Survey: Foreground Segmentation of Live Videos Using Boundary Matting Technology

Leena Kalbhor[1], Sachin V. Todkari[2]

[1]M.E (Computer) Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, India.
Savitribai Phule Pune University, Pune, Maharashtra, India -411028

[2]Professor (IT) as Head of Department of Information Technology, Jayawantrao Sawant College of Engineering, Pune, India.
Savitribai Phule Pune University, Pune, Maharashtra, India -411007

Abstract: *In this paper we to propose interactive methodology to extract foreground segmentation of objects from input videos using Boundary Matting Technology. Given an initial segmentation for the primary associated frame of a first and last video sequence, our objective is to propagate the segmentation from the middle frame in the sequence. Video frames are 1st divided into super pixels. As a final step the initial segmentation boundaries are improved to get an accurate object boundary. We then exploit these object boundaries in associate application for matting these videos. Here an effective algorithm for segmentation and then matting them is done accordingly which processes live videos with fuzzy boundaries. In our algorithm we generate Support Vector Machine (SVM's) and also algorithm which models local color distribution for both foreground and background video frames.*

Keywords: Foreground segmentation, Boundary Matting, Support Vector Machine (SVM„s).

1. Introduction

Image segmentation is the technique of separating an image into multiple parts. This is typically used to identify things or other allied data in digital images. Segmentation subdivides a picture into its constituent regions or objects. Segmentation may be a method of grouping along pixels that have similar attributes. Image Segmentation is that the process of dividing an image into non-intersecting regions such every region is consistent and also the union of no 2 adjacent regions is consistent.

Foreground segmentation, studies the way to extract objects from input videos. It's an elementary drawback in laptop vision and sometimes is a pre-processing step for different video analysis tasks like police work, teleconferencing, action recognition and retrieval. Over the years a big quantity of techniques are planned in each laptop vision and graphics communities. However, a number of them square measure restricted to sequences captured by stationary cameras, whereas others need massive coaching datasets or cumbersome user interactions. What is more, most existing algorithms square measure rather difficult and computationally strict. As a result, there still lacks an economical and powerful algorithmic rule capable of process difficult live video scenes with minimum user interactions.

This paper addresses the matter of accurately extracting a foreground layer from video in real time. A primary application is live background substitution in teleconference. This demands layer separation to close lighting tricks quality, together with transparency determination as in video-matting [2, 3], however with procedure potency comfortable to achieve live streaming speed. Intended by the on top of finding, we express a unique integrated foreground segmentation and boundary matting approach.

2. Literature Survey

In this we are going to make comparison of different papers. In Real time Background Subtraction from Dynamic Scenes [4] a basic problem in video content study is to detect moving foreground objects from background scenes, the suggested a method to work with the highly parallel graphics processors (GPUs). It can handle monotonous background motion, such as flowing water and waving plants, but none can deal with large camera motion. The frame-by-frame processing characteristics of the online approaches, on the other hand, make them more natural to work with live videos. Some of these approaches assume that foreground and background are properly and densely labeled in each frame by an existing algorithm [5], [6]. Foreground segmentation of live videos using locally competing 1SVMs [7], is a novel foreground segmentation algorithm is suggested in this paper that is able to efficiently and effectively deal with live videos. The formula is simple to implement, easy to use, and capable of handling a spread of inauspicious eventualities, like dynamic background, camera motion, topology variations, and fuzzy object boundaries. Our integrated foreground segmentation and boundary matting method, which is an extension to our preliminary work on foreground segmentation. Finally, compared to our preliminary work that focuses on foreground segmentation [7], the algorithm discussed here incorporates an additional matting step into the original train-relabel procedure, allowing both foreground segmentation and boundary matting problems to be solved in an integrated manner. To properly utilize the information extracted from matting calculation, the training process has been revised and more precise report on processing time are added throughout the paper.

Minglun Gong, Yiming Qian, and Li Cheng [1] proposes the key plan to coach and maintain 2 competitor one-class support vector machines at every constituent location, that model native color distributions for each foreground and

background, severally. The usage of 2 competitor native classifiers, as we've advocated, provides higher discriminative power whereas permitting higher handling of ambiguities. By exploiting this projected machine learning technique, And by addressing each foreground segmentation and boundary matting issues in an integrated manner, our algorithmic rule is shown to be notably competent at process a large vary of videos with advanced backgrounds from freely moving cameras. This is often typically achieved with minimum user interactions. What is more, by introducing novel acceleration techniques and by exploiting the parallel structure of the algorithmic rule, close to real-time operation speed (14 frames/s while not matting and eight frames/s with matting on a midrange computer & GPU) is achieved for VGA-sized videos.

M. Gong and Li Cheng [7] explains that there is a vast body of existing work for foreground segmentation. Here we've got to specialize in the foremost connected ones, that area unit is classified into unsupervised [20, 17, 16, 10, 12, 15] and supervised [13, 11, 19, 14, 18, 8, 9] approaches. Unsupervised ones try and generate background models mechanically and notice outliers of the models as foreground. Most of them, denoted as background subtraction methods, supposed that the input video is captured by a stationary camera and model background colors at every picture element [20, 17] or statistic [16, 10] ways. A number of these techniques will handle repetitive background motion, like wave water and waving trees, however none will touch upon camera motion. On the opposite hand, supervised ways enable users to produce training examples for each foreground and background then use them to be told classifiers. Variety of vital ways [8, 11, 19] on this line are developed with spectacular results, wherever multiple visual cues like color, contrast, motion, and stereo area unit utilized. These area units are integrated with the assistance of structured prediction ways like conditional random fields. Though operating all right for video conferencing applications, these algorithms need an outsized set of totally annotated pictures and considerable quantity of offline coaching, that observe several problems once applying to completely different scene setups.

Ting Yu Cha Zhang, Michael Cohen, Yong Rui Ying Wu [21] presents a new approach to segment monocular videos captured by motionless or hand-held cameras videoing enormous moving non-rigid foreground items. The foreground and background objects are showed using three-dimensional color Gaussian combination model (SCGMM), and segmented by means of the graph cut algorithm. Observing existence of a modeling gap between the existing SCGMMs and segmentation task of a new frame, one major role of paper is the overview of a new foreground/background SCGMM joint tracking algorithm to associate this space, which significantly increases the segmentation performance in situation of composite or rapid motion. Specifically, they have combined the two SCGMMs into a generative model of the complete image, and maximize the joint data probability using a inhibited Expectation-Maximization (EM) algorithm. The effectiveness of the proposed algorithm is established on a variety of sequences.

Liyuan Li, Weimin Huang, Irene Y.H. Gu, Qi Tian [22] states that for detection and segmentation of foreground objects from a video that contains each stationary and moving background objects and undergoes each gradual and unforeseen "once-off" change. A Baye's regulation for association of background and foreground from selected feature vectors is developed. Underneath this rule, different types of background objects are classified from foreground objects by selecting a correct feature vector. The stationary background object is delineating by the color feature, and also the moving background object is diagrammatic by the color co-occurrence characteristic. Foreground objects are extracted by fusing the classification results from each stationary and moving pixel. Learning methods for the gradual and unforeseen "once-off" background changes are projected to adapt to numerous changes in background through the video. The convergence of the training method is tried and a formula to pick out a correct learning rate is additionally derived. Experiments have shown promising ends up in extracting foreground objects from several complicated backgrounds as well as wavering tree branches, unsteady screens and water surfaces, moving escalators, gap and shutting doors, change lights and shadows of moving objects.

Yi Wang [38] described those two implementations that are capable of separating a foreground layer from video sequences. The first one is a method known as "background subtraction", and the other one is resultant from "Bilayer Segmentation of Live Video" [13] by Criminisi et al. It is a totally unique methodology for detection and segmentation of foreground objects from a video they may be used as building blocks for 2 of the continued analysis comes within the Vision place of work, i.e., the "fruit-fly" and so the "human action" comes. The fundamental methodology of background subtraction is to match |frame – background| with a pre-defined threshold. If the distinction of a component is larger, then classify it as foreground; otherwise, claim that it's background.

Paper ID: NOV152009

790

Table 1: Survey Table

Paper Name	Author Name	Proposed Work	Advantages	Disadvantages
Integrated Foreground Segmentation and Boundary Matting for Live Videos	Minglun Gong, Yiming Qian, and Li Cheng,	According to the user in the primary frame of the video, the rule is in a position to propagate labeling info to adjacent pixels through a straightforward train-relabeled-matting procedure, leading to a correct segmentation of the frame.	Easy to implement, Simple to use & able to handle a variety of complex scenarios and possesses similar or finer performance.	Handling: Low Resolution Images, Illumination changes, Large Semi-Transparent Objects, Novel Foreground Colors.
Foreground Segmentation of Live Videos using Locally Competing 1SVMs	Minglun Gong, Li Cheng	According to the user in the primary frame of the video, the rule is in a position to propagate labeling info to neighboring pixels through a straightforward train-relabeled-matting procedure, leading to a correct segmentation of the frame.	Easy to implement & efficiently deals with live videos.	Here, Matting is not performed.
Monocular Video Foreground/Back ground Segmentation by Tracking Spatial-Color Gaussian Mixture Models	Ting Yu Cha Zhang, Michael Cohen, Yong Rui, Ying Wu	This paper presents a new approach to segment monocular videos captured by static or hand-held cameras filming massive moving non-rigid foreground items. The foreground and background objects are modeled using spatial color Gaussian mixture models (SCGMM), and segmented using the graph cut algorithm, which minimizes a Markov random field energy function containing the SCGMM models.	The proposed method can not only handle difficult and rapid motions of the foreground / background objects, but also help determine occlusion / deocclusion issues caused by their motions.	For complex as well as fast foreground / background motions, the SCGMM well-read from earlier frames are not appropriate for the segmentation task of the existing frame due to large variations of the spatial components.
Foreground Object Detection from Videos Containing Complex Background	Liyuan Li, Weimin Huang, Irene Y.H. Gu, Qi Tian	In this paper proposes a novel method to extract foreground objects from a real-time complex video under the Bayes decision framework. A Bayes decision rule for classification of background and foreground from a general feature vector is formulated. For the meantime, an effective data structure to study and keep up the statistics of different feature vectors is established.	Using different features to model different parts of background is more accurate than using just one type of features for complex background.	The weak point of the proposed method is that it is prone to absorb foreground objects if they are motionless for a long time.
Bilayer Segmentation of Live Video	A. Criminisi, G. Cross, A. Blake, V. Kolmogorov	This paper presents an algorithm capable of real-time partition of foreground from background in monocular video sequences.	Good quality composite videos are developed. Achieves comparable levels of accuracy.	Accurate estimation of optical flow is computationally expensive.
Segmenting Foreground Objects from a Dynamic Textured Background via a Robust Kalman Filter	Jing Zhong and Stan Sclaroff	The objective is to segment the foreground objects in video given time varying, textured backgrounds. Foreground background segmentation algorithm that accounts for the non-stationary nature and clutter-like appearance of many dynamic textures. The dynamic quality is modeled by an Autoregressive Moving Average Model (ARMA). A Kalman filter algorithm describes the appearance of the dynamic texture and also the regions of the foreground objects.	Algorithm can successfully segment the foreground objects, even if they share a similar grayscale distribution with the background.	ARMA model only takes grayscale images as input.
Real-Time and Temporal-Coherent Foreground Extraction With Commodity RGBD Camera	Mengyao Zhao, Chi-Wing Fu, Jianfei Cai, and Tat-Jen Cham	In this paper, a CUDA-based coherent approach for real-time foreground object extraction from RGBD videos is done. High quality foreground extraction is achieved.	This method is fully automatic without any manual markup and is also capable of extracting more stable foreground objects.	When the foreground object moves very fast, object becomes blurred and so, nearby background pixels gets misrecognized. Also, if background region is enclosed by foreground object in image, our method gets misrecognized.

3. Proposed Work

Foreground segmentation is one of the major tasks in the field of Computer Vision whose aim is to detect changes in image sequences. In this we are going to extract foreground objects from input videos. There are limitations in an algorithm that can develop live videos of things with furry boundaries (e.g., hair) captured by freely moving cameras. The key idea is that we are going to use two competing one-class support vector machines at each pixel location, which gives confined color distributions for both foreground and background, provides higher discriminative power while allowing better handling of ambiguities. By using this competing one-class support vector machines and by addressing both foreground segmentation and boundary matting problem our algorithm is competent at processing a wide range of videos with complex backgrounds from freely moving cameras and with minimum user interaction. Furthermore we are also calculates near time processing speed by introducing novel acceleration techniques and by exploiting the parallel structure.

4. Conclusion

In this paper we conclude that there are many methods for foreground and background segmentation. The objective of foreground segmentation is to remove the required foreground item from input videos. Over the years, there have been remarkable amount of hard work on this topic. Yet, there still lacks a simple yet effective algorithm that can develop live videos of items with blurry boundaries (e.g., hair) captured by freely moving cameras. This algorithm is easy to execute, easy to use, and able to handle a variety of difficult scenarios, such as dynamic background, camera motion, topology changes, and fuzzy objects. The integrated boundary matting step can effectively pull the matte for fuzzy objects, allowing seamless composites over new backgrounds. Our algorithm possesses comparable or superior performance compared for specifically background subtraction, foreground subtraction and video matting.

References

[1] M. Gong, Y. Qian and Li Cheng, "Integrated Foreground Segmentation and Boundary Matting for Live Videos" in *IEEE Transactions On Image Processing*, Vol. 24, No. 4, April 2015.

[2] Y.-Y. Chuang, A. Agarwala, B. Curless, D. H. Salesin, and R. Szeliski, "Video matting of complex scenes," in *Proc. 29th Annu. Conf. Comput.Graph. Interact. Techn*., 2002, pp. 243–248.

[3] Y.-Y. Chuang, B. Curless, D. Salesin, and R. Szeliski. "A Bayesian approach to digital matting." In *Proc. Conf. Computer Vision and Pattern Recognition*, CD–ROM, 2001.

[4] L. Cheng and M. Gong, Gong, "Realtime background subtraction from dynamic scenes," in *Proc. IEEE 12th Int. Conf. Comput.* Vis., Sep./Oct. 2009,pp. 2066–2073.

[5] E. S. L. Gastal and M. M. Oliveira, "Shared sampling for real-time alpha matting," *Comput. Graph. Forum*, vol. 29, no. 2, pp. 575–584, May 2010.

[6] M. Gong, L. Wang, R. Yang, and Y.-H. Yang, "Real-time video matting using multichannel Poisson equations," in *Proc. Graph. Interf.*, 2010, pp. 89–96.

[7] M. Gong and L. Cheng, "Foreground segmentation of live videos using locally competing 1SVMs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2105–2112.

[8] X. Bai and G. Sapiro, "Geodesic matting: A framework for fast interactive image and video segmentation and matting" in *IJCV*, 82(2):113–132, 2009.

[9] X. Bai, J. Wang, D. Simons, and G. Sapiro, "Video snapcut: robust video object cutout using localized classifiers" in *SIGGRAPH*, 2009.

[10] L. Cheng and M. Gong, "Realtime background subtraction from dynamic scenes" in *ICCV*, 2009

[11] A. Criminisi, G. Cross, A. Blake, and V. Kolmogorov, "Bilayer segmentation of live video" in *CVPR*, 2006.

[12] E. Hayman and J. Eklundh, "Statistical background subtraction for a mobile observer" in *ICCV*, 2003.

[13] A. Criminisi, G. Cross, A. Blake, and V. Kolmogorov, "Bilayer segmentation of live video," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis.Pattern Recognit.*, Jun. 2006, pp. 53–60.

[14] Y. Li, J. Sun, and H. Shum, "Video object cut and paste" in *SIGGRAPH*, 2005.

[15] Y. Sheikh, O. Javed, and T. Kanade, "Background subtraction for freely moving cameras" in *ICCV*, 2009.

[16] Y. Sheikh and M. Shah, "Bayesian object detection in dynamic scenes" in *CVPR*, 2005.

[17] J. Sun,W. Zhang, X. Tang, and H. Shum, "Background cut" in *ECCV*, 2006.

[18] J. Wang, P. Bhat, A. Colburn, M. Agrawala, and M. Cohen, "Interactive video cutout" in *SIGGRAPH*, 2005.

[19] P. Yin, A. Criminisi, J. Winn, and I. Essa, "Tree-based classifiers for bilayer video segmentation" in *CVPR*, 2007.

[20] J. Zhong and S. Sclaroff, "Segmenting foreground objects from a dynamic textured background via a robust Kalman filter" in *ICCV*, 2003.

[21] T. Yu, C. Zhang, M. Cohen, Y. Rui, and Y. Wu, "Monocular video foreground/background segmentation by tracking spatial-color Gaussian mixture models," in *Proc. IEEE Workshop Motion Video Comput.*, Feb. 2007, p. 5.

[22] L.Li, W.Huang, I.Gu, and Q.Tian, "Foreground object detection from videos containing complex background" in *ACM International Conference on Multimedia*, 2003, pp. 2-10.

[23] Mengyao Zhao, Chi-Wing Fu, Jianfei Cai, and Tat-Jen Cham," Real-Time and Temporal-Coherent Foreground Extraction With Commodity RGBD Camera" in *IEEE Journal On Signal Processing*, Vol. 9, No. 3, April 2015

[24] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines", 2001. [Online]. Available: http://www.csie.ntu.edu.tw/~jlin/libsvm/

[25] G. Dalley, J. Migdal, and W. E. L. Grimson, "Background subtraction for temporally irregular

dynamic textures," in *Proc. IEEE Workshop Appl.Comput. Vis.*, Jan. 2008, pp. 1–7.

[26] J. Fan, X. Shen, and Y. Wu, "Scribble tracker: A matting-based approach for robust tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1633–1644, Aug. 2012.

[27] M. Gong and L. Cheng, "Real-time foreground segmentation on GPUs using local online learning and global graph cut optimization," in *Proc.19th Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.

[28] N. Joshi, W. Matusik, and S. Avidan, "Natural video matting using camera arrays," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 779–786, Jul. 2006.

[29] V. Kolmogorov, A. Criminisi, A. Blake, G. Cross, and C. Rother, "Bi-layer segmentation of binocular stereo video," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 407–414.

[30] S.-Y. Lee, J.-C. Yoon, and I.-K. Lee, "Temporally coherent video matting," *Graph. Models*, vol. 72, no. 3, pp. 25–33, 2010.

[31] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 228–242, Feb. 2008.

[32] M. McGuire, W. Matusik, H. Pfister, J. F. Hughes, and F. Durand, "Defocus video matting," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 567–576, 2005.

[33] C. Rhemann, C. Rother, J. Wang, M. Gelautz, P. Kohli, and P. Rott, "A perceptually motivated online benchmark for image matting," in *Proc.IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1826–1833.

[34] E. Shahrian, D. Rajan, B. Price, and S. Cohen,"Improving image matting using comprehensive sampling sets," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 636–643.

[35] J. Sun, J. Jia, C.-K. Tang, and H.-Y. Shum, "Poisson matting," in *Proc.ACM SIGGRAPH*, 2004, pp. 315–321.

[36] J. Wang and M. F. Cohen, "Image and video matting: A survey," *Found. Trends Comput. Graph. Vis.*, vol. 3, no. 2, pp. 97–175, Jan. 2007.

[37] L. Wang, M. Gong, C. Zhang, R. Yang, C. Zhang, and Y.-H. Yang, "Automatic real-time video matting using time-of-flight camera and multichannel Poisson equations," *Int. J. Comput. Vis.*, vol. 97, no. 1, pp. 104–121, 2012.

[38] Yi Wang ," Foreground-Background Segmentation of Video Sequences ",*Mentors: Pietro Perona, Claudio Fanti*

Author Profile

Leena Kalbhor, is currently pursuing M.E. (Computer Engineering) From Department of Computer Engineering, Jayawantrao Sawant College of Engineering, Hadapsar, affiliated to Savitribai Phule, Pune University, Pune, Maharashtra, India -411028. She received her B.E. (Computer Engineering) Degree from PGMCOE, affiliated to Savitribai Phule Pune University, Pune, and Maharashtra, India - 411007. Her area of interest is Image Processing and Data Mining.

Prof. Sachin V. Todkari, (Computer Science and Engineering) have completed his PG from MIT COE KOTHRUD PUNE. He received his B.E (Computer Science and Engineering) Degree from CSE college of Engineering Ambajogai. He is currently working as Professor as Head of Information Technology Department in Jayawantrao Sawant College of Engineering, Hadapsar, Pune, India 411028, affiliated to Savitribai Phule Pune University, Pune, Maharashtra, India -411007. His area of interest is Wireless Sensor Network etc