash

1. INTRODUCTION
These video surveillance systems are used at a very large scale everywhere. But nowadays the problem faced by security systems is,a huge amount of data generated but meaningful part in that data is very small as compared to the data which is stored during surveillance. So it is difficult for us to identify the meaningful part and it is also a time consuming process. The observation in the surveillance was that the most part of the video did not contain any substantial movements. Thereforeit is expected to extract only that part of video which has some meaningful movements. This approach makes us to concentrate only on the important events and ignore the meaningless part. By following this method we generate small segments of videos. Later on we merge these small segments into a single video which will be the summarized part for the whole video.

2. RELATED WORK
Researchers are working in the field of Video Analysis since years. Katkar and Barve [1] proposed a new technique for Content Based Video Retrieval (CBVR) of Animal Videos based on Antipole tree indexing algorithm. Tao Xiang and Shaogang [2]proposed that the aim of their work are two-fold: firstly a continuous video is segmented according to the activities captured in the video; secondly a model is constructed for the video content, based on which an unseen activity pattern can be recognized and any unusual activities can be detected. Mural and Girisha [3]suggested a development of an efficient adaptive object segmentation algorithm for color video surveillance sequences.McKoen, Navarro-Prieto et al [4]proposed that in evaluation of video of objective figures of merit, there is no standardized methodology for subjective evaluations of segmentation performance and these are currently perceived as too onerous. Lyons, Brodsky [5] et alproposed that the automatic detection of surveillance related events from video input is an important technology in CCTV security applications. ZhongJi, Yuting Su [6] et alsuggested that in order to browse the video data quickly and make full use of the surveillance video data, they propose a fully automatic and computationally efficient framework for analysis and summarization of surveillance videos with the techniques of moving object detection and trajectory extraction.Xiaohua He and JianLing[8] suggested a videosummarization method based on the TemporallyMaximum Occurrence Frame (TMOF).Aydemir,Ergul[9] et al proposed a novel approach forthe summarization problem which is based on humanmovement understanding. Proposed system presentsefficient video knowledge extraction, especially forsurveillance cases. Yang,Dadgostar [10] et al proposed two novel techniques for automaticsummarisation of lengthy surveillance videos, based on selectionof frames containing scenes most informative for rapid perusaland interpretation by humans.Almeida, Torres [11] et al suggested that themethod is based on bothexploiting visual features extracted from the video stream andon using a simple and fast algorithm to summarize the videocontent.Panchal, Merchant [12] et al introduce a proposed method to detect scene based on motionvector and occurrence rate of shot boundaries in video. Inwhich, motion vectors weights and directions addresses exactscene of video with prior action on occurrence rate of shotboundaries of video, it is possible to differentiate two scenes in amovie or video. Searching and access of scene of video from large database, it isa user requirement to have proper indexing, browsing. Hence, itbecomes a challenging problem of searching and accessing of relevant scene of end user, trivially.

3. PERCEPTUAL HASHING
P-Hash algorithm uses the features of images to generate distinct but not unique fingerprints. These fingerprints can be compared with each other to find similarities or dissimilarities between the two images. The steps for perceprual hashing are as follows:

1. **Reduce size -** Like Average Hash, pHash starts with a small image. However, the image is larger than 8x8; 32x32 is a good size.

2. **Reduce color -** The image is reduced to a grayscale just to further simplify the number of computations.

3. **Compute the DCT -** The DCT separates the image into a collection of frequencies and scalars. While JPEG uses an 8x8 DCT, this algorithm uses a 32x32 DCT.

4. **Reduce the DCT -** While the DCT is 32x32, just keep the top-left 8x8. Those represent the lowest frequencies in the picture.

5. **Compute the average value -** Like the Average Hash, compute the mean DCT value (using only the 8x8 DCT low-frequency values and excluding the first term since the DC coefficient can be significantly different from the other values and will throw off the average).

6. **Further reduce the DCT -** Set the 64 hash bits to 0 or 1 depending on whether each of the 64 DCT values is above or below the average value.

7. **Construct the hash -** Set the 64 bits into a 64-bit integer. The order does not matter, just as long as you are consistent.

The Detailed Description of this algorithm can be found in [7].

4. PROPOSED ALGORITHM FOR VIDEO SEGMENTATION AND SUMMARIZATION

In the proposed algorithm, a large video is being taken as an input and a summarized video is created as an output. Steps for the algorithm are as follows:

1. Input Video.
2. Create frames from the input video after every constant interval of the time till video ends.
3. Compare every consecutive frame with each other using Perceptual Hashing algorithm to detect change.
4. Calculate the time where change is detected as in format of start time and end time.
5. Create Segments (small clips) form input video using calculated start time and end time of change detected.
6. Combine all Segments (small clips) to crate summarized video as an output.

5. EXPERIMENTAL RESULTS

The above experiment based on the proposed approach has been achieved on DELL 2nd Generation core I-5 processor having 4GB of RAM. The used language is JAVA (JDK 7) in Windows 8 (64 Bit).
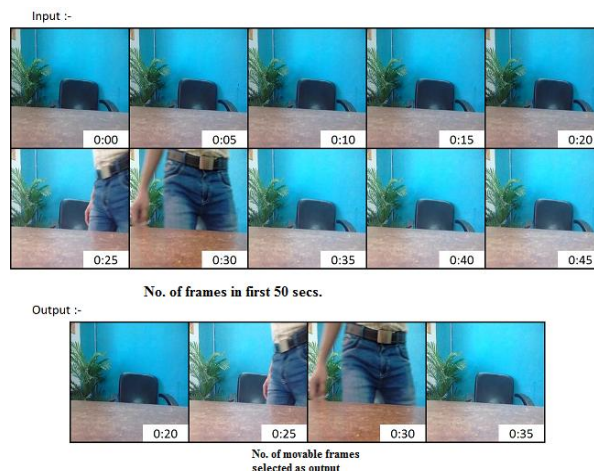

Fig.1: Experimental Result (I/O)

In the above picture we have described how the framing takes place. We have taken a video of 50 seconds and created 10 frames with a time interval of 5 seconds each. Further the frames which have some movement are extracted by using DCT algorithm and Xuggler API.

Table 1: I/O Frames at fixed time-intervals

Time (Mins.)	Input (No. of frames)	Output (No. of movable frames)
1	20	3
2	40	6
4	80	12
8	160	24
16	320	48
32	640	96

Video Frame Rate = 30 frames/sec.
Sampling Rate = 5 seconds.

The above table shows several time intervals (at 1min, 2 min, 4 min and so on) the number of frames generated is given as input and number of movable frames as output. On the basis of the above table we can further depict these input-output in the graphical form -
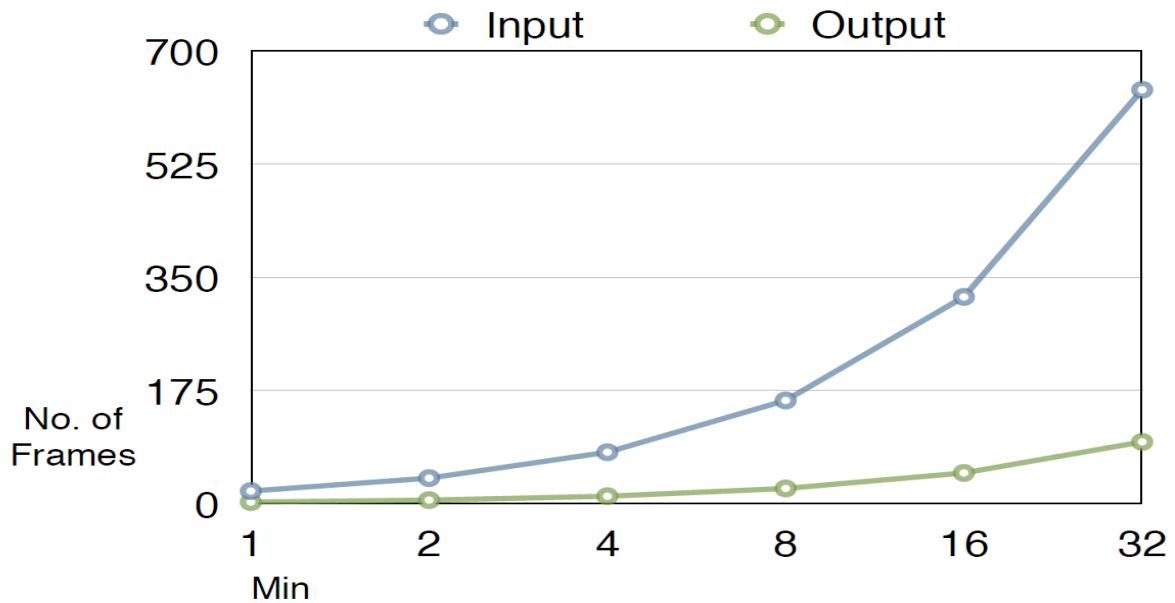
Fig.2: Frame Comparison (I/O)

6. CONCLUSION

This paper presents an overview of segmentation and summarization of a video through merging segments of different events. By the use of algorithm the wastage of resources like memory for video storage can be reduced. The summarized videos will be more meaningful and can fulfill the purpose of video surveillance. The algorithm can be used in real time scenario i.e. at the time of video recording to get summarised video. High level semantic information of the video to make better video summaries can be employed further.

7. REFERENCES

[1] VijayKatkar andAmitBarve; *"Efficient Content Based Video Retrieval for Animal Videos"*at the international conference on advances in computers, communication and control; Mumbai January 2011.

[2] Murali S and Girisha R*"Adaptive Object Segmentation from Surveillance Video Sequences"*; at 2010 First International Conference on Integrated Intelligent Computing; at Karnataka.

[3] Tao Xiang∗ and Shaogang Gong;*"Activity Based Surveillance Video Content Modelling"* at University of London, London.

[4] Kevin McKoen, Raquel Navarro-Prieto, Benoit Duc, EmrullahDurucan, Francesco Ziliani,TouradjEbrahimi; *"Evaluation Of Video Segmenattion Methods For Surveillance Applicatios";* at Geneva, Switzerland.

[5] D.M. Lyons, T. Brodsky, E. Cohen-Solal and A. ElgammalPhilips; *"Video Content Analysis for Surveillance Applications"*at NewYork USA.

[6] ZhongJi, Yuting Su, RongrongQian, Jintao; *"Surveillance video summarization based on moving object detection and trajectory extraction"*; at Tianjin, China.

[7] www.phash.org

[8] Xiaohua He and Jian Ling; *"A Video Summarization Method Based on Key Frames Extracted by TMOF";*at Hangzhou, China.

[9] M. Said Aydemir, UgurErgul, AdemGuclu, M. ElifKarsligil; *"Video Summarization Using Simple Action Patterns";* at 21st International Conference on Pattern Recognition (ICPR 2012) November 11-15, 2012. Tsukuba, Japan.

[10] Yan Yang, FarhadDadgostar, Conrad Sanderson, Brian C. Lovell*; "Summarisation of Surveillance Videos by Key-frame Selection"*; at the University of Queensland, School of ITEE, Australia.

[11] Jurandy Almeida, Ricardo da S. Torres, and Neucimar J. Leite; *"Rapid Video Summarization on Compressed Video"*;at 2010 IEEE International Symposium on Multimedia; Institute of Computing, University of Campinas – UNICAMP, Brazil.

[12] PradipPanchal, Sabbir Merchant and Nirav Patel; *"Scene detection and retrieval of video using motion vector and occurrence rate of shot boundaries"*; at2012 NirmaUniversity International Conference On Engineering, NUiCONE-2012, 06-08DECEMBER, 2012.