Rapid Technique to Eliminate Moving Shadows for Accurate Vehicle Detection

Kratika Garg Nirmala Ramakrishnan Alok Prakash Thambipillai Srikanthan Punit Bhatt
Nanyang Technological University, Singapore
kratika001@e.ntu.edu.sg

Abstract

Elimination of moving shadows is an essential step to achieve accurate vehicle detection and localization in automated traffic surveillance systems that aim to detect vehicles on road scenes captured by surveillance cameras. However, this is still a challenging problem as existing pixel based methods miss parts of vehicles and region-based methods, while accurate, incur higher computations. In this paper, we propose a highly accurate yet low-complexity block-based moving shadow elimination technique, which can effectively deal with varying shadow conditions. A novel shadow elimination pipeline is proposed that employs computationally lean features to quickly classify distinct vehicles from shadows, and uses a more sophisticated interior edge feature only for classification of difficult scenarios. Extensive evaluations on freely available and self-collected datasets demonstrate that the proposed technique achieves higher accuracy than other state-of-the-art techniques in varying scenarios. Additionally, it also achieves over 20 *times speedup on a low-cost embedded platform, Odroid XU-4, over a state-of-the-art technique that achieves comparable accuracy. Experimental results confirm the real-time capability of the proposed approach while achieving robustness to varying shadow scenarios.*

1. Introduction

Smart traffic law enforcement systems are an essential part of current and future smart cities to maintain smooth traffic flow and alleviate congestion on roads. Such systems need to employ complex video analytics pipelines to automatically process the video data captured from the extensive CCTV camera networks, in order to extract useful information such as traffic density on various roads, accident detection, vehicle breakdowns, etc..

A common necessity and typically the first step in all such application pipelines is to accurately detect the vehicles on road scenes. For traffic surveillance scenes with a stationary camera, background modeling techniques are typically used instead of the high-complexity generic vehicle detectors[20, 21] and segmentation techniques [16, 30]. Background modeling methods leverage on the availability of a static background (i.e. road) to model it and extract the moving objects as foreground. However, the extracted foreground not only contains the vehicles but also their moving cast shadows. This leads to several errors in the subsequent stages of the application pipeline that rely heavily on accurate vehicle detection and localization, such as over estimation of vehicle sizes and road occupancy rates, and erroneous vehicle tracking.

Several techniques have been proposed in the literature to deal with cast shadows. A number of background modeling techniques[28, 13, 27] have attempted to eliminate cast shadows during the background/foreground classification process. However, these techniques fail to handle "hard" cast shadows(dark shadows shown in Fig. 1) [29]. Thus, moving shadow elimination methods that strive to classify each pixel in the foreground mask into a vehicle or shadow pixel have been proposed in the literature [19, 25, 22]. These existing techniques can be broadly classified into two categories: pixel based and region based. The pixel based methods [8, 12] generally rely on the fact that shadows make the surface, on which they are cast, darker without changing the color. Therefore, the spectral information is exploited to distinguish the shadow pixels from the pixels belonging to the vehicles. However, these techniques wrongly classify parts of vehicles that have similar color as the road background, as shadows [19, 25].

In contrast, recently proposed region-based methods [24, 4, 7, 9, 23] first segment the foreground into regions and then classify these regions as either vehicle or shadow, based on texture or color similarity with the background. When compared to the pixel based methods, this approach ultimately provides a more accurately segmented vehicle from the foreground without losing parts of vehicles as shadows. However, the region based methods incur heavy computations in the initial pixel accurate region segmentation. These techniques, while more accurate, are hence not suitable for implementation on cheap embedded devices needed for large scale deployment in smart cities. Therefore, a computationally lean and rapid, yet robust, technique

978-1-7281-1975-5/19/$31.00 ©2019 IEEE
DOI 10.1109/WACV.2019.00214

for moving shadow elimination is essential for realizing the smart cities of the future.

In this paper, we propose a novel region based technique for moving shadow elimination that extracts predefined vehicle-sized blocks proposed in [10] as candidate regions and classifies them into vehicle or shadow. This top-down context-specific block assignment avoids the heavy computations that are otherwise needed, for pixel wise region segmentation, in the existing region based methods. The proposed technique ensures that we retain the entire vehicle after moving shadow elimination, while minimally allowing some shadows (that are very close to the vehicles) to be classified as vehicles. As will be shown later in our evaluations in Section 4, this strategy is able to achieve even higher accuracy in vehicle detection compared to the best performing region based methods.

The main contributions of this work are:

- A block-based evaluation and classification of pixels in the foreground mask, inspired by the earlier work on block based foreground detection for incident detection on road traffic scenes [10]. The main objective of our work is to remove large shadows that drastically impact the vehicle localization.
- We use a computationally lean set of features and cues in a cascaded fashion during the classification step, where simpler aggregate features are used in the earlier stages for fast detection of distinct vehicle blocks. A more complex, interior edge based feature, is only applied at the later stage on a much smaller subset of shadow candidates, to recover vehicle blocks that look very similar to shadows. We employ vehicle size constraints to further eliminate false shadow detections by using the novel block assignment method.
- We provide extensive evaluations of our proposed method on a wide range of road scenes in comparison to state-of-the-art shadow elimination techniques on a real low cost embedded platform. As existing datasets of road scenes [19], only contain side shadows (as shown in Fig. 1a), we have collected additional traffic data with front/back and long shadows as shown in Fig. 1b, 1c and 1d respectively, and include them in our evaluations. Additionally, there is a lack of annotated datasets with vehicle cast shadows (1 extremely low-angle video in CDNet[29] and 10 frames in UCSD dataset[19]). Thus, we propose a semi-automated ground truth generation method to generate vehicle-shadow masks for the datasets.

2. Related Work

Excellent surveys on existing work on detection of moving shadows can be found in [25, 22]. As we are interested only in moving shadow removal, all these methods assume

(a) (b) (c) (d)

Figure 1: Different types of Vehicle-Shadow scenarios

that the foreground mask has been generated and a background reference image is available. The basis of all methods is to distinguish the vehicle from the shadow by measuring the similarity of the shadow with the corresponding background image.

The first stage in moving shadow elimination is to limit the search for shadows to only those foreground pixels that are darker than the background. In [17], the expected location of the shadow with respect to the object is explicitly estimated to refine the shadow detection. However, in road scenes captured from stationary cameras, this relative shadow location can be calibrated offline. As the chromaticity is not affected under a shadow, the spectral information is used in [8, 12] to extract shadow regions with similar chromaticity as the background. However, in traffic scenes, color is not always a very strong feature, and therefore, these methods are known to suffer [25]. In [11, 33], the object shape with various shadow positions is modeled, using geometric features to first extract the object and then remove the unwanted shadow pixels. These methods are only applicable in very sparse traffic conditions, when the foreground masks of the vehicles are distinct from each other, and therefore unsuitable in realistic traffic scenes. A significant cue for detecting shadows is that the edges and texture on the background remain unchanged when covered by a shadow. Texture similarity is used for small patches in [15], and for larger regions after segmentation in [24, 4, 23]. Block based methods have also been proposed that operate on uniform sized blocks directly [7] or after clustering them, to form larger regions [9]. Although segmentation and block based methods that classify foreground regions have improved accuracy, they are still too heavy in computations for embedded implementations where shadow elimination needs to be performed real-time alongside other video analytics tasks. Recent methods [14, 18] employ thresholding of various wavelet coefficients to remove shadow pixels but are far from real-time even on high-end desktop platforms.

3. Proposed Approach

Our proposed approach for block-based moving shadow elimination is summarized in Fig. 2. First, the input pixel-wise foreground mask is used to generate vehicle-sized foreground blocks. Then, the pre-defined direction of shadows is used to extract candidate shadow blocks. Next, low-complexity aggregate features based on intensity and coarse texture, followed by a more sophisticated feature based on

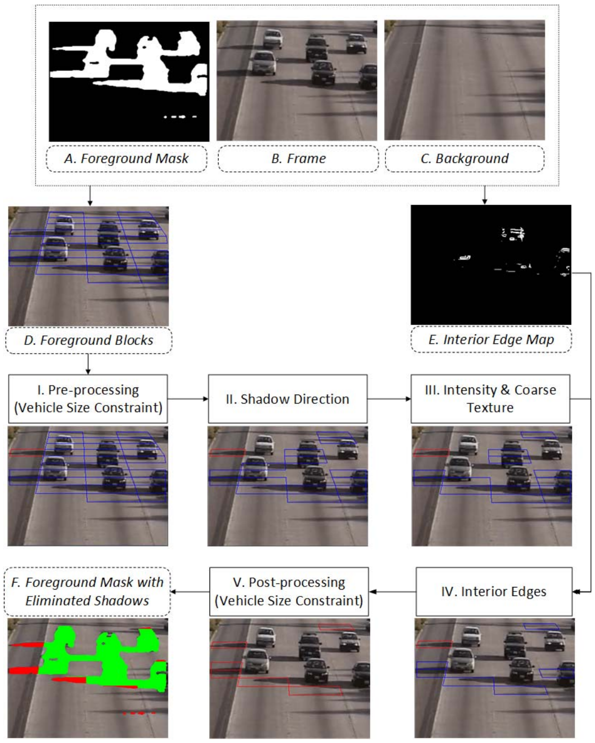


Figure 2: Overview of Proposed Approach [Red: Shadow, Blue: Candidate Shadow, Green: Vehicle, (Best viewed in color)]

interior edges, are successively employed to recover vehicle blocks from candidate shadow blocks. Additionally, pre and post-processing steps based on the size of the vehicle are applied to further refine the technique. Finally, a vehicle-shadow mask is generated by marking all foreground pixels inside a shadow block as shadow pixels and the rest as vehicle pixels. These steps are discussed in detail below.

3.1. Vehicle-Size Block Assignment

In order to extract the candidate shadow regions, we employ the block assignment proposed in [10] on the input pixel-wise foreground masks, and generate foreground blocks, with a width equal to the lane width and length equal to 1/3rd of a small vehicle (shown in Fig. 2-D). Only blocks containing at least 5% foreground pixels are considered.

3.2. Shadow Elimination

All the foreground blocks are now processed to be classified as vehicle and shadow blocks using our multi-feature cascaded shadow elimination technique as described below:

3.2.1 Pre-processing

The BoIs are divided in such a way that length of three blocks is equivalent to that of a small vehicle. Thus, single foreground blocks surrounded by background cannot

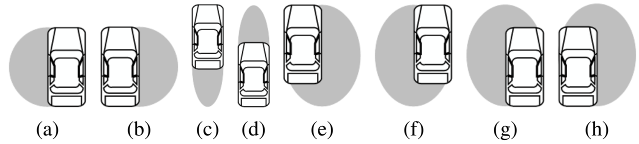


Figure 3: Relative Position of Shadows with respect to Vehicles: Shadows at (a) Left (b) Right (c) Back (d) Front (e) Back and Right (f) Back and Left (g) Front and Left (h) Front and Right

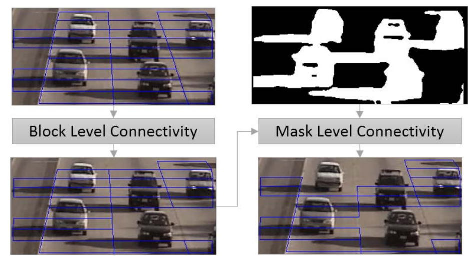


Figure 4: Candidate Shadow Block Filtering based on Shadow Direction

contain a vehicle. We employ this constraint on the expected size of vehicles, to classify such blocks as shadows, as shown in Fig. 2-I.

3.2.2 Shadow Direction

As the camera view is fixed for static traffic surveillance cameras, the expected direction of cast shadows with respect to the vehicle, can be pre-calibrated based on the time of day and the vehicle flow direction. At the block level, we consider 8 possible directions, in which shadows accompany vehicles (shown in Fig. 3). For a given direction, we employ the following connectivity checks to extract the candidate shadow blocks:

Block Connectivity: A foreground block is considered as a candidate shadow only if it is connected with other foreground block(s) in the given shadow direction. For right/left/front/back shadows there should be at least 1 foreground block (representing a possible vehicle) at the left/right/back/front of the candidate shadow block. For the remaining four orientations, a combination of the above rules is applied. For example, in Fig 3e, a combination of the rules for back and right shadows can be used.

Mask Level Connectivity: Block connectivity check results in candidate shadow blocks that represent these cases - (a) two vehicles moving side by side without any shadows, and (b) shadow connected to a vehicle. We therefore employ a check for mask level connectivity, by looking for foreground pixels along the boundary of connected foreground blocks to further refine the candidate shadow blocks (shown in Fig. 4).

3.2.3 Intensity

Shadows are always darker than the background. Therefore, brighter foreground blocks are considered vehicles. In addition, if the foreground block is too dark such that the color/texture of the background is not preserved, we classify them as vehicles as well (e.g. black colored vehicles). For each foreground block B_i, the mean of intensity difference, μ_i, between the input frame I, and the background M, for all the foreground pixels, F is computed as:

$$\mu_i^K = \frac{\sum_{x1}^{x2}\sum_{y1}^{y2}(I^K(x,y) - M^K(x,y)) * (F(x,y))}{255 * N_i} \quad for\ K = R, G, B \tag{1}$$

where $(x1,\ x2)$ and $(y1,\ y2)$ represent the boundaries of B_i and N_i is the total number of foreground pixels in B_i. A foreground block B_i is classified as a candidate shadow block, only if it meets the following criterion: $t_l < \mu_i^K < t_u \quad for\ K = R, G, B$. The thresholds $[t_l,\ t_u]$ are set to [-0.5, 0] based on empirical evaluations.

3.2.4 Coarse Texture

Shadows result in a uniform drop in intensity with respect to the background, unlike dark vehicles that will contain some texture. We distinguish this with a coarse texture feature. For each foreground block B_i, we compute the variance of intensity difference, Var_i between the input frame I, and the background M, for all the foreground pixels, F as:

$$Var_i^K = \frac{\sum_{x1}^{x2}\sum_{y1}^{y2}((I^K(x,y) - M^K(x,y)) * F(x,y))^2}{N_i} - (\mu_i^K)^2 \quad for\ K = R, G, B \tag{2}$$

A foreground block B_i is classified as a vehicle if it meets this criterion: $max(Var_i^R, Var_i^G, Var_i^B) > t_V$, where t_V was set to 1000.

3.2.5 Interior Edges

After employing the coarse feature, only some dark vehicle parts with low texture remain as candidate shadows, as shown in Fig. 2. As shadows do not add any additional textures, absence of edges inside the shadow regions (i.e. interior edges) is a significant cue for shadow elimination, as in [31]. Therefore, we compute an interior edge feature, to further recover remaining vehicle blocks from the shadow candidates, using an interior edge map as shown in Fig. 5. First, the gradient map for the input frame I is computed as E_I. This may contain background edges due to pavement markings/static shadows in addition to the edges due to the vehicle. Therefore, we also compute the gradient map for the background M, represented as E_M and subtract this from E_I, to generate a thresholded difference edge map E_D, that contains only edges contributed by the foreground. In addition, a boundary edge image E_B is generated by applying the Sobel filter on the foreground mask. Finally, the boundary pixels from the boundary edge map E_B are used to erode the foreground edge map, E_D. This removes all boundary edges leaving behind only interior edges. Foreground blocks that contain interior edge pixels ($> 5\%$ of the block) are classified as vehicles as shown in Fig. 2-IV. It should be noted that in Fig. 5, we show the interior edge map calculation for all the foreground areas, for illustration. However, for the proposed method, only the candidate shadow blocks remaining after the previous set of features are applied, are passed through this step.

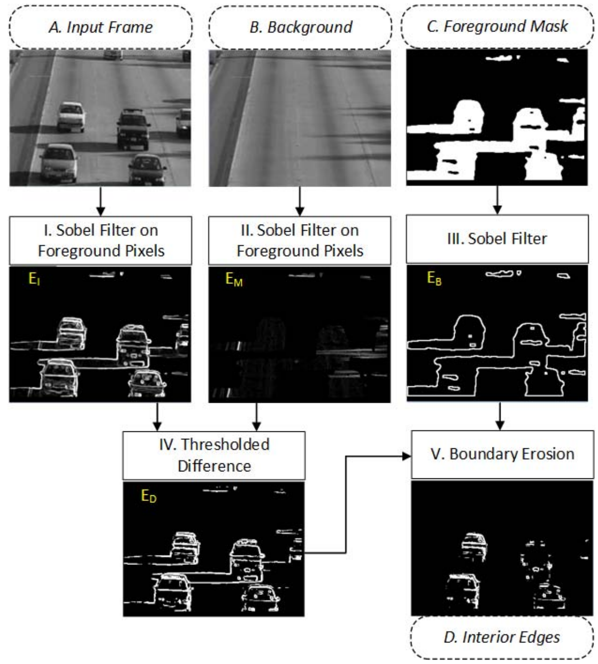


Figure 5: Interior Edge Map Generation

3.2.6 Post-processing

We apply a final post processing step to recover dark and non-textured vehicle blocks, as vehicles. We apply the vehicle size constraint that vehicles should be at least 3 connected blocks. For example, in Fig. 2-IV, a vehicle block in the rightmost lane is mis-detected as a shadow block. As this leads to the recovered vehicle being smaller than its expected size, this block is classified as a vehicle at this step.

4. Results

In this section, we present details about our dataset which covers diverse scenarios, followed by our novel semi-automated ground truth generation technique. Further, we provide details on evaluation metrics and experimental setup, followed by the quantitative and qualitative comparison of the proposed approach with state-of-the-art methods.

4.1. Datasets including Additional Videos for Diverse Scenarios

There is a significant lack of freely available datasets of road traffic scenes for the evaluation of moving shadow elimination techniques. The available datasets are also limited to side shadow cases as shown for the Highway dataset in Table 1. We collected additional videos of real traffic scenes with shadows in different directions, sizes and intensities. In terms of shadow sizes, we chose medium to long shadows that lead to large errors in the estimation of vehicle size. The additional videos also contain shadows in different directions with respect to the vehicle. Since the objects of interests for traffic scenes are vehicles on the road, we limit our region of interest to the road pixels. All these details have been summarized in Table 1.

4.2. Ground Truth Generation Technique

In addition to the lack of publicly available datasets for shadow elimination techniques for traffic scenes, the ground truth available with the datasets is also limited. For example, there are only 10 ground truth annotated frames available for the widely used Highway dataset [25]. In this paper, we propose a semi-automated technique to create pixel-level ground truth to evaluate shadow elimination for accurate vehicle detection, shown in 6.

We employ a deep-learning based object detector YOLO[20] to extract the vehicle bounding boxes. This is combined with the input foreground mask generated by a background modeling technique. The bounding boxes are then overlaid over the foreground masks and pixel-level ground truth mask is generated by marking all foreground pixels inside/outside the bounding box as vehicle/shadow. Since these two steps may not be 100% accurate, we take additional steps to mitigate their impact on the evaluation. We apply median filter on the foreground masks (to fill small holes and remove stray noise), manually remove the frames still containing noise and correct the YOLO bounding boxes.

For our evaluations and for generating ground truth, we use the foreground masks created by state-of-art background modelling techniques. Since the shadow elimination techniques require background image in addition to the foreground mask to operate, the following techniques were chosen from the BGS Library[26]. Independent Multimodal Background Subtraction[6] was used for all the videos except HighwayI as it was unable to generate the background model for this short video. Multilayer[32] was used for HighwayI. These can be treated as a realistic input to any shadow elimination engine. We would like to highlight that

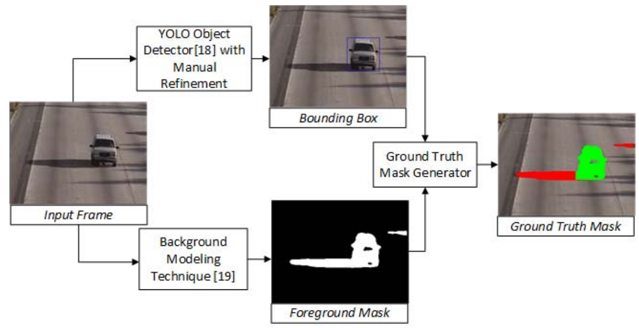


Figure 6: Ground Truth Generation Technique

the same foreground mask is given as an input to all techniques used for evaluation to ensure fair comparison.

4.3. Evaluation Metrics

For the evaluation procedure, similar to [5], we measure the vehicle detection accuracy to evaluate how each shadow elimination technique improves the reliability of vehicle detection. Precision, Recall and F-Measure are used as evaluation metrics for comparing the proposed technique with existing state-of-the-art techniques.

Thus, vehicle/shadow pixels are treated as positive/negative pixels respectively. Precision, Recall and F-Measure are used for evaluation.Recall in vehicle detection is the same as shadow discrimination rate [25], and ideally, a 100% recall (i.e. no missed vehicle parts) is necessary for higher level tasks that rely on vehicle detection and localization.

An average of these metrics over all frames from a video is used to compare with the state-of-the-art techniques. The number of frames used for each video is presented in Table 1. For the comparisons, the implementation in [25] has been used for the baseline shadow elimination techniques based on Chromaticity[8], Physical[12], Small Texture [15], Large Texture[24]. The implementation of the other two state-of-the-art techniques Color Constancy [4] and Tone Mapping [7] is taken from [1] and [3] respectively.

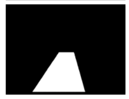
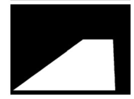
4.4. Experimental Setup

All the algorithms have been executed and verified on a state-of-the-art low cost mobile application development platform - Odroid-XU4 [2] from Hardkernel. This platform contains a Samsung Exynos 5422 mobile SoC that implements ARM big.LITTLE technology with a cluster of four ARM Cortex A15 cores (big cores) and a cluster of four ARM Cortex A7 cores (little cores). In our experimental setup, this platform runs on the ubiquitous Ubuntu 15.10 operating system.

4.5. Quantitative Evaluations

As seen in Fig. 7, the proposed approach achieves better performance than other state-of-the-art techniques for traf-

Table 1: Dataset used for Evaluations

	HighwayI[19]	**Lakeside1**	**Lakeside2**	**PayaLebar1**	**PayaLebar2**	**PayaLebar3**
Frame						
Region of Interest						
Size	320*240	640*480	640*480	640*480	640*480	640*480
Frames	390	1787	1855	1116	671	426
Direction	Left	Back+Left	Front+Left	Left+Back	Back+Right	Front+Right
Shadow Size	Long	Medium	Medium	Long	Medium	Medium
Intensity	Medium	Medium	Medium	Low	High	High

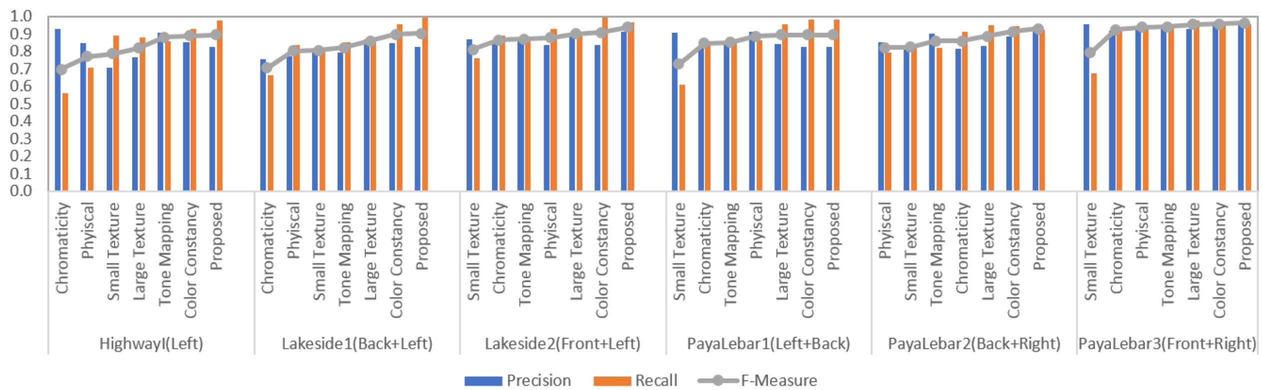


Figure 7: Quantitative comparison with state-of-the-art techniques

fic scenes with front/back+left/right shadows (Lakeside1, Lakeside2, Payalebar2, PayaLebar3), while maintaining a comparable performance for side shadow cases (Highway1, PayaLebar1).

It should be noted that in general, our block-based technique achieves a high recall similar to other region-based techniques like Large Texture [24] and Color Constancy [4]. This is achieved by multi-feature block-based approach. The vehicle sized block assignment and prior information available for our application helps eliminate vehicle blocks on the basis of size and shadow direction. They also provide large enough regions such that other features i.e. intensity, variance and interior edges can be applied effectively to achieve accurate performance.

We would also to highlight that the proposed approach achieves comparable precision to state-to-the-art techniques. We would like to highlight that this slightly low precision value is caused by shadow pixels in the vicinity of vehicle boundaries which do not affect the overall vehicle size estimation drastically.

In addition to the comparison with state-of-the-art techniques, we also evaluated the effectiveness of our cascaded multi-feature pipeline by performing an ablation study in Fig. 8a. This shows that Interior Edge is the strongest feature. Removing other features leads to a fall in recall value. This shows that each feature recovers parts of vehicles which cannot be recovered by the remaining features as shown in Fig. 8c-f. Additionally, Fig. 8b shows that our cascaded approach reduces the number of foreground blocks that need to be classified after each step, thus reducing overall complexity by computing the highly complex feature (Interior Edge) as the last step.

4.6. Qualitative Evaluations

As shown in Fig. 10, the proposed method achieves consistently accurate performance across varying scenarios unlike existing state-of-the-art techniques.

Our evaluations confirm that the pixel based techniques, Chromaticity [8] and Physical [12] (rows 3 and 4), fail in detecting many dark shadows and also detect dark vehicle parts as shadows. This is because they employ color-based features at pixel level, which are not very effective in sunny traffic scenes. In contrast, we do not face this issue, as the classification is performed at a larger vehicle sized block level, similar to region based methods.

Small texture based method [15] (row 5) fails to detect shadow pixels when the road background has low texture. Also, it eliminates parts of vehicles that appear similar to

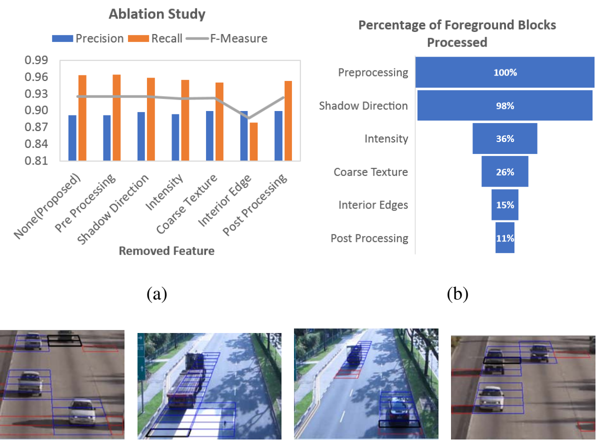


Figure 8: (a) Ablation Study (b) Percentage of Foreground Blocks Processed by each feature; Blocks that can only be recovered by (c) Shadow Direction: recovers parts similar to shadows in the intensity and texture, but not present in the shadow direction. (d) Intensity: recovers parts which have similar texture as the road, but different mean intensity. (e) Coarse Texture: recovers parts which have texture change near the boundary pixels, thus, not contributing to interior edges. (f) Post Processing: recovers parts that are present in shadow direction, look similar to shadows but violate the vehicle size constraint (Best viewed in color)

the background, as shadows. Similarly, the block-based technique, Tone-Mapping[7] (row 8) also misses out parts of vehicles. This is because the evaluation patches for these methods are too small compared to the size of a vehicle. Unlike these techniques, we use larger regions with a size relative to our foreground (i.e. vehicles) which makes us resilient to such classification errors.

As expected, techniques that evaluate larger regions Large Texture[24] and Color Constancy[4] achieve higher accuracy without missing parts of vehicle. However, they fail to detect shadows completely for some cases as seen in row 6 and 7. This is due to ineffective segmentation into regions. In contrast, the top down vehicle size block assignment leads to crude segmentation of the foreground, which can be - vehicle blocks, vehicle+shadow blocks and shadow blocks. Thus, the false positive detections, due to segmentation process, are limited to shadow pixels inside the vehicle+shadow blocks.

4.7. Accuracy vs Speed Trade-off

In addition to the robustness, we have also compared the speed (achieved frame rate) of the proposed method with existing state-of-the-art techniques. Fig. 9 shows the accuracy vs speed trade-off for all the discussed techniques. All the techniques were implemented on the Odroid-XU4 platform in C++ and the average performance across all testing videos is presented. Tone Mapping [7] is excluded from this evaluation as its implementation is in Matlab.

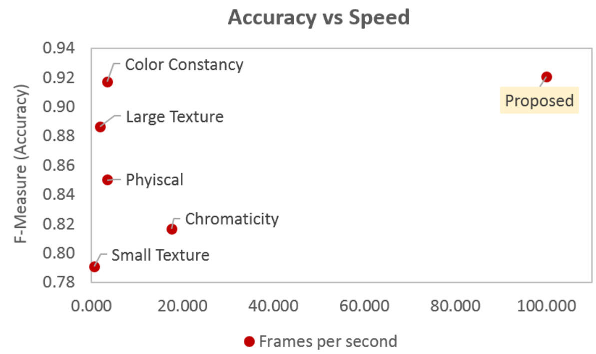


Figure 9: Accuracy vs Speed Trade-off comparsion

It can be seen that we are able to achieve the best trade-off in terms of accuracy and speed compared to all the other techniques. This is attributed to the block-level classification approach that does not involve the heavy computations incurred by explicit segmentation as in [24, 4]. In addition, the cascaded approach ensures that majority of the distinct vehicle blocks are classified using computationally efficient aggregate features. The robustness is still maintained by employing the more sophisticated internal edge feature for only those vehicle blocks that are similar to shadows.

5. Conclusion

In this paper, we introduced a novel technique to eliminate moving shadows for traffic surveillance scenes, which is an essential step for fast and accurate detection of vehicles on the road. The proposed method incorporates a block based shadow elimination technique that uses multiple features(i.e., size, color and texture) in a cascaded fashion to classify foreground blocks into vehicle and shadow blocks. We also proposed a technique to generate ground truth and evaluate shadow elimination techniques in a speedy yet effective manner. Extensive evaluations using multiple traffic videos on a low cost embedded platform confirmed that the proposed techniques not only achieve higher performance, in terms of accuracy, but also execute faster than existing state-of-the-art techniques for varying vehicle-shadow orientations. In future, we would like to further improve the precision of the proposed technique by eliminating shadow pixels in the vicinity of vehicle boundaries.

Acknowledgment

This research project is funded by the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme with the Technical University of Munich at TUMCREATE.

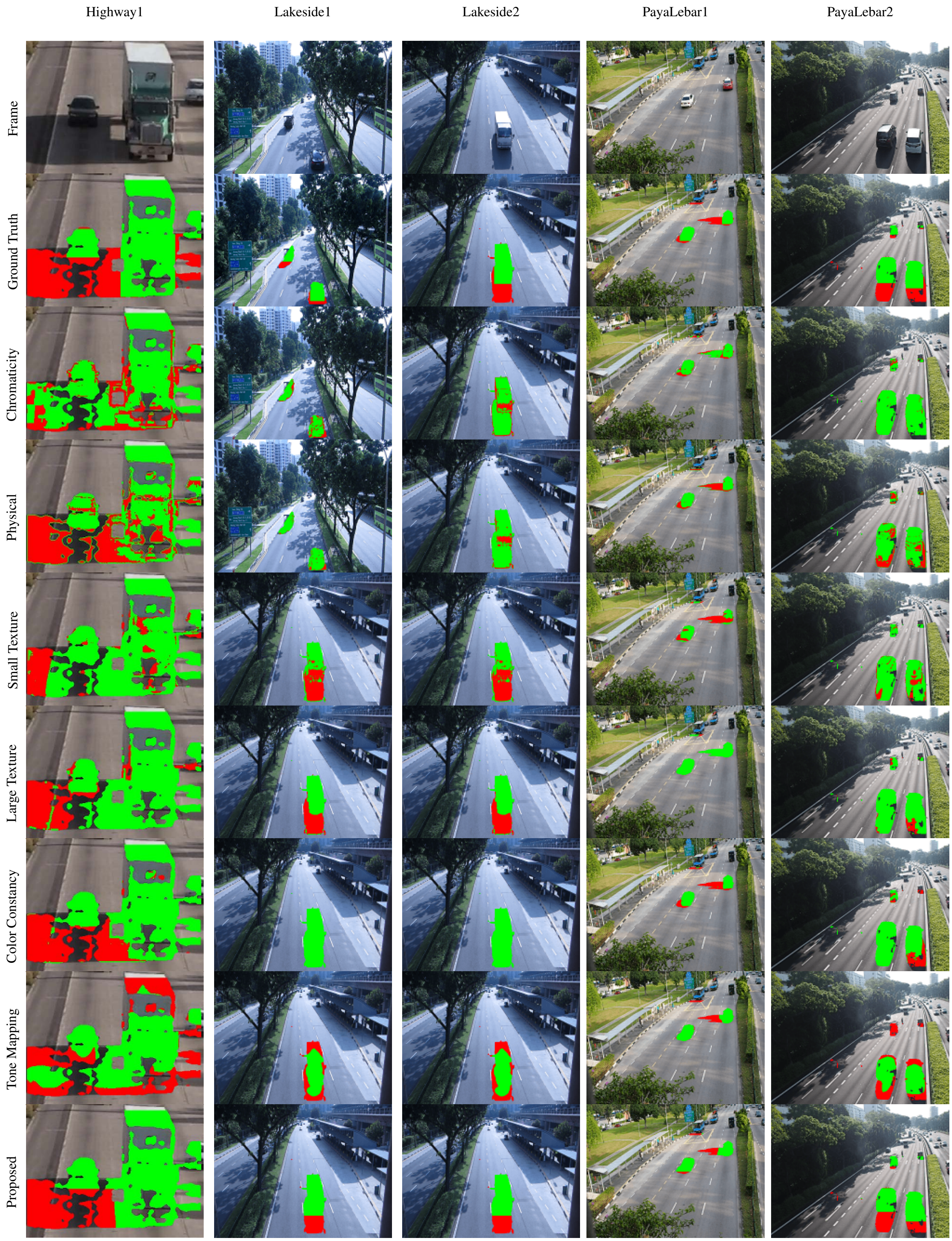


Figure 10: Qualitative Comparison with state-of-the-art techniques [Green:Vehicle, Red:Shadow (Best viewed in color)]

References

[1] Color Constancy Implementation. https://github.com/nicklhy/MovingShadowRemove.

[2] Odroid-XU4. http://goo.gl/Nn6z3O.

[3] Tone Mapping Implementation. https://www.mathworks.com/matlabcentral/fileexchange/36949-moving-shadow-detection.

[4] A. Amato et al. Accurate moving cast shadow suppression based on local color constancy detection. *IEEE Transactions on Image Processing*, 2011.

[5] P. Barcellos et al. Shadow detection in camera-based vehicle detection: survey and analysis. *Journal of Electronic Imaging*, 2016.

[6] D. Bloisi et al. Independent multimodal background subtraction. In *CompIMAGE*, 2012.

[7] E. Bullkich et al. Moving shadow detection by nonlinear tone-mapping. In *International Conference on Systems, Signals and Image Processing*, 2012.

[8] R. Cucchiara et al. Detecting moving objects, ghosts, and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003.

[9] J. Dai et al. Region-based moving shadow detection using affinity propagation. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2015.

[10] K. Garg et al. Low complexity techniques for robust real-time traffic incident detection. In *IEEE Conference on Intelligent Transportation Systems*, 2017.

[11] J.-W. Hsieh et al. Shadow elimination for effective moving object detection by gaussian shadow modeling. *Image and Vision Computing*, 2003.

[12] J.-B. Huang et al. Moving cast shadow detection using physics-based features. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[13] S. Jiang et al. WeSamBE: A weight-sample-based method for background subtraction. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.

[14] M. Khare et al. Shadow detection and removal for moving objects using Daubechies complex wavelet transform. *Multimedia Tools and Applications*, 2018.

[15] A. Leone and C. Distante. Shadow detection for moving objects based on texture analysis. *Pattern Recognition*, 2007.

[16] L. A. Lim and H. Y. Keles. Foreground segmentation using a triplet convolutional neural network for multiscale feature encoding. *arXiv preprint arXiv:1801.02225*, 2018.

[17] S. K. Meher et al. Efficient method of moving shadow detection and vehicle classification. *AEU-International Journal of Electronics and Communications*, 2013.

[18] K. Nagarathinam et al. Moving shadow detection based on stationary wavelet transform and Zernike moments. *IET Computer Vision*, 2018.

[19] A. Prati et al. Detecting moving shadows: algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003.

[20] J. Redmon et al. You only look once: Unified, real-time object detection. In *IEEE conference on computer vision and pattern recognition*, 2016.

[21] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017.

[22] M. Russell et al. An evaluation of moving shadow detection techniques. *Computational Visual Media*, 2016.

[23] M. Russell et al. Feature-based image patch classification for moving shadow detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017.

[24] A. Sanin et al. Improved shadow removal for robust person tracking in surveillance scenarios. In *International Conference on Pattern Recognition*, 2010.

[25] A. Sanin et al. Shadow detection: A survey and comparative evaluation of recent methods. *Pattern Recognition*, 2012.

[26] A. Sobral. BGSLibrary: An opencv c++ background subtraction library. In *IX Workshop de Viso Computacional*, Rio de Janeiro, Brazil, 2013.

[27] P.-L. St-Charles, G.-A. Bilodeau, and R. Bergevin. Universal background subtraction using word consensus models. *IEEE Transactions on Image Processing*, 25(10):4768–4781, 2016.

[28] P.-L. St-Charles et al. Subsense: A universal change detection method with local adaptive sensitivity. *IEEE TIP*, 2015.

[29] Y. Wang et al. CDnet 2014 :An expanded change detection benchmark dataset. *IEEE CVPRW*, 2014.

[30] Y. Wang et al. Interactive deep learning method for segmenting moving objects. *Pattern Recognition Letters*, 2017.

[31] M. Xiao et al. Moving shadow detection and removal for traffic sequences. *International Journal of Automation and Computing*, 2007.

[32] J. Yao et al. Multi-layer background subtraction based on color and texture. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[33] A. Yoneyama et al. Robust vehicle and traffic information extraction for highway surveillance. *EURASIP Journal on Advances in Signal Processing*, 2005.