Automatic Interesting Object Extraction from Images based on Edge Information and Texture Analysis

Adiljon Djurayev , Guzal Primova

Department of Telecommunication Engineering, Samarkand branch of Tashkent University of Information Technologies, Uzbekistan
Odil85dj@gmail.com

Department ofTelecommunication Engineering, Samarkand branch of Tashkent University of Information Technologies, Uzbekistan
GuzalPrimova@gmail.com

Abstract— Tracking moving objects in video sequence is an important problem in computer vision, with applications in several fields, such as video surveillance and target tracking. As the shadows attached along with the moving object also have the same motion that of objects, the detection of shadows as foreground objects is very common and produce large errors in object localization and recognition. This paper introduces an effective method which uses the edge information to detect moving cast shadows for traffic sequences. The proposed method initially removes the boundary of the cast shadow, preserving object's interior edges. The coarse object shapes are then reconstructed using the object interior edges. Finally, the cast shadow is detected by subtracting the reconstructed moving object from the change detection mask. The method is implemented and tested using three benchmark videos. The efficiency of the proposed method is compared with five other popular shadow detection methods and the results proved its superiority over others.

Keywords—*Saliency region extraction; visual attention; visually impaired people; saliency map; saliency cut; image enhancement;*

INTRODUCTION

Shadow is one of the common parts in the natural scenes and has become an important topic in computer vision. In many computer vision applications, shadows interfere with fundamental tasks such as moving objects segmentation and tracking. Thus it's necessary to suppress the effect of shadows. From the viewpoint of geometric relationship, shadow can be divided into umbra and penumbra [8]. The umbra corresponds to the background area where the direct light is almost totally blocked by the foreground object, whereas in the penumbra area of the background, the lighting is partially blocked. From the viewpoint of spatial relationship, shadow can be divided into static shadow and moving shadow. Static shadow is cast by the static object in the scene while the moving shadow is cast by the moving object.

A cast shadow is found mainly according to the results of change detection, static edge detection, shading change detection and penumbra detection. There are some problems in this approach. Some regions of a moving object, such as the facial part of a human, are easy to be misclassified as shadow regions because the uniform colors there present the same characteristics as the shadow regions. The regions that are always shadowed along the sequence cannot be detected by their algorithm, as pointed out by the authors. Moreover, the computation is quite complex. So, nowadays, it is more and more necessary to establish an efficient and effective method for the shadow problem.

Facing the problem that the moving cast shadow is often misclassified as part of the moving object in change detection based video segmentation, we propose an effective approach to the detection and removal of insignificant moving cast shadows in normal indoor scenes where the camera is stationary. It is especially appropriate for the applications of indoor video surveillance and conferencing. The main contribution of this paper is that we successfully remove cast shadows from moving objects by the conditional dilation operation, where the edge and region information are used in a unified framework. Compared with the method in our approach does not require the region uniform property, and thus seldom misclassifies the uniform moving object region as a shadow region. Moreover, it can detect insignificant shadows appearing along a whole image sequence. We have compared our approach with the gradient filter method which is the most recent state of the art related to our approach, and the experimental results show that our approach improves the detection performance.

Related Works

Of the various algorithms designed for shadow detection, one way of classifying the approach for detection is provided in[1] considering whether the decision process introduces and exploits uncertainty. *Deterministic approaches* use an on/off decision process, whereas *statistical approaches* use probabilistic functions to describe the class membership. In statistical methods (see [2], [3], [4], [5]), the parameter selection is a critical issue. Thus, statistical approaches are further divided to parametric and nonparametric methods. The study reported in [2] is an example of the parametric approach, whereas [4] is an example of the nonparametric approach. The deterministic class (see [6], [7], [8] and [9]) can be further subdivided. Sub classification can be based on whether the on/off decision can be supported by modelbased knowledge or not. The system described in [6] and [9] are examples of deterministic non model based approach. Another classification of shadow detection techniques is provided in [10]: *Model-Based techniques* and *Property-Based techniques.* Model-based techniques are based on matching sets of geometric features such as lines or corners to 3D object models, and rely on models representing the a priori knowledge of the geometry of the scene, the objects, and the illumination. In property based approach the properties such as the brightness, color or edge of shadows are used to detect shadows. Techniques based on the brightness or color properties ([11], [12], [13]) will have problems if the foreground contains objects having brightness or intensity values similar to that of shadow pixels.

In such cases the object points will be misclassified as shadow points. A technique using the edge information can overcome such limitations. Reference [10] explains an edge based method for the removal of shadows. The method solved the limitations of other color property based techniques but the method fails to detect thin shadows as well as the shadows which are far away from the camera position. These limitations are addressed and a modified edge-based shadow detection method is proposed in this paper. This paper pays attention to detection of moving cast shadows for traffic sequences, where the shadow suppression is very important to avoid misclassification and erroneous counting of vehicles on the road. The proposed method is detailed in the next section.

PROPOSED MOVING SHADOW DETECTION METHOD USING EDGE INFORMATION

In our algorithm, the edge information plays an important role for shadow removal. Canny edge is also applied for video segmentation in. However there are still several obvious differences between these two approaches. Canny edges are used to find the edge of the moving object, and then the filling algorithms can be applied to extract the whole moving object. Actually it is not an easy task to find the actual edge of the moving object by Canny edge and the initial change detection mask, so two model update schemes are proposed to handle slowly changing components and rapidly changing components respectively. Moreover, to handle the un-closed boundary, some complex filling algorithms are also implemented as the postprocessing step.However in our approach, the Canny edge is implemented based on the property of insignificant shadow, i.e. although the moving cast shadows appear in the initial change detection mask, however, the fact that the transition from the backgrounds to the shadow regions is gradual makes the edges caused by the shadow boundaries almost invisible. Thus Canny edges can be used to find some initial seed points for shadow region. So, in a word, the Canny edges are used for a different purpose, and thus in a different way.

It can be observed that the edge of cast shadows have two major properties:

-The object will have significant interior edges; however the corresponding shadow region does not have much interior edges.

-The edge of the cast shadow fastens on the boundary region of the moving foreground mask.

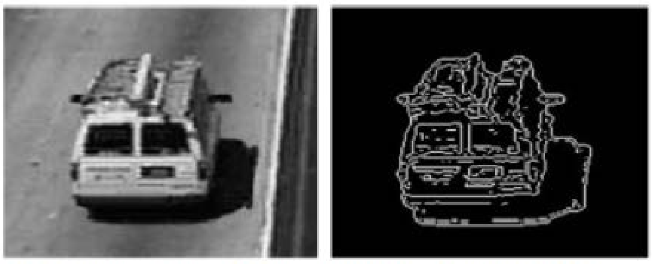


Figure 1. Input image and its corresponding foreground edge image.

Fig. 1.shows the input image and the foreground edge image corresponding to it.

The steps involved in the proposed shadow detection method are detailed in the following sections.

A. *Foreground Region Segmentation*

The easiest and common method, background subtraction, is used for segmenting the foreground region. The standard form of adaptive background is a time- averaged background image (TABI), where a background approximation is obtained by averaging a long time image sequences. The resulting background image is then used for background subtraction and the foreground image is obtained. The foreground image thus obtained contains the moving objects as well as its corresponding shadows. A binary mask, Mt, is also created from the foreground image by applying thresholding operation. The generated background image and the foreground region mask are shown in Fig. 2(b) and 2(c) respectively.

B. *Edge Detection for Foreground Region*

The edge image of the foreground region, E_t, is then obtained by using the classic Canny edge operator. E_t is shown in Fig. 2(d).

C. *Computing Interior Edge Image of Foreground*

All edges other than the boundary edges are interior edges. As the shadows do not have much interior edges, computing an interior edge image gives an edge image devoid of shadow edges.

The boundary of the foreground image, B_t, is obtained by applying Canny edge operator on M_t. And the interior edge image IE_t is obtained by

$$IE_t = E_t - DB_t \qquad (1)$$

where DBt denotes dilated B_t; dilated with a structuring element S1.
The interior edge image IE_t is shown in Fig. 2(e).

D. *Foreground Region Classification*

From the interior edge image IE_t which contain the interior edges of different objects, we have to identify which all interior edges belong to same object.

_ To accomplish this, IE_t is first filled horizontally to obtain IHE_t using a horizontal operation and the noise regions are removed. In horizontal operation the region inside the first and last edge points in each row is set to 1.

_ IE_t is then filled vertically to obtain IVE_t using a vertical operation. Here instead of row the column wise operation is performed. The small noise regions present in IVE_t have to be removed.

_ Then Blob Merging has to be done on both IHE_t and IVE_t to merge the erroneously split blobs which may occur during the above operations. For this, first the connected components in IHE*t* and IVE*t* have to be labeled. Any adjacent blobs are regarded as the same foreground, if the distance between two connected blobs is less than or equal to Td and they will be relabeled with the same number. When no further merging is possible, the procedure is terminated. The distance between two blobs in IHE*t* and IVE*t* is defined by number of rows and number of columns respectively. T*d* is given as

$$T_d = \alpha \,.\min\{\min_{i=1,2,\ldots,M}(HL^i),\ \min_{j=1,2,\ldots,N}(VC^j)\} \qquad (2)$$

where HLi is the number of rows of the blob labeled i in IHEt and VCj is the number of columns of the blob labeled j in IVEt. M and N are the total number of blobs of both labeled images. _ is a constant. IHEMt and IVEMt are the respective results of the merging operations on IHEt and IVEt.

_ The edge points corresponding to the regions which are common in both IHEMt and IVEMt are labeled with same number and they are considered to be that of the same foreground.

E. Reconstruction of the Moving Object

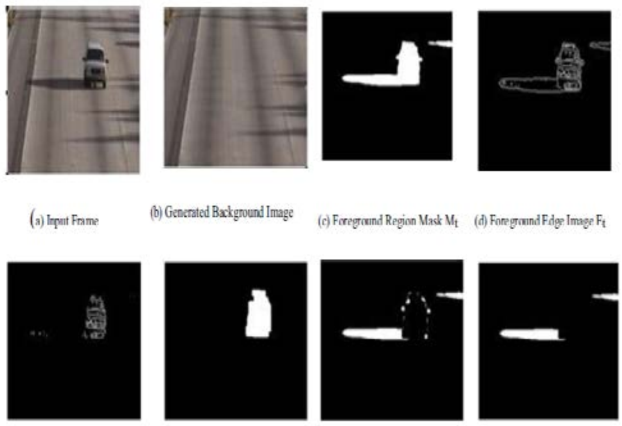
Each of the foreground objects has to be then reconstructed from those labeled interior edges. The procedure is as follows:

* The horizontal and vertical operation, explained above, has to be executed to the labeled edge image. Here instead of the entire image, the operation has to be applied on each of the labeled edges individually.

* The union of the components labeled 'i' in horizontally filled image and vertically filled

* The I th foreground object (i = 1,2..n, if there are n foreground objects). Similarly all 'n' foreground objects can be reconstructed.

* The above reconstructed objects will not contain the boundary portions as they are reconstructed from the interior edges. The boundary portions can be reconstructed by applying a dilation operation using the same structuring element S1 used in step C. The reconstructed foreground object image is denoted as Ft and is shown in Fig. 2(f).



(a) Input Frame (b) Generated Background Image (c) Foreground Region Mask M_t (d) Foreground Edge Image F_t

(e) Interior Edge Image IE_t (f) Reconstructed Object F_t (g) Shadow Regions IS_t (h) Detected Moving Cast Shadow Regions S_t

Figure 2. Illustration of the Proposed Shadow Detection Method.

F. Obtaining moving Cast Shadow Regions

The foreground mask M*t* contains both moving objects as well as shadows and F*t* contains only the foreground objects. So with M*t* and F*t* we can acquire the moving cast shadow region IS*t*.

$$IS_t = M_t - F_t \qquad (3)$$

As the perfect reconstruction of the foreground object is not possible, IS*t* will contain noise regions. So noise removal has to be done on ISt to obtain the accurate output St. IS*t* and S*t* are shown in Fig. 2(g) and 2(h) respectively. For the input frame if the detected shadow pixels are replaced with the corresponding background pixels, the shadow removed frame can be obtained.

EXPERIMENT AND EVALUATION

The proposed method has been implemented in a PC with INTEL PENTIUM IV, 3GHz processor using MATLAB 7.0.The efficiency of the method is analyzed using three different benchmark videos-*Highway I*, *Highway II* and *Campus* obtained from the CVRR database. Table 1 shows the properties of each of the benchmark videos used. The benchmark suite of video sequences and associated ground truth data (which is used for the empirical evaluation of the efficiency of the proposed method) is available at http://cvrr.ucsd.edu/aton/shadow (Courtesy: Andrea Prati). For simulation the constant _ mentioned in section III.D is taken as 0.5.

Fig. 3, Fig. 4 and Fig. 5 shows the results obtained for the proposed method for the video sequences *Highway I*, *Highway II* and *Campus* respectively. Fig. 3(a) corresponds to the input frames, where Fig. 3(a1) is frame no: 38 and Fig.3 (a2) is frame no: 115. Fig. 3(b) corresponds to the detected moving cast shadow regions, where Fig. 3(b1) is the result of Fig. 3(a1) and Fig. 3(b2) is that of Fig. 3(a2). Fig. 3(c1) and Fig. 3(c2) corresponds to the moving shadow removed frames of Fig. 3(a1) and Fig. 3(a2) respectively.

Fig. 4(a1) and Fig. 4(a2) corresponds to the input frames, frame no: 114 and 15 respectively of the *Highway II* video sequence. Fig. 4(b1) and Fig. 4(b2) are its corresponding detected moving shadow regions. Fig. 4(c1) and Fig. 4(c2) are the corresponding shadow removed frames.

TABLE I. PROPERTIES OF BENCHMARK VIDEO SEQUENCES

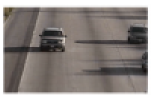
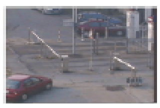
Video Sequence	Highway I 	Highway II 	Campus 
Sequence Type	Outdoor	Outdoor	Outdoor
Sequence Length	1074	1134	1179
Image Size	320X240	320X240	352X288
Shadow Strength	Medium	High	Low
Shadow Size	Large	Small	Very Large
Object Size	Large	Small	Medium
Object Speed (in pixels)	30-35	8-15	5-10
Noise Level	Medium	Medium	High

Figure 3. Results for *Highway I* video sequence.

Fig. 5(a1) and Fig. 5(a2) corresponds to the input frames, frame no: 53 and 66 respectively of the *Campus* video sequence. Fig. 5(b1) and Fig. 5(b2) are its corresponding detected moving shadow regions. Fig. 5(c1) and Fig. 5(c2) are the corresponding shadow removed frames. As the surface upon which the shadows are cast in *Campus* video sequence is highly textured, perfect shadow detection is not possible for this sequence. During edge detection this texture nature may cause interior edges to shadows which will affect the subsequent steps' efficiency.

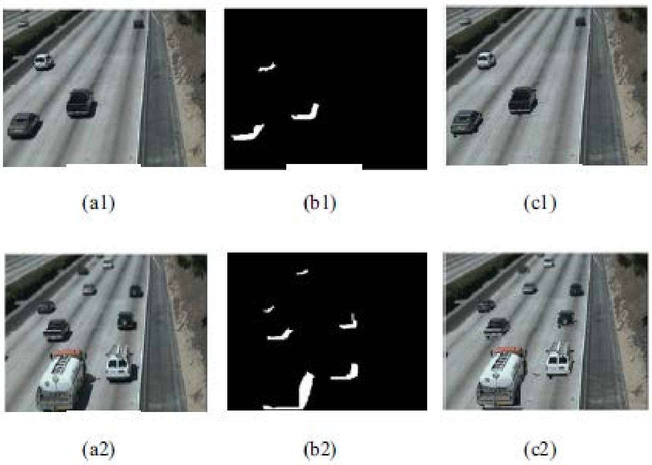


Figure 4. Results for *Highway II* video sequence.

From the results reported in Table 2 it is clear that the proposed moving shadow detection method has good superiority over other methods. In case of *Highway I* video sequence SNP and DNM2 have good detection rate, but their discrimination rates are not much impressive. The edge based method in [10] has got the highest discrimination rate but a comparatively lower detection rate. Compared to other methods, the proposed method shows good detection rate as well as high discrimination rate. *Highway II* is considered as a challenging test sequence for shadow detection methods. Almost all methods show their lowest detection rate for this video sequence. The highest detection rate for this video sequence is recorded by the proposed method along with good discrimination accuracy.

Figure 5. Results for *Campus* video sequence.

Even though *Campus* is a noisy sequence SP and DNM1 achieves good detection rate but low discrimination rate. Both detection and discrimination accuracies are low for DNM2. The proposed method has quite low detection rate for this sequence. This is due to the high texture nature of the background surface upon which the shadows are cast. But the proposed method exhibits the highest discrimination rate which proves that the method is robust even to strong noise.

CONCLUSION

In this paper, an efficient method for detecting moving cast shadows using edge information is proposed. As the proposed method initially performs motion segmentation, only the moving cast shadows are detected (not any background shadows). Unless other edge-based methods the proposed method can detect even thinner shadows and shadows that are far away from camera position.

The method is implemented and tested using three benchmark videos where the shadow sizes, shadow strength, shadow orientation, vehicle size, vehicle color etc varies. The method is then compared with five other popular shadow detection techniques, of which four of them are based on the color property of shadows and the other one is based on the edge property. As the results have shown, the method performs well in almost all test sequences. Though the method is mainly proposed for traffic video sequences, the shadow detection method is still valid if the shadow areas are not at all textured or less textured.

Future work

We further discuss possible applications to simplify inner edges of important object in the natural image. For the reason that, visually impaired people may not sure about object information just touching and feeling outer boundary of object, high complex images include a lot of unnecessary edges. Therefore, we will consider simplifying inner edges and add specific features to zoom in and zoom out image edges for more detailed information.

Acknowledgment

I wish to thank Senior Lecturer Guzal Primova under Department of Telecommunication Engineering for her support over the period in which this paper was written. I also express my gratitude to my parents for their encouragement and my colleagues for their constant cooperation. I would like to express my gratitude towards my college, Hanyang University, for giving me the opportunity to make its facilities available for me.

References

[1] Andrea Prati, Ivana Mikic, Mohan M. Trivedi and Rita Cucchiara, “Detecting Moving Shadows: Algorithms and Evaluation,” IEEE Trans in PAMI, vol. 25, no. 7, July 2003.

[2] Mikic, P. Cosman, G. Kogut, and M.M. Trivedi, “Moving Shadow and Object Detection in Traffic Scenes,”Proc. International Conference in Pattern Recognition, vol. 1. pp. 321-324, Sept. 2000.

[3] Zhi Liu, Oliver Le Meur and Shuhua Lua, “Superpixel-based saliency detection,”

[4] Sai Srivatsa R and R.Venkatesh Babu, “Salient object via objectness measure,” ICIP. IEEE, 2015.

[5] J. Stauder, R. Mech, and J. Ostermann, “Detection of Moving Cast Shadows for Object segmentation”, IEEE Trans. Multimedia, vol. 1, no. 1, pp. 65-76, Mar. 1999.

[6] R. Cucchiara, C. Grana, G. Neri, M. Piccardi, and A. Prati, “The Sakbot System for Moving Object Detection and Tracking”, Video- Based Surveillance Systems—Computer Vision and Distributed Processing, pp. 145-157, 2001.

[7] D. Koller, K. Daniilidis, and H.H. Nagel, “Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes”, Int’l J. Computer Vision, vol. 10, pp. 257-281, 1993.

[8] Noboku Tagaki and Jianjun Chen, “Character string extraction from scene images by eliminating non-character elements,” IEEE International Conference on Systems, Man, and Cybernetics, 2014.

[9] Sampada S Pathak, Prashant Dahiwale and Ganesh Padole, “A combined effect of local and global method for contrast image enhancement,” ICETECH. IEEE, 2015.

[10] Mei Xiao, Chong-Zhao Han,and Lei Zhang. “Moving Shadow Detection and Removal for Traffic Sequences”, International Journal of Automation and Computing vol. 04, no. 1, pp. 38-46, 2007.

[11] Jianping Shi, Qiong Yan, Li Xu, and Jiaya Jia, “Hierarchical Image Saliency Detection on Extended CSSD,” *IEEE Transactions On Pattern Analysis and Machine Intelligence,* Vol. 38, No. 4, April 2016.

[12] P. Felzenszwalb and D. Huttenlocher, “Efficient graph-based image segmentation,” *Int. J. Comput. Vis.,* vol. 59, no. 2, pp. 167– 181, 2004.

[13] K. Barnard and G. Finlayson, "Shadow Identification Using Color Ratios". In Proc. of the 8th Color Imaging Conference: Color Science, Systems, and Applications, pp. 97-101, 2000.

[14] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk, SLIC Superpixels, *EPFL Technical Report 149300*, June 2010.

[15] Junwei Han, Dingwen Zhang, Xintao Hu, Lei Guo, Jinchang Ren, and Feng Wu, “Background Prior-Based Salient Object Detection via Deep Reconstruction Residual,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 25, No. 8, August 2015.

[16] Quan Zhou, Ji Chen, Shiwei Ren, Yu Zhou, Jun Chen, Wenyu Liu, “On Contrast Combinations for Visual Saliency Detection,” *ICIP*, 2013.

[17] Gao Jie and Liu Ning, “An improved adaptive threshold canny edge detection algorithm,” *International Conference on Computer Science and Electronics Engineering,* 2012.

[18] Y. Zhai and M. Shah, “Visual attention detection in video sequences using spatiotemporal cues,” in Proc. *14th Annu. ACM Int. Conf. Multimedia,* 2006, pp. 815–824.

[19] Rita Cucchiara, Costantino Grana, Massimo Piccardi, and Andrea Prati, "Detecting Moving Objects, Ghosts, and Shadows in Video Streams", IEEE Transactions on PAMI, vol. 25, no. 10, Oct. 2003.

[20] Michael Donoser, Martin Urschler, Martin Hirzer and Horst Bischof, “Saliency Driven Total Variation Segmentatio,” *2009 IEEE 12th International Conference on Computer Vision (ICCV*)

[21] Quan Zhou, Nianyi Li, Yi Yang, Pan Chen and Wenyu Liu, “Corner-Surround Contrast for Saliency Detection,” *21st International Conference on Pattern Recognition (ICPR),* 2012.

[22] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk. Salient region detection and segmentation. International Conference on Computer Vision Systems, 2008.

[23] T. Horprasert, D. Harwood, and L.S. Davis, “A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection”, Proc. IEEE Int’l Conf. Computer Vision ’99 FRAMERATE Workshop, 1999..

AUTHORS

First Author - Adiljon Djurayev, Senior Lecturer, Department of Telecommunication Engineering, Samarkand branch of Tashkent University of Information Technologies, Uzbekistan, Odil85dj@gmail.com.

Second Author - Guzal Primova, Senior Lecturer, Department of Telecommunication Engineering, Samarkand branch of Tashkent University of Information Technologies, Uzbekistan, GuzalPrimova@gmail.com.