*International Journal of Engineering &Technology*, 7 (4) (2018) 7063-7067

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.20422
***Research paper***

# An advanced approach for distortionless seamcarving in video analysis

**G. Thirumalaiah [1] *, S. Immanuel Alex Pandian [2]**

*[1] Research Scholar, Department of ECE, Karunya University, Coimbatore, Tamil Nadu, India*
*[2] Assistant professor, Department of ECE, Karunya University, Coimbatore, Tamil Nadu, India*
**Corresponding author E-mail: tiru5502@gmail.com*

**Abstract**

Video synopsis is a technique that creates summary of the video or it can converts the abstraction of selected frames. This innovative approach permits the organizations to review long videos in minutes. It is more convenient to superimpose objects on to the static background and parallel displaying events to generate video synopsis. In this, paper a propelled noiseless video synopsis technique, which utilizes object-extracting method for vital objects. Along this technique spatial and temporal coherence cost is used to maintain time and position of the important objects. The proposed method will generate video spots and seam craving method to reduce the input (original) video. Finally, experimental results gives that our proposed method can produce a large reducing ratio, while preserving all the important objects of choice. Therefore, this noiseless approach can facilitate users to watch the surveillance video with greater accuracy.

***Keywords***: *Seam Carving; Spatial Coherence; Temporal Coherence; Video Analysis.*

## 1. Introduction

In the last decade a lot of police work cameras have been deployed and employed in transportation hubs, ATMs and many alternative public or non-public facilities. Thanks to the decreasing cost of deploying cameras, it's a lot of easier and cheaper to surveille a selected location. With the event of the web, millions of police work videos are transmitted through the web. This would need viewing in time period to see if there are any vital events and additionally establish any suspicious behavior for an outsized quantity of captured video by corporations and security organizations. However, while not the assistance of extended man power support, most police work videos are ne'er watched or analyzed. Thus, the most challenge has been a way to process police work video so one will browse and retrieve important elements within the handiest and timely manner.

## 2. Related works

One example may be a video abstract based mostly technique, that has the flexibility to preserve the integrity of moving objects within the compressed domain [1]. Another viable approach is image retargeting. Due to its distinctive feature of fixing the image resolution while not affecting the necessary elements, this approach has been extensively studied within the past [2], [3], together with its extension to stereoscopic images [4], and videos [5]. Retargeting can be effective for protective necessary data.
Various approaches to calculate video digests are proposed, among that the only is quick forward [6] [8]. It is based on skipping individual frames in keeping with a selected rate; so limiting its capability as solely entire frames is removed. As a result, the condensation rate of quick forward is relatively low. In distinction, video account, which extracts key frames and presents them as a plot, has a very high condensation rate [9]. However, it's going to lose all the dynamic content of the initial video as solely key frames are presented. Another approach, video paste-up, extracts relevant spatial-temporal segments of the video and combines them into a digest video [10]–[13]. The limitation of such Associate in Nursing approach is that it's going to cause content loss and its quality is kind of high.
Video abstract may be a compact illustration of video that permits efficient browsing and retrieval. this method is in a position to generate digest video from the initial version. Generally, video synopsis foremost defines objects of interest and handles them as tubes in frame of reference volume. Then every object is temporary shifted to avoid collision between totally different objects. during this method, a condensed digest video is generated. It ought to be noted that in distinction to different techniques, condensed video generated by video abstract is in a position to specific the whole dynamics of the scene. Additionally, video abstract might amendment the relative timing between objects so as to cut back temporal redundancy as much as doable [14].
Authors in [15]–[17] have projected a technique to condense video from frame of reference video volume by extending seam-carving [3] in 2-D pictures, that was recently projected for image retargeting. [15] made Associate in Nursing economical spatio-temporal group to try and do seam carving for outline video. [16] incised sheet from the frame of reference volume of video. In [17], the authors proposed the new construct of ribbon-carving (RC) in frame of reference video volume. totally different from intensity-based price operate for seam-carving, RC uses Associate in Nursing activity-based price operate. In this way, RC is in a position to recursively carve ribbons so as to get digest video by minimizing Associate in Nursing activity-aware price operate with dynamic programming. additionally, ribbon-carving uses a flex-parameter so as to trade-off the anachronism of events and video condensation rate [17]. The trade off is that the condensation rate of RC is

kind of restricted since it solely carves the pixels of the static background.

In [5], key casings are figured in light of unsupervised learning for video recovery and video rundown by mix of shot limit identification, intra-shot-bunching and key frame "meta-grouping". It misuses the Shading Design Descriptor (CLD) [6], on back to back casings and process contrasts between them characterize the limits of each shot. As of late, dynamic programming methods have been proposed in the writing, for example, the MINMAX approach [7] of to extricate the key casings of a video succession. In this work, the issue is unravelled ideally in O (N2 · Kmax), where Kmax is identified with the rate-contortion streamlining. In [8], a video is spoken to as an entire undirected chart and the standardized slice calculation is done to all around and ideally parcels the chart into video groups. The subsequent groups shape a coordinated worldly diagram and a briefest way calculation is proposed for video synopsis.

In this paper, we propose a novel video synopsis approach to compute video digest. Based on seam carving [3] video retargeting, our proposed approach is consisting two modules. The first is Video retargeting using seam carving able to reduce the non-object content, as well as the redundancy in the movement of objects, and second is video synopsis by comparison the distortion between the sequential frames in the video. In this way, we may condense a video in a high ratio while preserving the important motion of objects.

The remainder organized as follows. Firstly, we introduce our proposed video synopsis method in Section II, in which we include video retargeting by using spatial and temporal coherence followed by minimization of sequential distortion between the frames. Then we evaluate the performance of our proposed method in Section III. Finally, we draw our conclusions in Section IV.

## 3. Proposed method

In video retargeting, the effort needs to traverse through spatial and temporal coherence.

### 3.1. Spatial coherence cost

In this, the seam carving operator which was proposed by Shai Avidan et al. by calculating the energy of the image we look forward to seams to be carved is used. Saliency map is used to identify the significant contents of the image. The saliency regions can be found out by considering a range of frequencies of different pixels i.e. using band pass filters. The saliency map will high lighten the significant objects in the frame with perfect borders as shown in the fig 1.

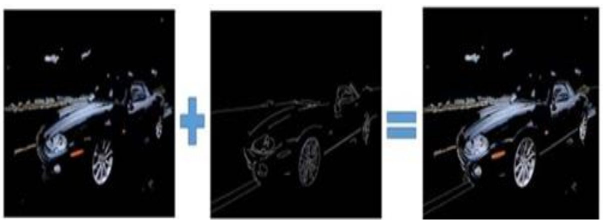

**Fig. 1:** (A) Saliency Map (B) Canny Detection (C) Output Image with Clear Edges.

In some cases, the edges were formed by the noise. This problem can be rectified by using canny detection. This technique uses gradients of Sobel and finds the strong edges. The weak edges in the image are suppressed based on thresholding and the noise can be reduced by using the Gaussian filter. The addition of saliency map and canny detected image gives the image with significant contents with strong edges so as to preserve them.

$$C_L(I,J) = \begin{cases} \text{if interior}, |I(i,j+1) - I(i,j-1)| + \\ \quad |I(i-1,j) - I(i,j-1)| \\ \text{if border}|I(i-1,j) - I(i,j-1)| \end{cases}$$

$$C_V(i,j) = I(i,j+1) - I(i,j-1)$$

$$C_R(i,j) = \begin{cases} \text{if interior}, |I(i,j+1) - I(i,j-1)| + \\ \quad |I(i-1,j) - I(i,j+1)| \\ \text{if border}, |I(i-1,j) - I(i,j+1)| \end{cases}$$

Seam carving algorithm is applied after calculating the compound energy. Dynamic programming is used to solve the complex issues. This procedure requires more time and may not produce better results. So the frame under goes quality checks before carving the seam right after seam removal computation. The similarity of the processed frame and the original frame defines the cost. To measure the similarity, bi-Directional warping can be used especially in video retargeting. The cost may be increased due to the computation of similarity of the complete image. To reduce the cost we calculate the window where the seam is removed and measure the similarity. There is no need to evaluate the similarity at the remaining positions because those are unaffected by the algorithm.

### 3.2. Temporal coherence cost

As said earlier the seam carving algorithm must be applied to individual frames because there may be chances of occurrence of artifacts if applied directly to a video. We identify the videos which are recorded from a static camera arrangement or dynamic camera arrangement.

In static camera arrangement, the recorded video background remains same throughout the video. For this, the sum of the difference between the frames can be calculated with equation (2).

$$\sum_{i=0}^{l-1} F_i(x,y) - F_{i+1}(x,y)$$

This gives the difference map that consists of the overall motion of the object in the image and the seams calculated for the first frame are directly applied to next frames. Here the cost is effective.

**Fig. 2:** A) Seams in Static.

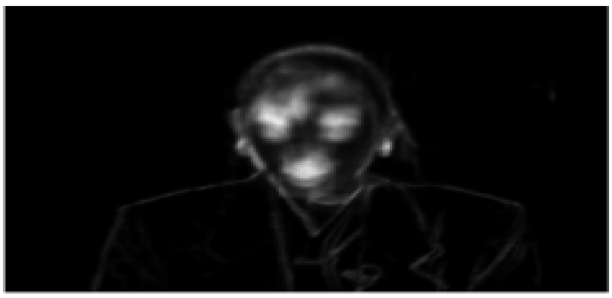

**Fig. 2:** B) Difference Map.

In dynamic camera arrangement, the video background changes from time to time. Here we are not using difference map because the object motion tracking is difficult. So we are using the previous seam in order to find the present seam of the frame. The searching for the pixels near the previous seam is based on thresholding.

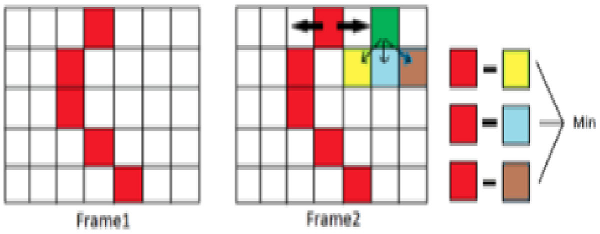


**Fig. 4:** Using Frame1 Seam to Frame2.

If the seam is not found the new seam is calculated for the present frame. Before calculating the new seam the coherence cost $T_c(i,j)$ is added to the energy map of the frame. These techniques can reduce the artifacts in the video.

$$T_c = \left|I(i,j) - S^{l-1}(j)\right|$$

### 3.3. Methodology for video synopsis

The issue of video abstract has a place with video synopsis issues. Its objective is to make another video, shorter than the underlying video as indicated by guaranteed parameter α, without noteworthy loss of substance between the two recordings (the mutilation between the first video and the video rundown is limited). The proportion between the transient span of the video abstract and the underlying video is equivalent to β ∈ [0, 1]. Give N a chance to mean the number edges of the unique video. At that point, the video summary comprises of β:N outlines. In this manner, we need to choose the·N delegate key casings. The telecom of the video abstract is finished with the first edge proportion implying that the genuine speed of the new video has been expanded by the factor of 1 α all things considered. For instance, we have a video with 5 sec term with 25 outlines/sec, so the entire video is comprised of 5× 25 = 125 outlines and the given parameter α = 0.2 the last video will have 125×0.2 = 25 outlines. At the end of the day, the last length will be one sec, which is 20% of starting the video. Let $L_i$, i ∈ {1... K} mean the visual descriptor of I-edge of the unique video. Let S ⊂ {1... K} indicate the edges of video abstract. As indicated by the issue definition, it holds that the quantity of edges of video abstract (|S|) is equivalent to β: K. At that point, the mutilation D ({1... K}, S) between the first video and video rundown is given by the accompanying condition:

$$D(\{1 \ldots \ldots K\}, S) = \sum_{i=1}^{s(1)} d\big(i, S(1)\big) + \sum_{i=S(|S|)+1}^{N} d(i, S|S|) + \sum_{i=S(1)+1}^{s(|s|)} \min_{S(j) \le i \le S(j+1)} (d(i, S(j)), d(i, S(j+)$$

Where d (i, S (j)) signifies the separation between the visual descriptor of I-outline what's more, S (j)- outline. S (j) and S (j + 1) are two progressive edges of video summary with the goal that S (j) ≤ i ≤ S (j + 1), this implies S(j) is controlled by the list I. The first and the second parts of this entirety concern the cases that the casing i is situated before the primary key casing S (1) or after the last key casing S(|S|), separately. In this manner, the utilized bending that is characterized by the entirety [4]of visual separates between the casing of unique video and the "nearest" relating edge of video summary, can be considered as an augmentation of the meaning of Iso-Content Contortion rule in the space of shots.

### 3.4. Low noise model (LDM)

Fig. 1 gives a method of the proposed system block diagram. This method has several steps to execute. First, we calculate the gray layout includes for all the frames in the original video. Based on the shot detection results and to the given parameter β we estimate the number of frames per shot that the video synopsis. Finally, the video contortion is consecutively limited by the proposed strategies coming about to the video outline.

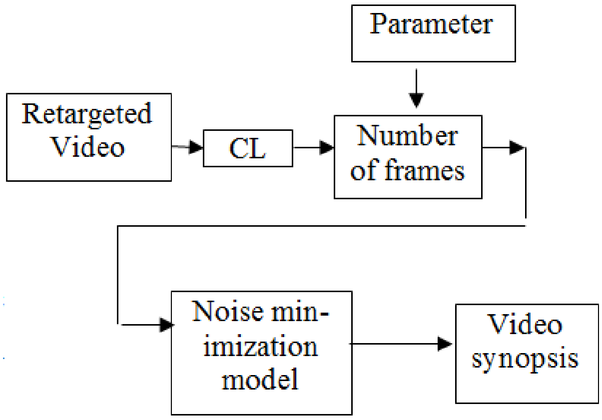


**Fig. 5:** Block Diagram of the Proposed Work.

The proposed technique can be executed under any decision or blend of sound/visual substance descriptors. Descriptors in view of picture division come about or, on the other hand on camera movement estimation procedures are computationally costly. Additionally, there is no certification that their outcomes will be exact for any video content variety. To conquer these issues, we embrace the MPEG-7 visual descriptors as proper highlights, for example, the Shading Format Descriptor (CLD), a minimal effort and minimized descriptor, which gets the job done to portray easily the progressions in the visual substance of a shot. These descriptors have been effectively utilized on our pre mentioned take a shot at key casings extraction issues. The CLD is a smaller descriptor that utilizations agent hues on a network took after by a DCT and encoding of the subsequent coefficients. We utilized the accompanying semi -metric work D to quantify the substance separation of two CLDs, {DY, DCb, DCr} and {DY', DCb', DCr'}

$$D = \sqrt{\sum_i (DY_i - DY_i')^2} + \sqrt{\sum_i (DCb_i - DCb_i')^2} + \sqrt{\sum_i (DCr_i - DCr_i')^2}$$

Where (DY, DCb, DCr) represent the $i^{th}$ DCT coefficients of the respective elements.
This segment displays the proposed Low Noise model for video summary creation. This technique chooses $b_k$ outlines for the k shot, with the goal that the mutilation between the first video and the video rundown is successively limited. The requesting of key edges choice compares to their criticalness on content depiction. Give $Ck_n$ a chance to signify the arrangement of competitor edges of shot k for video outline. At first, we set $CK_n = SH_n$. Give $S_k$ a chance to be the casings of video abstract of shot k. At first, we set $S_n = \emptyset$. For each shot k, we iteratively select the edge f from $CK_n$ so that in the event that we incorporate it in set $S_n$ the present video contortion of shot n is limited (see Condition 2). Next, we expel it from set $CK_n$ and we include it on set $S_n$:

$$f = \mathrm{argmin}_{u \in CK_n} \sum_{i \in SH_k} D(SH_n, S_n \cup u)$$

$$CK_n = CK_n - \{f\}, S_n = S_n \cup f$$

At the point when the quantity of key casings of shot k progress toward becoming $b_k$ discharge set ($CK_n = \emptyset$), since we cannot choose more casings from this shot. The procedure proceeds to the point that the quantity of key edges of video summation move toward becoming α.N. concerning the computational cost, this technique can be actualized in P (K2). The most pessimistic scenario is showed up when the video comprises of one shot. In this case, it holds that (K = |SH1|). In the begin (clench hand step), the finding of worldwide minima of D ({1... K}, ∅) needs P (K2) (see

Condition 1). In the n-venture of the strategy, we need to figure D ({1... N}, S ∪ u) just when the past or the next key casing of u is the last key edge that have been included S in past step (n−1). Else, it holds that D ({1, ...,k}, S∪U) = D({1, ...,N}, S). This needs O (K2 n2), since the video content changes "easily" as in the chosen outlines are about similarly disseminated amid the time. Let T (.) mean the calculation cost of the calculation. It holds that T (1) = P (K2). In the nstep, we need to locate the base of D(., .) that can be given in O(K) and to refresh the particular estimations of D(…) in P(K2 n2 ). In this way, the aggregate computational cost is P (K2).

# 4. Results

We processed different videos taken from static and dynamic arrangements using distortion less seam carving and compared our results with the seam crop technique. Bunny video consisting of 64 frames are reduced from 150x150 to 120x150 the objects frame are preserved and when differentiated with the seam crop the squirrel lost in the frame is preserved by video retargeting(fig 6). Golf video consisting of 83 frames are reduced from 256x150 to 200x150 the objects frame are preserved and when differentiated with the seam crop the tree lost in the frame is preserved(fig 7). The video from motion camera, water -ski was processed and differentiated (fig 8). We end up with the LDM based video synopsis.
The proposed strategies LDM method has been contrasted and the substance equidistant and time equidistant calculations in similar informational collections and same arrangement of parameters α =0.1 and α =0.3.The substance equidistant calculation (CEA) is roused by the work, where the iso-content rule has been proposed to appraise the key edges that are equidistant in video content.

**Fig. 6:** A) Original Image B) Retargeting Using Seam Crop C) Our Approach with Reduction of Aspect Ratio -20%.

**Fig 7:** A) Original Image B) Retargeting Using Seam Crop C) Our Approach with Reduction of Aspect Ratio -25%.

**Fig. 8:** A) Original Image B) Retargeting Using Seam Crop C) Our Approach with Reduction of Aspect Ratio -44%.

**Fig. 9:** Snapshots of Videos That We Have Used in the Paper.

**Table 1:** The Distortion D ({1... N}, S) between the Original Video and Video Synopsis

| Data set | α=0.1 LDM | α=0.1 CEA | α=0.1 TEA | α=0.3 LDM | α=0.3 CEA | α=0.3 TEA |
|---|---|---|---|---|---|---|
| Toy | 19209 | 21814 | 22069 | 6755.1 | 7738.2 | 8992.2 |
| Coast-guard | 6962.7 | 7486.6 | 7079.9 | 2562.4 | 2669.5 | 4146.4 |
| golf | 3913.8 | 4309.1 | 4444.4 | 2137 | 2228.1 | 3863 |

Table 1 shows the bending D ({1... N}, S) between the first video and video outline of LDM, CEA, TEA techniques under the ten utilized video groupings with α = 0.1 and α = 0.3.
As per these tests, LDM yields the most astounding execution comes about, beating alternate calculations, since in 95% of cases (19 out of 20) gives the least bending.
At the point when α = 0.3 is dependably the second most astounding execution strategy. At the point when α = 0.1, in 70% of cases is the second most elevated execution strategy. Furthermore, in foreman.avi, LDM gives the least twisting when α = 0.1.

## 4.1. Limitations and future scope

There are some of the limitations. The Computation time of the process is more. If the advancements in the image retargeting technique clear the problem of computation time. If there is more number of important objects in the frame we are helpless to get a good result. Handling the shadowing problem is difficult.
In the LDM video synopsis technique, the results and comparisons with existing methods are compared. Our LDM method on a large set of data is executed which consists real time videos. The real time videos are taken from indoor or outdoor environments.
In future, we will work on the 3D videos which will become a burning issue. The computation time must be reduced depends on the advancements in the image anlysing techniques.

# References


[1] *Avidan, S. and Shamir, A. 2007. Seam carving for contentaware image resizing. ACM Transactions on Graphics, vol. 26, no. 3.* https://doi.org/10.1145/1276377.1276390.

[2] *Achanta, R. and Su¨sstrunk, S., "Saliency Detection for Content-aware Image Resizing," in [IEEE Intl. Conf. on Image Processing], (2009).* https://doi.org/10.1109/ICIP.2009.5413815.

[3] *Rubinstein, M., Shamir, A., and Avidan, S. 2009. Multioperator media retargeting, ACM Transactions on Graphics, vol. 28, no. 3* https://doi.org/10.1145/1531326.1531329.

[4] *Grundmann, M., Kwatra, V., Han, M., and Essa, I. 2010. Discontinuous seam-carving for video retargeting. Computer Vision and Pattern Recognition (CVPR), 569-576* https://doi.org/10.1109/CVPR.2010.5540165.

[5] *Kiess, J., Guthier, B., Kopf, S., and Effelsberg, W. 2012. SeamCrop: Changing the size and aspect ratio of videos. Workshop on Mobile Video, 13-18.* https://doi.org/10.1145/2151677.2151681.

[6] *Miaomiao Zhao, Hongxia Liu and Yi Wan, An Improved Canny Edge Detection Algorithm Based on DCT, 2015 IEEE.* https://doi.org/10.1109/PIC.2015.7489844.

[7] *Shamir, A. and Sorkine, O., "Visual media retargeting." in [SIGGRAPH ASIA Courses], ACM (2009)* https://doi.org/10.1145/1665817.1665828.

[8] *M. Rubinstein, A. Shamir, and S. Avidan. Multi-operator media retargeting. ACM Transactions on Graphics, SIGGRAPH 2009, 28(3):1–11, 2009.* https://doi.org/10.1145/1531326.1531329.

[9] *M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir. A comparative study of image retargeting, 2010. ACM* https://doi.org/10.1145/1866158.1866186.

[10] *N. Petrovic, N. Jojic, and T. S. Huang, o fast forward,"Multimedia Tools and Applications, vol. 26, no. 3, pp. 327–344, 2005.* https://doi.org/10.1007/s11042-005-0895-9.

[11] *B. H¨oferlin, M. H¨oferlin, D. Weiskopf, and G. Heidemann, "Information-based adaptive fast-forward for visual surveillance," MultimediaTools and Applications, vol. 55, no. 1, pp. 127–150, 2011* https://doi.org/10.1007/s11042-010-0606-z.

[12] *M. A. Smith and T. Kanade, "Video skimming and characterization through the combination of image and language understanding," in Content-Based Access of Image and Video Database, 1998. Pro-*

*ceedings, 1998 IEEE International Workshop on. IEEE, 1998, pp. 61–70.*
[13] *C. Kim and J.-N. Hwang, "An integrated scheme for object-based video abstraction," in Proceedings of the eighth ACM international conference on Multimedia. ACM, 2000, pp. 303–311.* https://doi.org/10.1145/354384.354512.
[14] J. Oh, Q. Wen, S. Hwang, and J. Lee, "Video abstraction," Video data management and information retrieval, pp. 321–346, 2004. https://doi.org/10.4018/978-1-59140-571-9.ch014.
[15] Y. Li, T. Zhang, and D. Tretter, "An overview of video abstraction techniques," Technical Report HPL-2001-191, HP Laboratory, Tech. Rep., 2001.
[16] *M. M. Yeung and B.-L. Yeo, "Video visualization for compact presentation and fast browsing of pictorial content," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 7, no. 5, pp. 771–785, 1997.* https://doi.org/10.1109/76.633496.
[17] *C. T. Dang and H. Radha, "Heterogeneity image patch index and its application to consumer video summarization," Image Processing, IEEE Transactions on, vol. 23, no. 6, pp. 2704–2718, 2014.* https://doi.org/10.1109/TIP.2014.2320814.
[18] *S. Chakraborty, O. Tickoo, and R. Iyer, "Adaptive keyframe selection for video summarization," in Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on. IEEE, 2015, pp. 702–709.* https://doi.org/10.1109/WACV.2015.99.
[19] *S. Lu, Z. Wang, T. Mei, G. Guan, and D. D. Feng, "A bag-of-importance model with locality-constrained coding-based feature learning for video summarization," Multimedia, IEEE Transactions on, vol. 16, no. 6, pp. 1497–1509, 2014.* https://doi.org/10.1109/TMM.2014.2319778.
[20] *Y. Pritch, A. Rav-Acha, and S. Peleg, "Nonchronological video synopsis and indexing," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 30, no. 11, pp. 1971–1984, 2008.* https://doi.org/10.1109/TPAMI.2008.29.
[21] *J. Zhu, S. Feng, D. Yi, S. Liao, Z. Lei, and S. Z. Li, "High-performance video condensation system," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 25, no. 7, pp. 1113–1124, 2015.* https://doi.org/10.1109/TCSVT.2014.2363738.
[22] Y. Nie, C. Xiao and H. Sun, "Compact video synopsis via global spatiotemporal optimization", *IEEE Transactions on Visualization and Computer Graphics*, Vol.19, pp.1664–1676, 2013. https://doi.org/10.1109/TVCG.2012.176.
[23] K. Choeychuen, "An efficient implementation of the nearest neighbor based visual objects tracking", *International Symposium on Intelligent Signal Processing and Communications*, 2006. https://doi.org/10.1109/ISPACS.2006.364723.
[24] C.R. Huang, P.C. Chung and D.K. Yang, "Maximum a posteriori probability estimation for online surveillance video synopsis", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.24, pp.1417–1429, 2014. https://doi.org/10.1109/TCSVT.2014.2308603.
[25] Shao-Ping Lu, Song-Hai Zhang, Jin Wei and Shi-Min Hu, "Timeline editing of objects in video", *IEEE Transactions on Visualization and Computer Graphics*, Vol.19, No.7, pp.1218– 1227, 2013. https://doi.org/10.1109/TVCG.2012.145.
[26] Tong, Y., Maosen, X., Caiwen, M., *et al.*: 'Object based video synopses. 2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA), Ottawa, Canada, September 2014, pp. 1138–1141. https://doi.org/10.1109/WARTIA.2014.6976479.
[27] Zivkovic, Z.: 'Improved adaptive Gaussian mixture model for background subtraction'. Proc. 17th Int. Conf. on Pattern Recognition, 2004. ICPR 2004, 2004, vol. 2, pp. 28–31. https://doi.org/10.1109/ICPR.2004.1333992.
[28] Xin, L., Kejun, W., Wei, W., *et al.*: 'A multiple object tracking method using Kalman filter'. 2010 IEEE Int. Conf. on Information and Automation (ICIA), Harbin, China, June 2010, pp. 1862–1866.
[29] Lei, S., Junliang, X., Haizhou, A., *et al.*: 'A tracking based fast online complete video synopsis approach'. 2012 Int. Conf. on Pattern Recognition (ICPR), Tsukuba, Japan, November 2012, pp. 1956–1959.