Trends in Engineering and Technology (NCTET-2K17)
International Journal of Advanced Engineering Research and Science (IJAERS) Special Issue-4
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.19 ISSN: 2349-6495(P) | 2456-1908(O)

An Effective Approach to Streaming Video Segmentation with Sub-Optimal Low Rank Decomposition

[1]V.Sujatha., [2] P.Suresh babu
[1]M.Tech, Dept.of ECE, (DSCE), PG Scholar, SV College of Engineering

[2] Assoc Proff, Dept.of ECE, SV Engineering college for women

Abstract— ***This paper explores how to perform powerful and proficient video division while smothering the impacts of information commotions and additionally debasements, and a viable approach is acquainted with this end. Initial, a general calculation, called problematic low-rank decay (SOLD), is proposed to seek after the low-rank portrayal for video division. Given the information framework shaped by supervoxel elements of a watched video succession, SOLD looks for a problematic arrangement by making the lattice rank unequivocally decided. Specifically, the portrayal coefficient grid with the settled rank can be disintegrated into two sub-networks of low rank, and after that we iteratively streamline them with shut frame arrangements. Additionally, we join a discriminative replication earlier into SOLD in light of the perception that little size video designs have a tendency to repeat regularly inside a similar question. Second, in light of SOLD, we introduce an effective deduction calculation to perform gushing video division in both unsupervised and intuitive situations. All the more particularly, the obliged standardized cut calculation is embraced by joining the low-rank portrayal with other low level signs and fleeting reliable imperatives for spatiotemporal division. Broad analyses on two open testing informational collections VSB100 and Seg Track propose that our approach beats other video division approaches in both exactness and effectiveness.***

Index Terms—***Video processing, streaming segmentation, low-rank representation, spectral clustering.***

I. INTRODUCTION

Video division is to parcel the video into a few semantically predictable spatio-transient locales. It is a key PC vision issue in numerous applications, for example, protest following, action acknowledgment, video investigation, outline and ordering. Be that as it may, it is as yet a testing research zone because of its computational many-sided quality and intrinsic

challenges, similar to the vast intra-class varieties and the substantial between classification similitudes.

As per the measure of manual explanation, late video division calculations can be classified into four gatherings, i.e., unsupervised, intuitive, semi-administered and regulated.

1) The unsupervised strategies deliver reasonable spatial-worldly areas from the base up design, and have been presented running from mean-move, otherworldly bunching, chart based preparing and super pixel

Following. In addition, a few benchmarks have likewise been given to assess existing strategies and help additionally think about

2) A little measure of human toward the begin casing or edges is required in the intuitive ways to deal with fragment the frontal area from the foundation. Some of these methodologies are emphatically intelligent that enable the client to amend any errors on top of it if necessary.

3) The semi administered forefront proliferation approaches acknowledge a casing named physically with the closer view area and spread it to the rest of the edges.

4) Methods for the administered setting endeavor to portion a similar protest or question classification of enthusiasm as frontal area by taking in a protest show from named models. In this paper, we examine the issue of gushing video Segmentation under the Low-Rank Representation (LRR) framework.1 Although LRR had been extremely fruitful in picture division, there exists a few residual issues for applying LRR to video division. To begin with, most LRR calculations unwind the rank limitation with the atomic standard to make the goal tractable. The casual target for the most part is streamlined by ALM strategy which meets gradually, making it computationally wasteful for video division. Second, it is

International Journal of Advanced Engineering Research and Science (IJAERS) Special Issue-4
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.19 ISSN: 2349-6495(P) | 2456-1908(O)

demonstrated that inside video measurements is useful to enhance division execution, however it remains not all around contemplated for fusing inner video insights into LRR.

At long last, to adapt to discretionarily long video, transiently reliable imperatives is key for spilling video division. Gone for tending to these issues and inspired by the advances in subspace grouping, particularly LRR strategies for picture division, we propose a successful approach for spilling video division with a Sub-Optimal Low-rank Decomposition (SOLD) calculation, which seeks after the low-rank portrayal by abusing the low-rank structure of low-level supervoxel highlights. It is notable that the rank requirement can stifle the impacts of serious clamors as well as debasements, which is critical for strong video division. Rather than pixels or super pixels in past works, we take super voxels as diagram hubs to deduce their ideal affinities. Super voxels can protect nearby spatiotemporal intelligibility and in addition great limits, and significantly enhance division productivity. We expect that the intra-class super voxels are drawn from one indistinguishable low-rank element subspace, and all super voxels in a worldly window lie on a union of various subspaces. In this, a worldly window is characterized as various adjoining outlines. In this manner, we can speak to each supervoxel descriptor as a direct blend of other supervoxel descriptors, and look for the low-rank portrayal of all super voxels in a joint mold. Additionally, we likewise incorporate discriminative replication earlier in the plan for improving its discriminative capacity. These earlier, nearby little size video shapes (e.g., 6×6×6 voxels) with certain appearance designs have a tendency to repeat fre quently inside the semantic district, however may not show up in the diverse semantic areas, abuse the little non-neighborhood repeating locales to refine affinities among supervoxels. Thus, a semantic district is characterized as an arrangement of spatio-fleeting pixels of a similar question. It additionally can be seen as the augmentation of inward picture measurements for video information, yet can significantly diminish the computational multifaceted nature. Not at all like unwinding the rank minimization to the atomic standard minimization in different works, the rank of the portrayal coefficient network in SOLD is expressly decided for better portrayal. Specifically, the portrayal coefficient network with the settled rank can be disintegrated into two low rank sub-grids. Accordingly, we productively upgrade the low-rank portrayal by iteratively taking care of a few sub-issues with shut shape arrangements. The enhancement arrangement is then utilized to characterize affinities among super voxels. In view of SOLD, two unique undertakings, unsupervised and intelligent video division, are tended to in our system. To begin with, we join the low-rank portrayal grid with other low-level signals to characterize the liking lattice. At that point, we straightforwardly apply compelled NCut calculation [24] on the characterized fondness network to accomplish the unsupervised division. In intelligent assignment, we characterize the appearance models of closer view and foundation by client associations, and join with the low-rank portrayal and the spatiotemporal smoothness limitations to precisely fragment the objective question. We define it as the Markov Random Field (MRF) issue, which can be effectively understood by the Primal-Dual technique.

Fig. 1 illustrates the unsupervised
and interactive segmentation results of our approach.

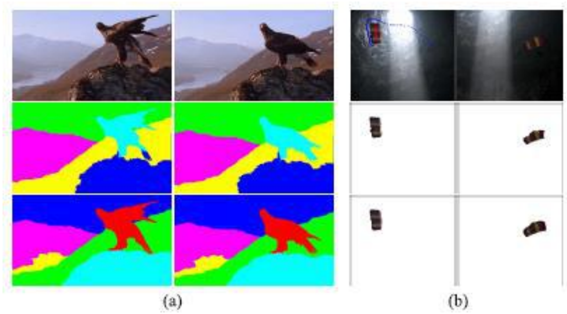


Fig. 1. The unsupervised and interactive segmentation results of our approach are shown in (a) and (b), respectively. The different colors indicate the different regions in (a). The first row shows the quintessential frames in video sequences, and the dash lines in (b) indicate the user interactions, in which the red denotes the foreground and the blue denotes the background. Our results and the corresponding ground truth are shown in the middle and last row, respectively.

This paper makes the following contributions to video processing and related applications. It presents an effective approach for segmenting videos into consistent spatio-temporal regions, which pursues the low-rank representation of the video supervoxel feature matrix. Our approach is able to deal with both unsupervised and interactive scenarios and outperforms other video segmentation methods on the standard benchmarks. It presents a novel low-rank decomposition method with the fixed-rank representation coefficient matrix, achieving a very efficient sub-optimal solution by iteratively solving three closed-form sub-problems. This proposed method can be extended to other similar tasks for pursuing low-rank representations. I t utilizes an internal replication prior for enhancing discriminative ability between super voxels, which is naturally incorporated into SOLD. Moreover, we utilize several temporal consistent

International Journal of Advanced Engineering Research and Science (IJAERS) Special Issue-4
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.19 ISSN: 2349-6495(P) | 2456-1908(O)

constraints during the inference of streaming video segmentation, effectively improving the robustness.

II. LITERATURE REVIEW

Some of the relevant state-of-the-art methods on the unsupervised and interactive video segmentation are reviewed in this section. Recent advances in hierarchical methods, streaming methods [5], [28] and related benchmarks [7], [8] have shown that unsupervised supervoxel segmentation has gained potential as a first step in early video processing. Hierarchical video segmentation provides a rich multi scale decomposition of a given video. Grundmann *et al.* [4] proposed Hierarchical Graph-Based video segmentation (HGB) algorithm based on local properties. It iteratively merged nodes in a region graph to produce a hierarchical segmentation. To process arbitrary long video, Xu *et al.* [5] proposed a streaming hierarchical video segmentation framework and instantiated HGB within this framework (SHGB). This method enforced a Markov assumption on the video stream, which leveraged ideas from data streams. Galasso *et al.* [28] proposed a spectral graph reduction algorithm for efficient streaming video segmentation. In this method, the reduced superpixel graph was reweighted such that the resulting segmentation was equivalent to the full graph under certain assumptions. Xu and Corso [7] presented a thorough evaluation of five supervoxel methods on a suite of suitable metrics designed to access supervoxel desiderata. A united video segmentation benchmark was provided by Galasso *et al.* [8] to evaluate effectively over- and under-segmentation of current video segmentation methods. These benchmarks not only allow to analyze the current state-of-the-art in video segmentation, but encourage the progress on new aspects of the video segmentation methods.

Recent works on video segmentation focus only on salient moving objects by analyzing point trajectories, while taking background as a single cluster [2], [29]. Some other works [3], [6] over-segment frames into superpixels, and partition them spatially and match them temporally. These methods provide a desirable computational reduction and powerful within-frame representation [30]. For instance, Galasso *et al.* [3] proposed a robust Video Segmentation approach with Super pixels (VSS) to explore various within and between-frame affinities. In addition, Tarabalka *et al*[31] presented a more efficient method for joint segmentation of monotonously growing or shrinking shapes in a time sequence of noisy images, and this method was applied to three practical problems to validate its performance and practicality. Different from unsupervised video segmentation, interactive video segmentation focused on extracting foreground object in clutter background with simple user interventions (often just one scribble for the object and one for the background). Wang *et al.* [9] introduced a hierarchical mean-shift preprocess to reduce the number of nodes for efficient computation, and extended 2D alpha matting scheme to 3D video volumes. Bai *et al.* [10] presented an interactive framework for soft segmentation and matting of natural images and videos. The proposed technique was based on weighted geodesics distance functions, which can be solved in computationally optimal linear time. It also allowed additional constraints into the distance definition to efficiently handle occlusions. A learning based method was proposed by Price *et al.* [11] to automatically weighted multiple features by learning from the previous implicitly-validated frame or the user corrections required in the previous frame. The above methods segmented or matted object frame by frame, and may require additional supervision in more complex videos. The long video intervals (up to 100 frames) were considered by Dondera *et al.* [12] on the basis of occlusion and long term spatio-temporal structure cues. Their system obtained good results quickly by running spectral clustering on superpixels.

III. SUB-OPTIMAL LOW-RANK DECOMPOSITION

Given an arbitrarily long input video, we adopt the overlapping sliding temporal window approach to save memory and space. In this section, we focus on the proposed model to obtain the low rank coefficient matrix **Z** of the supervoxel feature matrix of a temporal window.

A. Formulation

The proposed low rank decomposition model is imposed on the supervoxel for better tradeoff of efficiency and accuracy. We over-segment a temporal window into supervoxel by employing unsupervised video segmentation method [4], where each supervoxel comprises an ensemble of voxels that are coherent both spatially and temporally, and perceptually similar with respect to certain appearance features (*e.g.* color). Generally speaking, multi-level supervoxel representation can provide more appearance and motion features. However, as shown in Fig. 2

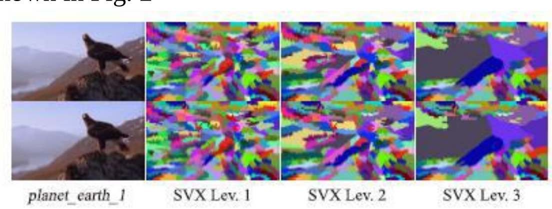


Fig. 2. Sample supervoxels at level 1 (200, where 200 indicates the number bof supervoxels), 2 (150) and 3 (100) extracted from a hierarchical video segmentation [4]. The different colors indicate the different supervoxels. the finest-level supervoxels have good spatiotemporal coherence and boundaries whilst the coarse-level supervoxels usually introduce large under-segmentation errors. Therefore, our

https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.19 ISSN: 2349-6495(P) | 2456-1908(O)

model is formulated in the finest-level supervoxels to avoid error propagation. Each temporal window of the video is segmented into n supervoxels. For each supervoxel, a set of appearance and motion features are extracted and combined into one single d-dimensional feature vector $\mathbf{x}i$ for supervoxel representation. Then, all the feature vectors of the n supervoxels form the data matrix $\mathbf{X} = [\mathbf{x}1, \mathbf{x}2, \ldots, \mathbf{x}n] \in \mathrm{R}d{\times}n$. We assume that supervoxels belonging to the same semantic region are all drawn from the same low-rank subspace, and all supervoxels in one temporal window lie on a union of multiple subspaces. Even SOLD is nonconvex and sub-optimal, in our experiments, such formulation can deliver both efficient algorithms and promising video segmentation accuracy. The Sub-Optimal Low-rank Decomposition (SOLD) model is then formulated as, where β is a regularization parameter that controls over fitting.

Optimization

To optimize above equation we adopt the alternating optimization method, and denote

Algorithm 1 Optimization Procedure to Eq. (7)

Input: The supervoxel feature matrix $\mathbf{X}$, the discriminative replication prior matrix $\mathbf{Q}$, the low-rank r, the parameter λ, β and γ;
Set $\mathbf{E} = 0$; $\varepsilon = 10^{-8}$, $maxIter = 500$.
Output: $\mathbf{A}$, $\mathbf{B}$, $\mathbf{E}$.

```
while not converged do
    Update A by Eq. (11);
    Update B by Eq. (9);
    Update E by Eq. (12);
    Check the convergence condition: the maximum element change of A, B, and E between two consecutive iterations is less than ε or the maximum number of iterations reaches maxIter.
end while
```

Our optimization delivers a more efficient algorithm. Some matrices (computing $\mathbf{S}1$ and $\mathbf{X}T\mathbf{X} - \gamma\mathbf{Q}$) in our algorithm can be pre-computed, and we require compute the top r generalization eigenvectors, where r is the desired rank. Note that r is generally much smaller than the size n of coefficient matrix $\mathbf{Z}$, making our algorithm more efficient to be optimized.

It should be noted that, both [38] and our SOLD adopt the alternating minimization algorithm, but the algorithm in [38] alternates between updating **A** and **B** while our SOLD alternates between updating **E** and {**A**,**B**}. Moreover, even for our specific sub problem on {**A**,**B**}, instead of the AltMin algorithm by [38], we suggest a generalized eigen value decomposition algorithm which can directly obtain the closed form solutions to **A** and **B**.

C. Implementation

Some important implementation details are briefly introduced. In this work, we utilize the hierarchical graph-based method (HGB) [4] to generate one layer supervoxels. HGB performs well on all the metrics of the unified video segmentation benchmarks [7], [8], and only involves one input parameter, *i.e.*, the total number n of supervoxels. Note that the supervoxel number n should not be set too small (large under-segmentation errors) or too large (heavy computational cost). On one hand, as shown in Fig. 2, the supervoxel segmentation result with $n = 100$ usually introduce large under-segmentation errors. On the other hand, as demonstrated in [7], when the supervoxel number n is between 200 and 900, the 3D under-segmentation error of HGB on the Seg Track [42] dataset only changes a little, and so do the other performance metrics including boundary recall, segmentation accuracy and explained variation. Therefore, it is reasonable to set $n \geq 200$.

Moreover, considering that the computational complexity of SOLD (two SVD operations in each iteration) is $O(n3)$, we should let n as small as possible, and thus set $n = 200$ to balance the accuracy-efficiency tradeoff. For robust supervoxel description, four low-level features are extracted from supervoxels and normalized with unit _2 norm. These feature vectors, including 12-dimension color histogram in each channel of RGB, 58-dimension Local Binary Pattern (LBP), 31-dimension Histogram of Oriented Gradient (HOG) and 18-dimension Histogram of Optical Flow (HOF), are concatenated into a single descriptor vector. To reduce computational complexity, we perform PCA [43] on the feature matrix **X** to remove insignificant components. The same operation is employed on computing the discriminative replication prior matrix **Q**.

IV. STREAMING VIDEO SEGMENTATION

In this section, we will deploy the optimized low-rank representation in Sect. III to perform streaming video segmentation in both unsupervised and interaction scenarios. First, the coefficient matrix **Z** is combined with other low-level cues to define the affinity matrix. Then, we apply NCut with temporal consistent constraints for clustering supervoxels. Finally, both unsupervised and interactive video segmentation can be conducted based on the supervoxel clustering results.

An effective streaming (sometimes called online as a synonym) algorithm can enable us to process an arbitrary-long video with limited memory and computational resources. Thus, it is essential to perform video segmentation in a streaming way. To this end, we segment the video in overlapping sliding windows. In particular, we consider both the temporal consistent constraints and low-rank representations to improve the long-range consistency and segmentation accuracy of the inference algorithm.

A. Affinity Definition

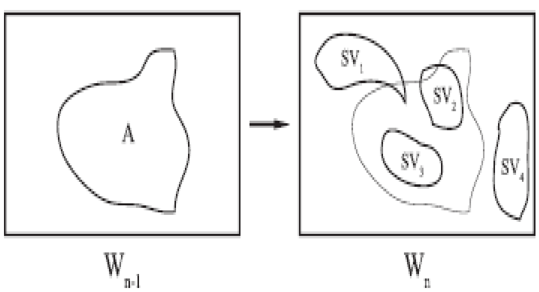
International Journal of Advanced Engineering Research and Science (IJAERS) Special Issue-4
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.19 ISSN: 2349-6495(P) | 2456-1908(O)

We define the affinity between two supervoxels as a linear combination of three cues. Herein, supervoxel-pair *(i, j)* is decomposed into the within-frame super pixel-pairs. The *Edge(x)* is the edge strength computed by gradient at location x and $\alpha 1$ is a tuning parameter. $\varphi 2$ is the smoothness kernel

B. NCut With Temporal Consistent Constraints

The temporal consistent constraints are introduced to properly propagate solutions between neighboring windows. We utilize some reasonable constraints to propagate the segmentation labels, while avoiding some bad results should not affect the quality of segmentation in the future frames. To this end, we divide the supervoxels into two categories as follows.

Given segmentation labels of the current window, the supervoxels in the next are divided into the deterministic supervoxels and the non-deterministic supervoxels. More specifically, the deterministic supervoxel is defined completely belonging to one specific label, and the non-deterministic supervoxel is defined as partly belonging to some label. Then, the partial grouping supervoxel set is composed by only the deterministic supervoxels. Figure shows this process.



The generation of the temporal consistent constraints between two neighboring sliding windows Wn−1 and Wn. A denotes one segmentation region in Wn−1, and provides some constraints to the segmentation of Wn. For clarity, four typical supervoxels are shown here, which stand for four typical supervoxel types based on their relationship to the region A: complete (SV3),

almost (SV2), part (SV1) and none (SV4). Thus, only SV2 and SV3 compose a partial grouping supervoxel set and generate a constraint due to A.

Algorithm 2 Temporal Consistent Constraint Matrix Computation Between Two Neighboring Windows

Input: Label set $\mathscr{T}$ from the previous window; Supervoxel set $\mathscr{S}$ in the current window.

Output: Temporal consistent constraint matrix $\mathbf{U}$.

for $t = 1 : |\mathscr{T}|$ **do**
  Find the deterministic supervoxel set $\mathscr{U}_t$ ($\subseteq \mathscr{S}$) for the label $\mathscr{T}(t)$ according to the overlap ratio of overlapping frame(s);
  $k = 0$;
  for $s = 1 : |\mathscr{U}_t| - 1$ **do**
    $k = k + 1$;
    $\mathbf{U}(k,\ \mathscr{U}_t(s)) = 1$;
    $\mathbf{U}(k,\ \mathscr{U}_t(s+1)) = -1$.
  end for
end for

C. Unsupervised Video Segmentation

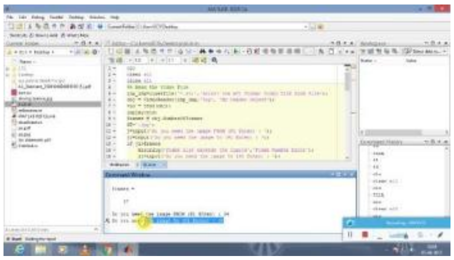
The eigenvectors $\hat{}\mathbf{V}[K]$ can be discretized by spectral rotation or k-means (spectral rotation in this paper) to obtain the discrete solutions of graph partition. In order to create new labels or remove old labels when the objects enter or leave the camera view, we utilize a reasonable strategy to determine the label mapping by their spatial overlap [6]. An overlap of one frame between neighboring windows is used to determine whether current labels are new ones or mapped from previous ones.

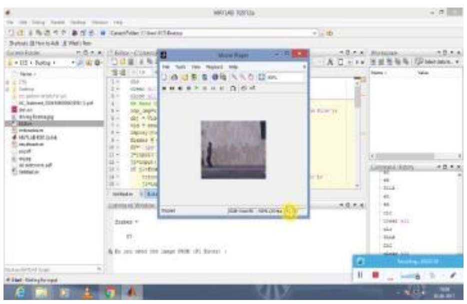
D. Interactive Object Segmentation

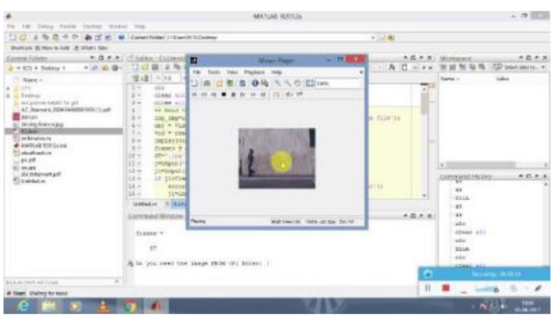
For the applications which utilize priors from user interactions, we employ an energy minimization approach to achieve interactive object segmentation. Since objects are spatially compact and temporally consistent, we integrate the appearance model of foreground and background by user interactions, the spatio-temporal smoothness constraints and the low-rank representation into our framework to accurately segment the target object. To this end, we formulate it as the MRF model, We adopt the Primal-Dual solver in MRF framework introduced in [25] to optimize it due to its high accuracy and efficiency. For obtaining reliable interactive segmentation in later sliding windows, we propagate the user interactions over time by supervoxel propagation in overlapping frame(s), instead of optical flow propagation because of its incorrect estimation. Furthermore, we divide the penalty node set *Ck* into the interactive node set (reliable) and the propagated node set (less reliable) from preceding segmentation results, and empirically set $\xi 1$ to be 1010 and 102, respectively. Herein, the process of supervoxel propagation is the same as generation of the temporal consistent constraints.

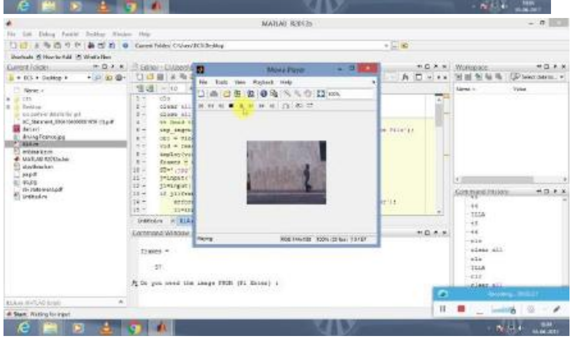
https://dx.doi.org/10.22161/ijaers/nctet.2017.ece.19 ISSN: 2349-6495(P) | 2456-1908(O)

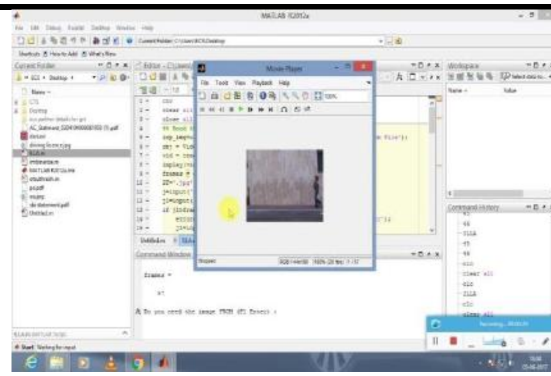
V.RESULTS

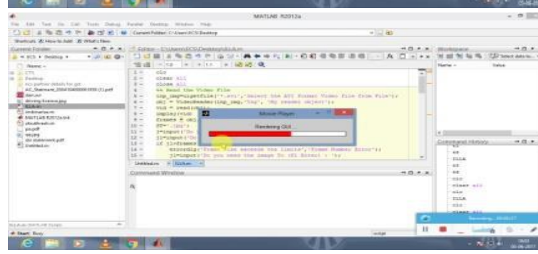


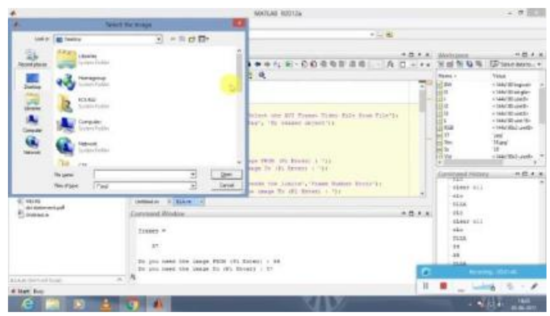












VI. CONCLUSION

In this paper, we have proposed a general algorithm for low rank representation pursuit by decomposing the matrix with the fixed rank and proved that a sub-optimal solution can be achieved by alternating closed-form optimization. Based on this algorithm, we have developed an effective and efficient approach that automatically segments streaming videos in both unsupervised and interactive way. In future work, we will improve our video segmentation framework by introducing more robust video features or deep feature learning methods. Our low-rank decomposition algorithm can be also extended to other vision tasks such as multi-object tracking and saliency detection.

REFERENCES

[1] S. Paris, “Edge-preserving smoothing and mean-shift segmentation of video streams,” in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 460–473.

[2] K. Fragkiadaki, G. Zhang, and J. Shi, “Video segmentation by tracing discontinuities in a trajectory embedding,” in *Proc.*

IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 1846–1853.

[3] F. Galasso, R. Cipolla, and B. Schiele, "Video segmentation with superpixels," in *Proc. Asian Conf. Comput. Vis.*, 2012, pp. 760–774.

[4] M. Grundmann, V. Kwatra, M. Han, and I. Essa, "Efficient hierarchical graph-based video segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2141–2148.

[5] C. Xu, C. Xiong, and J. J. Corso, "Streaming hierarchical video segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 626–639.

[6] A. Vazquez-Reina, S. Avidan, H. Pfister, and E. Miller, "Multiple hypothesis video segmentation from superpixel flows," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 268–281.

[7] C. Xu and J. J. Corso, "Evaluation of super-voxel methods for early video processing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1202–1209.

[8] F. Galasso, N. S. Nagaraja, T. J. Cardenas, T. Brox, and B. Schiele, "A unified video segmentation benchmark: Annotation, metrics and analysis," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3527–3534.

[9] J. Wang, P. Bhat, R. A. Colburn, M. Agrawala, and M. F. Cohen, "Interactive video cutout," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 585–594, Jul. 2005.

[10] X. Bai, J. Wang, D. Simons, and G. Sapiro, "Video SnapCut: Robust video object cutout using localized classifiers," *ACM Trans. Graph.*, vol. 28, no. 3, Aug. 2009, Art. no. 70.

[11] B. L. Price, B. S. Morse, and S. Cohen, "LIVEcut: Learning based interactive video segmentation by evaluation of multiple propagated cues," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 779–786.

[12] R. Dondera, V. Morariu, Y. Wang, and L. Davis, "Interactive video segmentation using occlusion boundaries and temporally coherent super pixels," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Mar. 2014, pp. 784–791.

[13] L. Lin, W. Yang, C. Li, J. Tang, and X. Cao, "Inference with collaborative model for interactive tumor segmentation in medical image sequences," *IEEE Trans. Cybern.*, doi: 10.1109/TCYB.2015.2489719, 2016.

[14] V. Badrinarayanan, F. Galasso, and R. Cipolla, "Label propagation in video sequences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3265–3272.

[15] K. Tang, R. Sukthankar, J. Yagnik, and L. Fei-Fei, "Discriminative segment annotation in weakly labeled video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2483–2490.

[16] B. Cheng, G. Liu, J. Wang, Z. Huang, and S. Yan, "Multi-task lowrank affinity pursuit for image segmentation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2439–2446.

[17] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.

[18] X. Liu, Q. Xu, J. Ma, H. Jin, and Y. Zhang, "MsLRR: A unified multiscale low-rank representation for image segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2159–2167, May 2014.

[19] Z. Lin, A. Ganesh, J. Wright, M. Chen, L. Wu, and Y. Ma, "Fast convex optimization algorithms for exact recovery of a corrupted lowrank matrix," Univ. Illinois Urbana–Champaign, Champaign, IL, USA, Tech. Rep. UILU-ENG-09-2214, 2009.

[20] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 1–8.

[21] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2790–2797.

Author's details:

Miss.V.Sujatha: she has received his Bachelor degree in Electronics and Communication Engineering from JB Women's Engineering College at tirupati and currently pursuing her Postgraduate in DSCE in SV College of Engineering at tirupati.

Mr.P.Suresh babu: He has received his M.Tech degree in LICS from SV University at tirupati.. He completed his UG in Electronics and Communication Engineering from KSRMCE ,Affiliated to SV University at tirupati.He pursuing PhD in signal processing from SV University at tirupati. He is currently working as a assoc proff in sv college of engineering.