Surveillance Video Synopsis Techniques : A Review

Shefali Gandhi
Assistant Professor, Dharmsinh Desai University, Nadiad
and
Tushar Ratanpara
Assistant Professor, Faculty of Technology, Dharmsinh Desai University, Nadiad

This is the era of video surveillance, not just security. The arrival of inexpensive surveillance cameras and increasing demands of security has caused an explosive growth of surveillance videos, which are used by government or other organizations for prevention or investigation of crime. As browsing such lengthy videos is very time consuming, most of the videos are never watched and analyzed. The video synopsis is a technique to represent such lengthy videos in a condensed way by showing multiple activities simultaneously. The purpose of this paper is to explore development stages, various algorithms of it, framework and tools used to implement them, challenges and limitations of existing video synopsis techniques and its application in the field of surveillance video analysis.

Keywords: video synopsis, video abstraction, surveillance video, video summary, video indexing, background modeling, object stitching.

1. INTRODUCTION

Hampapur et al. [2005] stated that in the recent years, due to increased demand of ensuring high level security at public places like, airports, railway stations, banks, parking lots etc., a huge number of surveillance cameras are mounted all over the world and enormous volume of data is captured day and night. The development of smart cities brings in more responsibilities too and video surveillance and its investigation, for quick decision making, is one of them 1. The task of analyzing such large videos is a waste of manpower and very time consuming too. Sometimes, the important content might be missed by the viewer. Moreover, storage and retrieval of such infinite video sequences is also very difficult. The user has to browse through a long video to locate the salient parts of it. This led to the need to develop video summarization techniques for faster browsing of videos.

2. LITERATURE REVIEW

Li et al. [2001], Li et al. [2006] and Truong and Venkatesh [2007] explained about Video abstraction, the mechanism to produce a short summary of a long video and represent it in a compact way. Kansagara et al. [2014] classified video abstracts in two categories: Still image abstracts and moving image abstracts. The collection of key images extracted from the long video sequence is called still-image abstract or summary of video. The moving-image abstract or video skim, consists of a collection of image sequences, as well as the corresponding audio abstract extracted from the original sequence. [1] (normal) In video summary, generally only visual information is considered. The skimming is also classified in two types: Summary sequence, which provides users an impression about the whole video content and highlight, which includes only the most salient parts of the original video. Benjamin et al. [2010] outlined the approach of adaptive fast forwarding based on information in it. It allows browsing the video data quickly by skipping

[1] 1http://www.varindia.com/news/ip-based-cameras-and-video-analytics-set-to-take-the-next-leapsthash.eQbhvdnr.dpbs

individual frames or groups of frames at fixed or adaptive intervals. It is simple but removes complete frames from the video which may result in ignoring important information and ratio of video condensation is relatively low. The Space-time Video Montage is another video summarization approach suggested by Kang et al. [2006], where both the spatial and temporal information of a video stream are simultaneously analyzed and packed together. Pritch et al. [2007] and Pritch et al. [2008] have presented an approach to create temporally compact representation of the long surveillance video, called video synopsis. It creates a short summary of long input video while preserving essential activities of the source video by simultaneously showing multiple events occurred at different timestamps in source video by keeping the spatial details intact (See fig 1). Video synopsis itself is a video and it reduces spatial as well as temporal redundancies which may change the chronological order of the events. It addresses user queries like Show synopsis of the input video of the past 10 hours in 5 minutes. It can be used as a powerful tool by government or private security agencies to quickly browse a large number of surveillance videos to speed up case analysis and saving a lot of time and human resources cost. It can also be used for remote monitoring of home for a child or elder care, in which, video synopsis of CCTV camera videos are sent to cloud servers via cellular wireless networks to the users who can view it on their smart equipment, like smart phones and tablets. It is an effective tool for browsing and indexing long surveillance videos. Zhang et al. [2014] and Yogameena and Sindhu [2015] outlined the method to detect human crowd anomalies or abnormal activities from the synopsis of the video. The surveillance video synopsis is further extended for automatic ranking of videos based on their content by Shah and Ratanpara [2017].

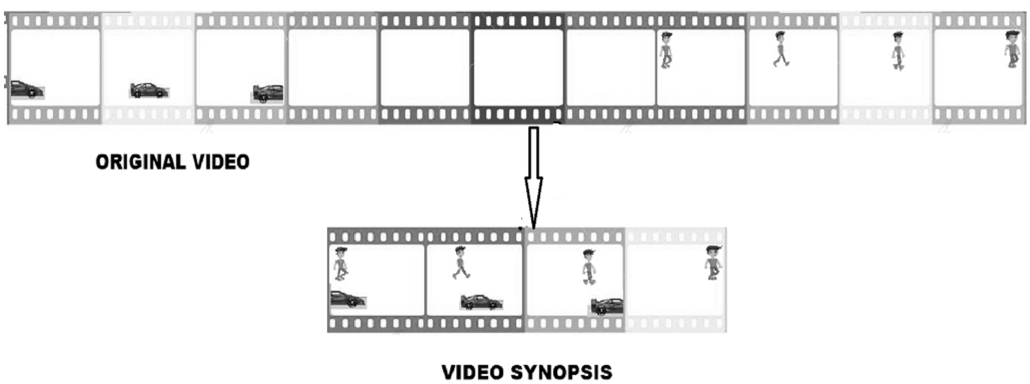

Figure 1. Video Synopsis

3. VIDEO SYNOPSIS PROCESS

The video synopsis can be broadly classified in to two types: Region-based and object-based. In region based approach, synopsis of a video is produced using Markov Random Fields optimization. But it is complex and results in high computational cost. In order to overcome these limitations, Pritch et al. [2008] presented an object-based approach where energy minimization is employed on higher level cost functions.

The video synopsis framework consists of five basic stages: Background modeling and moving objects detection, Space-time activity tube generation, temporal shifting optimization, time-lapse background generation and seamless stitching of objects to the background. The system to generate synopsis of an endless video (webcam or online synopsis) is divided in two phases. During the first online phase, the moving objects are detected and added to an object queue. And time-lapse video generation, selection of tubes to appear in synopsis and object stitching are done offline during the response phase. The hierarchical classification of different approaches of video synopsis generation is shown in figure 2.

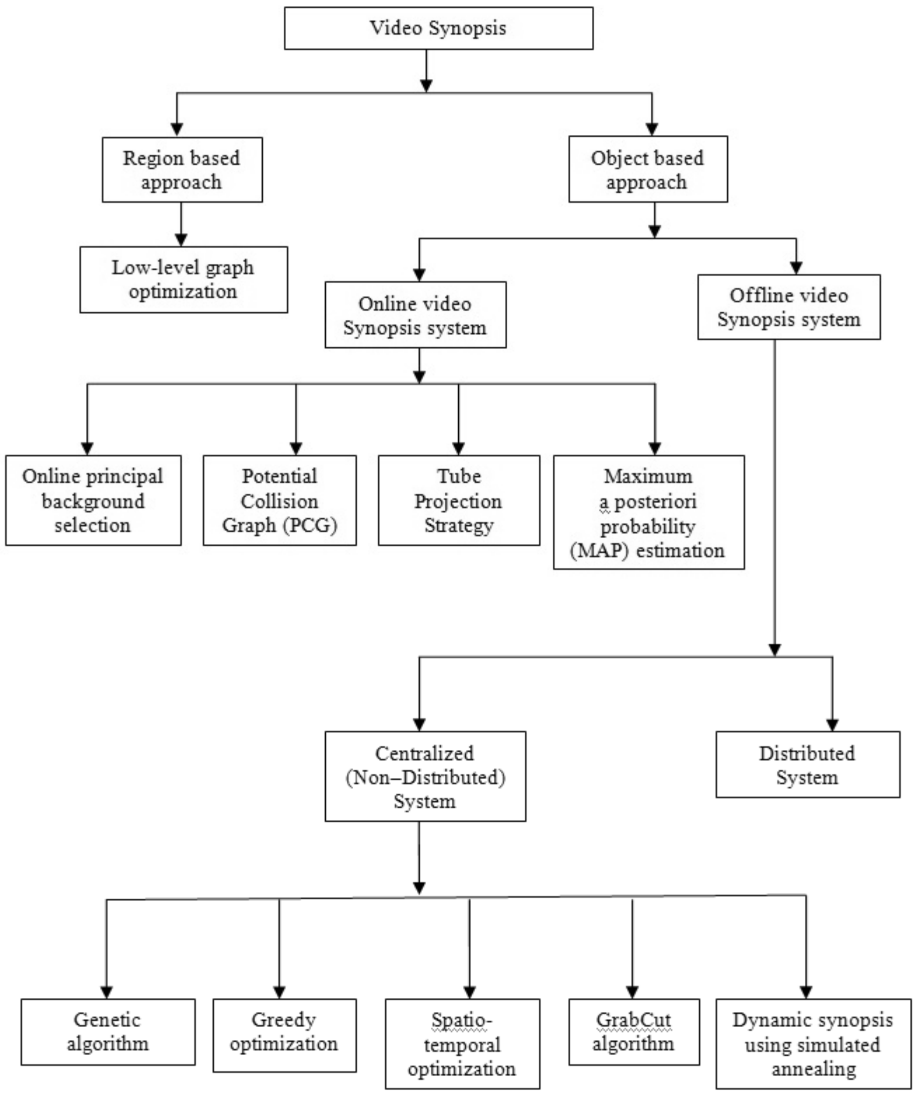

Figure 2. Hierarchical classification of video synopsis approaches

3.1 Background Modeling And Moving Objects Detection

In order to build effective video synopsis, moving objects and events should be detected from the input video. The detection of moving object starts with background modeling. The first frame of a video or temporal median over the entire clip can be considered as background in short videos. But background changes over time in case of surveillance videos due to variation in illumination (sunlight, clouds, shadows, etc.), day night transition, changes in background objects, etc. Considering these issues, Pritch et al. [2008] and Gandhi and Ratanpara [2017] suggested temporal median over few minutes before and after the each frame to construct the dynamic background. In Feng et al. [2010], Hsia and Chiang [2013], Zhong et al. [2014], and Chou et al. [2015], background is estimated using GMM(Gaussian Mixture Model), while Yao et al. [2014] used the frame difference method in order to model background, where frames before and after the current frame are subtracted to generate frame difference images. These frame difference images are then accumulated from average. Hassanpour et al. [2011] lists many other statistical and non-statistical approaches to model background of a video. Nie and Xiao [2013] has elaborated the simplest way to segment foreground objects by computing frame difference pixel by pixel and then applying 2D morphological operations to get larger foreground masks. Background cut (Sun et al. [2006]) can be used to detect objects more precisely but it is computationally expensive. The background subtraction algorithm is proposed by Huang and Huang [2014], Chen et al. [2007], Bagheri and Zheng [2014], and Chou et al. [2015], as an efficient method to extract foreground. Background subtraction can be combined with min-cut for smooth extraction of foreground (Huang and Huang [2014], Chen et al. [2007], Bagheri and Zheng [2014], and Chou et al. [2015]). Sometimes shadows of the moving objects can affect the correct detection of moving objects. Lu et al. [2013] presents a novel approach combining GMM and LBP feature based method to get more compact foreground without shadows. While Hao et al. [2013] employed a Graph cut algorithm based approach, called Grabcut, as an effective object segmentation technique, which considers color and border information of the frame. Zhu et al. [2014] explains Visual Background Extractor (ViBe) method to model background and extract foreground. Zhong et al. [2014] presented a 3D graph cut based approach for segmentation of moving objects. It assigns binary label to each processing unit, as foreground or background. In Zhu and Liao [2015] for joint synopsis of multiple camera network, background subtraction based on Scale Invariant Local Ternary Pattern is employed. In order to generate more efficient and condensed synopsis, Li et al. [2016] presented an optical flow based approach for extracting moving objects. And when seam carving is applied on the extracted tubes, only fast moving pixels are preserved and pixels with low motion are removed.

3.2 Space-Time Activity Tubes Generation

The parts of the detected objects are connected together and tracked to generate space-time activity tubes. The tracking of objects is very important and error prone because of noise, collision, occlusion, especially tracking multiple interacting objects. All the detected foreground object masks are connected to generate 3D tubes. The speed at which objects are moving may vary and objects may occlude. The real-time tracking with shadow removal based on the real time robust tracker approach is explained by Kaewtrakulpong and Bowden [2001] and Nie and Xiao [2013] to overcome these issues. If one object is segmented into number of parts and not tracked properly, the flickering of moving objects can be there in the synopsis video. Lu et al. [2013] solves this issue using particle filtering algorithm that produces fluent tubes. Yao et al. [2014] and Gandhi and Ratanpara [2017] tracks the moving objects using Kalman filter. It first predicts the position of moving object by considering its location in previous frame and then determines the matching model using size, position, area, velocity and acceleration to decide the matching sequence of moving object. If the objects are interacting with each other, kalman filter loses its effectiveness. The issue can be handled by particle filter.

Zhu et al. [2014] explains the approach for generating video synopsis based on tracking of multiple objects and it combines Kalman filter with particle filter for the purpose of tracking moving objects. Parekh et al. [2014] explains other kernel based tracking methods like, simple template matching, mean shift method, layer based tracking to track moving objects in a video. Examples of extracted tubes are shown in the figure 3 below, where fig. 3(a) shows a tube of a moving person from video captured by indoor surveillance camera of a billiard club (Pritch et al. [2008]) and (b) shows a moving person tube from an outdoor surveillance video (Gandhi and Ratanpara [2017]).

3.3 Temporal Shifting Optimization

The synopsis of a video should be considerably shorter than the input video and should cover maximum events from the original video without any visible seams or fragmentation of objects. Moreover, the collisions among the objects should be least and chronological order of events should be maintained as much as possible. In order to achieve these properties, Rav-acha et al. [2006] andPritch et al. [2008] explained the optimal temporal shifting of tubes, which can be obtained by minimizing an energy function consisting of following cost terms:

(1) Collision cost: It is a penalty imposed when two or more objects collide in the synopsis video and it is computed between each pair of tubes.
(2) Activity cost: If an object from original video is not mapped at all or partially mapped in the synopsis video, activity cost is imposed.
(3) Chronological disorder cost: It encodes the cost imposed when the chronological order of the events in original video is changed in the synopsis video for all pair of tubes.
(4) Background consistency cost: It represents the cost of not stitching tube to its original background in the synopsis video.

The simplest approach to minimize the collision cost is to keep shifting the tubes repeatedly until there is acceptable or no collision at all. In Pritch et al. [2008], the Markov Random Field(MRF) based techniques is used to minimize energy function. A greedy optimization technique also works well to minimize the energy function. In Nie and Xiao [2013] alpha-beta swap graph cut method is employed to obtain global spatiotemporal optimization for generating compact video synopsis. A graph is constructed first, where each node represents corresponding activity tube. Each node is labeled as per the activity cost and edge connecting two nodes is labeled considering collision cost and spatiotemporal consistency cost. Minimum cut on graph minimizes energy function.

Figure 3. Sample tubes (a) A billiard club video (Pritch et al. [2008]) (b) An outdoor surveillance video (Gandhi and Ratanpara [2017])

Gandhi and Ratanpara [2017] presented a novel approach using Genetic Algorithm to optimize energy function, which is proved to be less time consuming and more efficient compared to previous approaches. In the aforementioned approaches rearrangement of the tubes is achieved by minimizing global energy function iteratively, which is time consuming and memory demanding in case of large videos. Because of this reason, the task of tubes rearrangement is generally done offline and sometimes by breaking long video into parts.

Jin et al. [2016] proposed a low-complexity and high efficiency method of online video synopsis. In this algorithm, rearrangement of tubes is done with the help of projection matrix which is updated frequently to avoid collisions of tubes. It requires the collision costs to be computed only once. The another approach is presented by He et al. [2017] for rearranging tubes online by constructing a Potential Collision Graph based on two types of relationships, Collision Potential and Collision Free.

The rearrangement of tubes in the temporal domain sometimes destroys interaction among the objects which may possess vital information like thefts and accidents. Recently, the approach proposed by Badal et al. [2017] resolves this issue by merging interactive tubes and considering them as a single tube. In order to achieve high computing speed, Lin et al. [2017] employed distributed processing model for it. The original video is split into segments and distributed on different processors to increase computing speed.

3.4 Time-Lapse Background Generation

The time-lapse background is required to be generated before stitching activity tubes on it to obtain final synopsis video. It should represent the changes in background over the time. Feng et al. [2010] proposed online principal background selection(OPBS) method to generate time-lapse background for building efficient online video synopsis systems. In this approach, n principal backgrounds are selected from N backgrounds. The background selection is done in an online fashion to reduce memory cost. The background selection not only considers changes in background over time but also considers video activities. It presents two algorithms of OPBS, one is without prediction and another is with prediction. The OPBS with prediction selects more backgrounds from highly active time periods but does not ignore backgrounds from time durations having less activity. As explained in Pritch et al. [2008], another approach to generate time-lapse background is based on two histograms: a uniform activity histogram and a temporal activity histogram. The third histogram is obtained by doing interpolation of these two histograms and background frames are selected based on this interpolated histogram. The method selects more frames from active time period and a few frames from inactive time period as well. Some of the surveillance videos have a narrow path of motion and large area of background is free, having no active object moving in it. Temporally condensed synopsis of such video may result in number of collisions and undesirable artifacts. Nie and Xiao [2013] presented an approach to construct large virtual background motion space to reduce the number of collisions effectively, making synopsis video look more realistic. Cho et al. [2008] and Yogameena and Sindhu [2015] employed patch relocation and multilevel relocation methods to achieve this goal. Patch relocation is also used to generate compact background by relocating patches of extracted background.

3.5 Object Stitching

In order to construct final video synopsis, tubes from different time periods are stitched on the time-lapse background images as per the new shifted time periods, obtained in optimization stage. Pe'rez et al. [2003] explains Poisson image editing, an approach to stitch objects on the background. It is very effective algorithm but time consuming because of the Poisson equation. Tanaka et al. [2012] presents an alternate approach based on an approximated solution of the Poisson equation, which is computationally more effective and faster. Alternatively, a coordinate-based method (Farbman et al. [2009]) can be used for seamless image cloning. This approach is based on MVC (mean value coordinates), which is beneficial in terms of computational speed and is easy to implement.

A typical frame from synopsis of a lengthy video is shown in the figure 3, in which activities occurring at different time stamps can be observed in a single frame (Gandhi and Ratanpara [2017]). While fig. 5 shows three frames from a video captured over two days by an indoor camera placed in a billiard club, along with a frame from its video synopsis of 1 min (citeNyz).

4. TOOLS AND FRAMEWORKS USED

Gandhi and Ratanpara [2017] used Matlab 2013a for the implementation of proposed algorithm for video synopsis. Three video sequences are used to evaluate the work, one is from ViSOR data repository and remaining two are captured using real outdoor surveillance system. In order to evaluate video synopsis process, Yogameena and Sindhu [2015] used Matlab 10 and UMN test datasets. Nie and Xiao [2013] implemented their proposed algorithm for generating compact video synopsis in C++ and executed it on Intel core 2 duo 2.1 GHz processor with 2GB RAM.

Figure 4. A typical frame from synopsis of a lengthy video (Gandhi and Ratanpara [2017])

Figure 5. Sample frames from long video and a frame from its video synopsis (yz)

While Lin et al. [2017] tested their distributed system for generating video synopsis on C++ in visual studio 2008 with OpenCV 2.4.9 librabry.

5. COMPARATIVE REVIEW

Table 1 shows the comparative review of related work done in the field of video synopsis.

6. CHALLENGES

Even though video synopsis is very useful technique for the quick browsing of surveillance videos, it still has following challenges:

(1) If the original video itself is already very dense, like the surveillance videos captured at busy

Table I: Comparision of Video Synopsis Technologies

Author and Title	Main feature	Foreground Extraction	Performance	DataSet
Pritch et al.,[2009] Clustered Synopsis of Surveillance Video	Clustered,synopsis based on similar activities using SVM classifier.	Background, subtraction with, min-cut.	70% objects labeled correctly in unsupervised clustering and 93.6% using SVM.	Training set of 100 tubes.
Feng et al.,[2010], Online Principal Background Selection for Video Synopsis"	A technique to select principal backgrounds in online fashion to eliminate the need to store all backgrounds.	GMM based background modeling and Graph cut algorithm for foreground extraction.	Total processing time for OPBS with prediction on 24 hours video is about 562 secs, i.e., 0.43ms per frame.	Artificial data,of 4000 frames and Read Surveillance video of 24 hours (15 fps)
Hsia et al. [2013], A Complexity Reduction Method for Video Synopsis System	A range tree,based algorithm to browse and search object of interest quickly in the synopsis.	GMM based background subtraction algorithm.	Accuracy rate of 97% and processing speed 32 fps.	Video sequence of 2,659 frames with 32.42 fps
Nie et al. [2013], Compact Video Synopsis via Global Spatiotemporal Optimization	An approach to create synthesized compact background that provides larger virtual motion space and alpha-beta swap graph cuts method is employed for energy,minimization.	Temporal median over a short period to model background and frame difference for extracting foreground.	The proposed method eliminates collisions and creates more condensed synopsis video.	Five input videos of 10mins, 3mins, 30mins, 5mins & 10mins (30 fps)
Lu et al. [2013], Generating fluent tubes in video synopsis	Kalman filter and particle filter are combined for moving objects tracking and so, more fluent tubes are generated.	GMM and a texture based methods are combined for compact foreground detection.	Number of object detection is improved by 20.6% & 23.3% in playground & road clips respectively.	Four videos of Playground (11410 frames), Corridor (1078 frames), Road1(2716 frames), Road2 (2431 frames)
Chou et al. [2015], Coherent Event-Based Surveillance Video Synopsis Using Trajectory Clustering	Coherent events containing similar actions of objects with different moving speeds are combined using longest common subsequence algorithm.	GMM for background modeling and Background Subtraction algorithm for detecting foreground pixels.	93% to 96% saving of space.	30 minutes surveillance videos of 4 indoor scenes.
Zhong et al. [2014], Fast Synopsis for Moving Objects Using Compressed Video	A new graph cut algo. to extract object tubes which gives a fast solution to minimize energy function in compressed domain.	GMM based background modelling and 3D graph cuts algorithm for fine extraction of moving objects.	58.79% and 67.10% time-saving.	Standard sequences named Hall Monitor Daytime and F-building.
Zhu et al. [2015], Multi-Camera Joint Video Synopsis	A novel multi-camera joint video synopsis algorithm for multi camera surveillance video.	Averaging frames in a fixed time interval to produce background model and SLITP method to detect foreground pixels.	JVS method has less Chronological Disorder Time Stamp Pairs between different camera views.	One outdoor dataset (2 cam-2664 frames) and two outdoor/indoor datasets (3 cam-6000 frames and 5 cam-5000 frames)
Yogameena et al.[2015], Synoptic Video Based Human Crowd Behavior Analysis for Forensic Video Surveillance	A technique to model human blobs in crowd to detect anomalies in the crowd in surveillance videos.	GMM based background subtraction and three-frame temporal differencing is used.	60% less time is taken to detect abnormalities in lengthy videos than existing methods.	Three UMN datasets with 1450, 2145 and 4415 frames (30fps)

railway station or bus station, where all the areas are active all the time, then generating a comparatively shorter video synopsis becomes difficult.

(2) Sometimes the synopsis video may confuse the viewer by simultaneously displaying multiple activities.

(3) The process of generating video synopsis is still limited to the videos captured by stationary camera.

7. CONCLUSION

This paper presents a conceptual framework for the process of video synopsis. The process consists of five stages: Background modeling and moving object detection, Space-time activity tube generation, Temporal shifting optimization, Time-lapse background generation and object stitching. It reviews and exploits the existing developments and different types of algorithms used in each stage of the process. The state-of-the-art of existing methods at each stage and their key issues are discussed as well. It also includes comparative review of the work done the field of video synopsis in form of a table and hierarchical classification of different techniques of video synopsis process. The framework and tools used to implement the system are also mentioned. This article provides valuable insight into the challenges in generating video synopsis and encourages new research in this area of computer vision.

8. FUTURE WORK

The reviews in this paper address only videos captured by stationary cameras. Video synopsis should also consider the videos captured by moving cameras. This can be achieved by algorithms to compute motion of camera while tracking objects. The synopsis of video can also be provided as input to the system used for face recognition, detection of abnormal activities, etc.

References

BADAL, T., NAIN, N., AND AHMED, M. 2017. Surveillance video synopsis while preserving object motion structure and interaction. In *In proceedings of international conference on computer vision and image processing*. pp.197–207.

BAGHERI, S. AND ZHENG, J. 2014. Temporal mapping of surveillance video. In *22nd International Conference on Pattern Recognition*. pp.4128–4133.

BENJAMIN, P., MARKUS, H., DANIEL, W., AND GUNTHER, H. 2010. Information-based adaptive fast forward for visual surveillance. *Multimedia Tools and Applications*.

CHEN, Y., CHEN, C., HUANG, C., AND HUNG, Y. 2007. Efficient hierarchical method for background subtraction. *Pattern Recognit. Vol.40*.

CHO, T. S., BUTMAN, M., AVIDAN, S., AND FREEMAN, W. 2008. The patch transform and its applications to image editing. In *IEEE conference on computer vision and pattern recognition*. pp.1–8.

CHOU, C., LIN, C., CHIANG, T., CHEN, H., AND LEE, S. 2015. Event-based surveillance video synopsis using trajectory clustering. In *IEEE international conference on multimedia expo workshop*. pp.1–5.

FARBMAN, Z., HOFFER, G., LIPMAN, Y.AND COHEN-OR, D., AND LISCHINSKI, D. 2009. Coordinates for instant image cloning. In *ACM transactions on graphics (TOG) - proceedings of ACM*. pp.67.

FENG, S., LIAO, S., YUAN, Z., AND LI., S. Z. 2010. Online principal background selection for video synopsis. In *International conference on pattern recognition*. pp.17–20.

GANDHI, S. AND RATANPARA, T. 2017. Object-based surveillance video synopsis using genetic algorithm. *Applied Video Processing in Surveillance and Monitoring Systems*.

HAMPAPUR, A.AND BROWN, L., CONNELL, J., EKIN, A., HAAS, N., LU, M., MERKL, H.AND PANKANTI, S., SENIOR, A., SHU, C., AND TIAN, Y. 2005. Smart video surveillance. *IEEE signal processing magazine*.

HAO, L., CAO, J., AND LI, C. 2013. Research of grabcut algorithm for single camera video synopsis. In *IEEE fourth international conference on Intelligent control and information processing*. pp.632–637.

HASSANPOUR, H., SEDIGHI, M., AND MANASHTY, A. R. 2011. Video frames background modeling: Reviewing the techniques. *Journal of Signal and Information Processing*.

HE, Y., QU, Z., GAO, C., AND SANG, N. 2017. Fast online video synopsis based on potential collision graph. *IEEE Signal Processing Letters*.

HSIA, C. AND CHIANG, J. 2013. A complexity reduction method for video synopsis system. In *International symposium on intelligent signal processing and communications systems*. pp.163–168.

HUANG, C., C. P. Y. D. C. H. AND HUANG, G. 2014. Maximum a posteriori probability estimation for online surveillance video synopsis. In *IEEE transactions on circuits and systems for video technology*. pp.1417–1419.

JIN, J., LI, F., GAN, Z., AND CUI, Z. 2016. Online video synopsis method through simple tube projection strategy. In *Wireless communications signal processing (wcsp), 2016 8th international conference*. pp.1–5.

KAEWTRAKULPONG, P. AND BOWDEN, R. 2001. An improved adaptive background mixture model for real-time tracking with shadow detection. *Video-based surveillance systems*.

KANG, H., MATSUSHITA, Y., TANG, X., AND CHEN, X. 2006. Spacetime video montage. In *IEEE computer society conference on computer vision and pattern recognition*. pp.1331–1338.

KANSAGARA, R., THAKORE, D., AND JOSHI, M. 2014. A study on video summarization techniques. *International journal of innovative research in computer and communication engineering Vol. 2*.

LI, K., YAN, B., WANG, W., AND GHARAVI, H. 2016. An effective video synopsis approach with seam carving. *IEEE Signal Processing Letters*.

Li, Y., Lee, S., Yeh, C., and Kuo, C. 2006. Techniques for movie content analysis and skimming: tutorial and overview on video abstraction techniques. *IEEE Signal Processing Magazine*.

Li, Y., Zhang, T., and Tretter, D. 2001. An overview of video abstraction techniques. Technical Report HPL.

Lin, L., Lin, W., Xiao, W., and Huang, S. 2017. An optimized video synopsis algorithm and its distributed processing model. *Soft computing*.

Lu, M., Wang, Y., and Pan, G. 2013. Generating fluent tubes in video synopsis. In *IEEE international conference on acoustics, speech and signal processing*. pp.2292–2296.

Nie, Y. and Xiao, C. 2013. Compact video synopsis via global spatiotemporal optimization. In *IEEE transactions on visualization and computer graphics*. pp.1664–1676.

Parekh, H. S., Thakore, D. G., and Jaliya, U. K. 2014. A survey on object detection and tracking methods. *International Journal of Innovative Research in Computer and Communication Engineering Vol.2*.

Pe'rez, P., Gangnet, M., and Blake, A. 2003. Poisson image editing. In *ACM Trans. Graphics (TOG)*. pp.313–318.

Pritch, Y., Rav-acha, A., Gutman, A., and Peleg, S. 2007. Webcam synopsis: Peeking around the world. In *IEEE 11th international conference on computer vision*. pp.1–8.

Pritch, Y., Rav-acha, A., and Peleg, S. 2008. Nonchronological video synopsis and indexing. In *IEEE transactions on pattern analysis and machine intelligence*. pp.1971–1984.

Rav-acha, A., Pritch, Y., and Peleg, S. 2006. Making a long video short: Dynamic video synopsis. In *Proc. IEEE conf. computer vision and pattern recognition*. pp.435–441.

Shah, R. and Ratanpara, T. 2017. An automatic content based video ranking from surveillance videos. *International journal of engineering and technology Vol.9*.

Sun, J., Zhang, W., Tang, X., and Shum, H. 2006. Background cut. In *Proc. European Conference Computer Vision (ECCV 06)*. pp.628–641.

Tanaka, M., Kamio, R., and Okutomi, M. 2012. Seamless image cloning by a closed form solution of a modified poisson problem. *SIGGRAPH Asia 2012*.

Truong, B. and Venkatesh, S. 2007. Video abstraction: A systematic review, classification. In *ACM transactions on multimedial computing, communications and applications*. pp.3–3.

Yao, T., Xiao, M., Ma, C., Shen, C., and Li, P. 2014. Object based video synopsis,. In *IEEE workshop on advanced research and technology in industry applications*. pp.1138–1148.

Yogameena, B. and Sindhu, P. 2015. Synoptic video based human crowd behavior analysis for forensic video surveillance. In *IEEE eight international conference on advances in pattern recognition*. pp.1–6.

Zhang, Y., Lin, W., Zhang, G., Luo, C., Jiang, D., and Yao, C. 2014. A new approach for extracting and summarizing abnormal activities in surveillance videos. In *IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. pp.1–5.

Zhong, R., Hu, R., and Wang, Z. 2014. Fast synopsis for moving objects using compressed video. *IEEE signal processing letters Vol.21*.

Zhu, B., Liu, W., and Yuan, G. 2014. A method for video synopsis based on multiple object tracking. In *5th IEEE international conference on software engineering and service science*. pp.414–418.

Zhu, J. and Liao, S. 2015. Multi-camera joint video synopsis. *IEEE transactions on circuits and systems for video technology Vol.PP*.

Ms Shefali Gandhi is working as assistant professor at Dharmsinh Desai University, Nadiad, Gujarat, India. She has academic experience of more than 10 years. She has done significant work in the field of image and video processing, especially surveillance videos. Ms. Gandhi pursued her engineering at Dharmsinh Desai University, from where she has BE (Computer engineering) and MTech (Computer engineering) degrees. Her research interests are in video processing, operating system and computer networks. A chapter entitled "Object-Based Surveillance Video Synopsis Using Genetic Algorithm", based on her research work on surveillance video synopsis is published in research handbook "Applied Video Processing in Surveillance and Monitoring Systems" by IGI Global.

Mr Tushar Ratanpara is working in the computer engineering department at Dharmsinh Desai University, Nadiad, India. He has published more than 15 research papers in the international journals and conferences like IEEE, Springer, IGI-Global, Inderscience and ACM. Mr Ratanpara has persued his engineering studies from Dharmsinh Desai University, Nadiad. His main research areas includes Image Processing, Video Processing, Audio Processing and Internet of Things.