

A Mathematical Morphology based Approach for Vehicle Detection in Road Tunnels

Andrés Frías-Velázquez, Jorge Oswaldo Niño-Castañeda, Vedran Jelača, Aleksandra Pižurica and Wilfried Philips

Ghent University-TELIN-IPI-IBBT, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium

ABSTRACT

A novel approach to automatically detect vehicles in road tunnels is presented in this paper. Non-uniform and poor illumination conditions prevail in road tunnels making difficult to achieve robust vehicle detection. In order to cope with the illumination issues, we propose a local higher-order statistic filter to make the vehicle detection invariant to illumination changes, whereas a morphological-based background subtraction is used to generate a convex hull segmentation of the vehicles. An evaluation test comparing our approach with a benchmark object detector shows that our approach outperforms in terms of false detection rate and overlap area detection.

Keywords: Vehicle detection, Road tunnels, Mathematical morphology, Uneven illumination.

1. INTRODUCTION

In spite of the fact that the number of accidents is lower in tunnels than in open stretches of roads, the risk of being killed in a tunnel accident is twice as high according to a study conducted by the Austrian Road Safety Board.¹ The study reveals that driving through a confined space is the cause of increasing the fatality risk in a tunnel. This statement is based on the fact that vehicle collisions turn to be more violent when they end up against the tunnel walls, causing severe injuries. Fires are also a major concern, since explosions and toxic gases may yield catastrophic consequences.

Closed-Circuit Television (CCTV) systems are normally deployed to enforce the security in tunnels. However, the tunnel length may reach up to 25 km, which may require a large number of monitoring personnel. In order to better assist tunnel operators and prevent accidents, Incident Detection Systems (IDS) are becoming essential in tunnel surveillance. The technology behind incident detection systems heavily depends on image processing and computer vision techniques such as vehicle detection.

A vast number of vehicle detection methods have been proposed so far for different applications related to traffic surveillance. These methods have been surveyed by Kastrinaki et al.² and more recently by Buch et al.³ It is worth to mention that most of the reviewed approaches were tested under relatively low traffic density, day-light and normal weather conditions. In recent years, on the other hand, vehicle detection techniques under real scenarios and challenging conditions have been studied more extensively. Poorly illuminated environments have been subject to analysis for different applications, such as traffic monitoring⁴⁻⁶ and on-vehicle driving assistance.⁷⁻⁹ In overall, these techniques strongly depend on background subtraction and template matching approaches to extract the head and rear lights from the background, and thus detect the candidate vehicles. Geometrical information of the vehicles such as the symmetry of the rear lights, as well as the height/width aspect ratio is also exploited. Furthermore, lane and spatial calibration information also provide relevant data for the case of stationary camera monitoring. This extra evidence is often used by the classifier or reasoning engine in order to obtain better detection results.

Non-uniform and poor illumination conditions prevail in road tunnels making difficult to achieve robust vehicle detection with either standard object or motion detectors. Apart of the scarce tunnel lighting, the head and rear lights of the vehicles are other important sources of illumination, which may drastically change the

Further author information: (Send correspondence to Andrés Frías-Velázquez)
A.F.V.: E-mail: Andres.FriasVelazquez@telin.ugent.be, Telephone: +32 9 264 4226

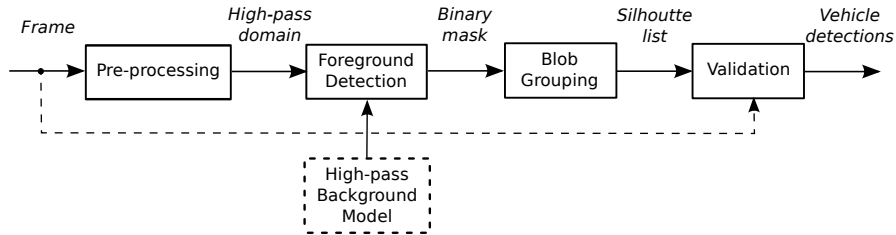


Figure 1: Proposed vehicle detection system.

scene conditions, and thus the appearance of the vehicles. Also, shadows and reflections are an issue, since they cause a considerable number of false detections.

The block diagram of the proposed vehicle detection system is depicted in Fig. 1. An image pre-processing technique is the first and key step of our approach, which consists on applying a local higher-order statistic filter in order to cope with the illumination issues and pass the image into a high-pass domain. The resulting image serves as input to a foreground detector, which performs a morphological-based background subtraction between the pre-processed image and a high-pass background model of the tunnel. This last step returns blobs of the vehicles detected, which are clustered by using information like centroid coordinates, area, and lane position. The vehicle’s silhouettes are approximated by getting the convex hull of the clustered segments. Finally, a vehicle validation test is run to discard false detections.

The rest of the paper is organized as follows: In Section 2, we present an analysis of the photometrical tunnel conditions. Section 3 details the processing blocks of the proposed system. While in Section 4, an evaluation of the system is presented. Finally, the conclusions of this work are given in Section 5.

2. ANALYSIS OF THE TUNNEL CONDITIONS

In order to provide a more specific context of a tunnel environment, we present in Fig. 2 some frames that exhibit typical problems encountered when vehicle detection is performed over a tunnel. We can start analyzing Fig. 2a, where some illumination and vehicle recognition issues are highlighted. For instance, the ovals A and B stand out the rear and lateral lights of a cargo truck, respectively. These detections are often mistakenly recognized as two different vehicles by template-based matching approaches.¹⁰ The lack of success of these methods under these conditions is normally caused by their strong dependence on the rear lights and the context loss of the vehicle’s bodywork. It is worth to remark that due to the constrained height of the tunnels, cameras and vehicles (especially trucks) are relatively close each other. Consequently, the robustness of these methods at different scales is put to test, since large vehicles cover a large percentage of the field of view.

On the other hand, the image segments labelled as C and D show some illumination issues that are common in tunnels. For instance, the large metallic bodywork of trucks generates light reflections from other vehicles (e.g. segment C), which in turn look like rear lights and thus being mistakenly recognized as a vehicle itself. Shadows, as shown in segment D, are also a typical phenomenon that disturbs the scene conditions. The main consequence is the overestimation of the detected area when traditional background subtraction techniques are employed.

Other common problem encountered in tunnels is illustrated in Fig. 2b. In violation of traffic regulations, some drivers do not switch their head and rear lights on, which makes really difficult to detect them under poor illumination conditions. Moreover, some vehicles can be missed when the vehicle color resembles to that of the road. Conversely, in the segment A (Fig. 2c), we can observe that the headlights heavily lit certain areas of the road creating potential false positive detections. Finally, segment B shows that the shadows created by large trucks affect the appearance of neighbouring vehicles.

Due to the minimal illumination conditions in tunnels, the vehicle head and rear lights play a key role in vehicle detection. On the one hand, both head and rear lights help to better detect and track the vehicles along the tunnel. On the other hand, they introduce local illumination variations on the scene, reflections over the road and other vehicles, turning the vehicle detection problem more challenging.

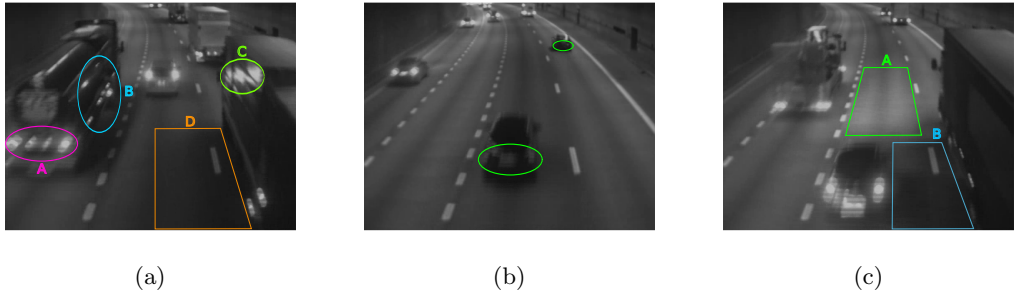


Figure 2: Several frames demonstrating the illumination conditions in a tunnel.

3. PROPOSED SYSTEM

The surveillance system considered consists of several cameras disposed along a tunnel to capture the rear side of the vehicles, as shown in Fig. 2. The cameras retrieve 8-bit grayscale video at a resolution of 720 x 576 @25fps. The intention is to solely detect the incoming vehicles that enter the field of view of the camera, and then use this detection as initialization for a vehicle tracking system.

3.1 Image Pre-processing

In general, traffic surveillance images such as those captured in tunnels, are characterized by large flat grayscale zones. As a result, the background scene contains little texture information, which allows relatively easy detect the presence of richer texture content in the scene. By giving a great importance to the fact that texture or high frequency information allows determining the presence of an object, it is thus intuitive to translate our workspace into a high-pass domain. In order to extract the texture content, we propose a local higher-order statistic filter, which is inspired on a statistical feature called kurtosis. The kurtosis is defined as follows

$$\beta_2 = \frac{\mu_4}{\sigma^4} \quad (1)$$

where μ_4 is the fourth central moment and σ is the standard deviation. Such feature measures the *peakedness* of a probability distribution. For instance, a uniform distribution has kurtosis of 1.8, while for the Gaussian and Laplace distribution is 3 and 6, respectively. Hence, the kurtosis value increases as sharper the peak that models the distribution. Unfortunately, we can not employ this definition straightforwardly because the kurtosis of flat zones returns an indeterminate form (i.e. 0/0). Since the intention is to apply the kurtosis in a local fashion, the sliding kernel will often find flat zones due to the nature of the background scene. Consequently, we adopt a definition of the kurtosis based on *non-central moments* as follows

$$\theta_2 = \begin{cases} 0 & \text{if } \forall i, j : x_{ij} = 0 \\ \frac{\sum_{i=1}^n \sum_{j=1}^m w_{ij} x_{ij}^4}{(\sum_{i=1}^n \sum_{j=1}^m w_{ij} x_{ij}^2)^2} - 1 & \text{otherwise} \end{cases} \quad (2)$$

where x_{ij} represents the grayscale value of a pixel located at the indices i and j of the evaluation window of size $n \times m$. While w_{ij} is a weighting kernel. In this way, indeterminate forms are avoided, with the sole exception when $\forall i, j : x_{ij} = 0$. Since the proposed statistical filter does not employ central moments, the resulting values of θ_2 should not be compared with those obtained using (1). Furthermore, the response of the proposed filter is inverse to the standardized kurtosis. As a result, θ_2 returns higher values as the texture richness increases. In spite of these differences, the ability of our approach to discriminate between poor and rich texture is kept.

In order to provide a glance at the proposed filter response, we present Fig. 3. In this figure, we also compare our approach with a high-pass and another statistical filter with a 5×5 size kernel. For instance, in Fig. 3a we can observe the high-pass response of Fig. 2a by using a well-known derivative approach called Laplacian of Gaussian. Due to the uneven illumination conditions, the response of the Laplacian of Gaussian is far to be uniform. As a result, the blob detection is forced to be done using critical thresholds. The side effects of this

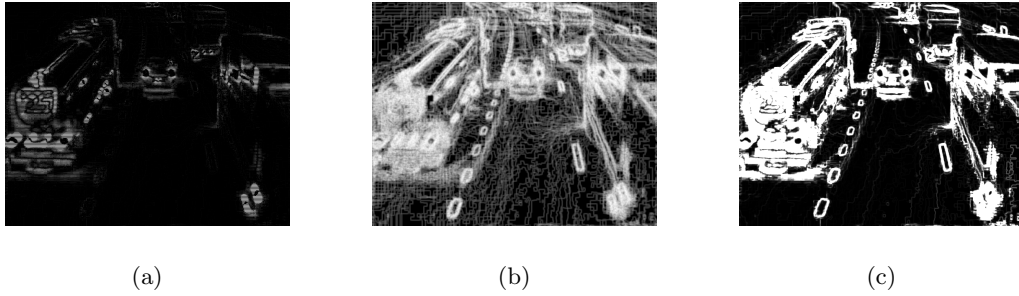


Figure 3: Comparison of several blob detectors.



Figure 4: Background model.

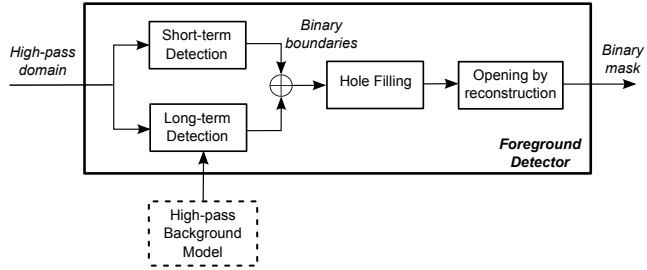


Figure 5: Block diagram of the foreground detector.

approach are false detections caused by noise and/or loss of informative parts of the vehicle. On the other hand, in Fig. 3b we present the resulting response of using a local entropy filter. As is known, the entropy measures the average information contained in a random variable, which is influenced by the shape of its distribution similarly as the kurtosis. Note that this statistical filter is more robust against uneven illumination than the Laplacian or Gaussian. Moreover, the vehicle's edges and texture response is more consistent; however, the background's texture information is also brought out. As a result, the blob detection still depends on a critical threshold that trade-off between salient information and noisy responses. Finally, in Fig. 3c we present the response of the proposed filter described in (2). By visually comparing the three presented approaches, we can easily notice that the kurtosis-based approach outperforms in terms of illumination invariance and noise suppression. The outcoming high-pass image is practically binary, and the blob detection can be obtained *without* a critical threshold. In this way, the edge and texture information is well preserved while the noisy responses of the background are suppressed, and thus easily removed by thresholding. It is worth to note that in comparison to the entropy filter, the proposed filter returns a value that does not depend on the parameters of the distribution model. This fact might be one of the causes that explains the good performance of the filter, since the kurtosis actually measures the shape of the distribution independently of the data dispersion.

3.2 Foreground Detection

3.2.1 Background Model

The grayscale background model is generated by averaging several frames with no traffic in its field of view. The intent is to generate an static scene as well as reducing the noise and temporal lighting variations. Then, the statistical filter described above (2) is applied to the background to translate it into a high-pass domain. The response of the background model is depicted in Fig. 4. Note that most of the high-frequency content detected comes from the road marks, while the rest of the background returns a very low response given that these areas are locally composed of quasi-flat regions. It is worth to point out that given the camera view is fixed, and the background model is now robust enough to illumination variations, there is no need to update the background model.

3.2.2 Algorithm

As discussed in previous sections, the proposed foreground detection algorithm operates in a high-pass domain due to the nature of the background model (Fig. 4) and the pre-processed input frames (e.g. Fig. 3c). The response in the high-pass domain is basically composed by edges and textured regions. Consequently, it is of utmost importance to preserve the edges since they convey salient information related to the shape of the vehicles.

In Fig. 5, we present the block diagram of the foreground detection system. The core of the algorithm consists in fusing the frame-by-frame motion detection (*short-term detection*) with the motion resulting from the difference of the incoming frames respect to a static background (*long-term detection*). The short-term detection is basically derived from frame differencing as shown in the lines 1 to 3 of Algorithm 1. Where G_k is the pre-processed frame at the discrete-time instance k , while the *hyst* function denotes the hysteresis thresholding with lower and upper threshold values of α_1 and τ_1 , respectively. Note that we use a three-frame differencing approach based on the forward and backward detection in order to estimate the motion D_{short} at instance k .

On the other hand, the long-term detection is described in lines 4 to 9 of Algorithm 1. In line 4, the subtraction between the pre-processed frame G_k and the static background B_s is performed. By inspecting this operation with Fig. 3c and Fig. 4, we can notice that the road marks will be removed, but at the same time some parts of the vehicle detected. This last effect is not desirable because it breaks the link of connected components, such as edges. Consequently, it is not possible to determine the foreground in the high-pass domain when a vehicle detection and a road mark intersect. We have to appeal to the original grayscale images to solve this ambiguity. Firstly, we create a binary mask W (line 5) telling us where the ambiguity problem is. This mask is computed by thresholding $\mathcal{T}_\lambda(\cdot)$ the intersection of the frame and the static background. Then, in line 6 we apply the binary mask over the difference between the original frame F_k and the grayscale background model B_g . In this way, the union of L_1 and L_2 reconnects the broken edges by the ambiguity and the resulting image is passed to binary by applying the hysteresis thresholding (line 7).

Unfortunately, previous steps also lead to false detections; however, they can be easily removed by using morphological reconstruction. Specifically, L_{msk} works as the *mask*, while the *marker* L_{mrk} is obtained by morphological opening γ_{se} as shown in line 8. Finally, we apply the morphological reconstruction by dilation \mathcal{R}^δ to remove the false detections and return the long-term detection D_{long} .

Having obtained the short and long-term detections, both sources of motion are merged, as shown in line 10. Then, in line 11, the closed contours are converted into blobs by using the *fillhole* morphological operator (FILL). Finally, the foreground estimation D_{blob} is computed by applying an opening by reconstruction $\gamma_{\mathcal{R}}^{(q)}$ with size q . This last step aims to ensure the removal of false detections.

Algorithm 1 Foreground detection algorithm

1: $U \leftarrow G_{k-1} - G_k \wedge G_{k+1} - G_k $	<i>Begin short-term detection</i>
2: $V \leftarrow (G_{k-1} \wedge G_k) - (G_{k+1} \wedge G_k) $	
3: $D_{short} \leftarrow \text{hyst}(U \vee V, \alpha_1, \tau_1)$	<i>End short-term detection</i>
4: $L_1 \leftarrow G_k - B_s $	<i>Begin long-term detection</i>
5: $W \leftarrow \mathcal{T}_{\lambda_1}(G_k \wedge B_s)$	
6: $L_2 \leftarrow F_k - B_g W$	
7: $L_{msk} \leftarrow \text{hyst}(L_1 \vee L_2, \alpha_2, \tau_2)$	
8: $L_{mrk} \leftarrow \mathcal{T}_{\lambda_2}(\gamma_{se}(L_1))$	
9: $D_{long} \leftarrow \mathcal{R}_{L_{msk}}^\delta(L_{mrk})$	<i>End long-term detection</i>
10: $D \leftarrow D_{short} \vee D_{long}$	
11: $D_{fill} \leftarrow \text{FILL}(D)$	
12: $D_{blob} \leftarrow \gamma_{\mathcal{R}}^{(q)}(D_{fill})$	

3.3 Blob Grouping

After generating the blobs resulting from the foreground detection, we can now proceed to group them to compose the final vehicle detections. Despite the efforts made by the foreground detection algorithm to preserve the connectivity of the vehicle’s detections, it may turn out that some of them are fragmented in several blobs. This problem is mostly experienced by truck detections because they occupy a larger space in the field of view than passenger cars. Although the foreground detector returns fragmented blobs, these detections are more less compactly distributed, which allows to group them easily. In other words, fragmented vehicle detections are normally composed by one or two large blobs surrounded by smaller detections.

In order to merge the blobs that compose a vehicle detection, we use a hierarchical clustering¹¹ approach. Specifically, an agglomerative method that allow us to first merge the small blobs with the larger ones, and then identify the disjoint blobs that correspond to the actual vehicles. The clustering approach is built with features like centroid coordinates, area, and lane position extracted from the blobs. Note that the hierarchical clustering is suitable for this application since the number of clusters are not known a priori. Moreover, the compactness of the clusters enable using a dissimilarity measure as a parameter of coherence to identify disjoint groups of detections. Finally, the vehicle boundaries are obtained by finding the convex-hull resulting from the blobs merged by the hierarchical clustering.

3.4 Validation

As reviewed in Sec. 2, there are several illumination issues that may cause false positive/negative detections. Regardless the illumination invariance provided by the pre-processing filtering, some false positives may still take place mostly due to the reflections on the road. In order to discard these false detections is necessary to be certain that the clustering output actually correspond to a vehicle. As a result, we trained a Bayesian classifier with projection signatures extracted from good and bad exemplar detections. When testing, the classifier is biased to discard as much false positive detections as possible without dismissing good detections (false negative detections).

4. EXPERIMENTAL RESULTS

Some experiments were carried out in order to evaluate the performance of the proposed system and an object detector¹² that was trained to recognize vehicles in a tunnel. The test sequence contains 260 vehicles, where 186 are cars and 74 are trucks. The detection results are analyzed on the basis of correct detections, false positives and negatives. The assessment not only takes into account the number of detections, but also the detected area respect to its ground truth. In Table 1, the performance of the vehicle detectors is evaluated using the precision and recall scores, which are defined as follows

$$Precision = \frac{N_c}{N_c + N_{fp}} \times 100 \quad (3) \qquad Recall = \frac{N_c}{N_c + N_{fn}} \times 100 \quad (4)$$

where N_c stands for the number of correct detections, while N_{fp} and N_{fn} represent the number of false positive and negative detections, respectively. A false positive occurs when a vehicle detection is considered as valid when in fact is not. On the other hand, false negative detections represent the missed vehicles.

Assessing the recall score displayed in Table 1, we can see that our approach achieves a better performance than the object detector. The recall score asserts that the object detector failed to recognize several vehicles, while our approach did not miss any of them. Similarly, our approach also stood out in terms precision since the number of false positives was much lower than the object detector. The reason why our approach outperforms in both scores is due to the combination of the kurtosis filter, which deals with illumination issues, and the Bayesian classifier that minimizes the false positive detections. In this way, the kurtosis filter unveils salient information that other approaches discard due to uneven illumination, while reducing the risk of false negatives.

Solely evaluating the number of correct detections does not reveal whether the detected area is close to the actual detection (ground truth) or not. For this reason, it is necessary to undergo an assessment that provides information on the tightness of the detections. As a result, such information is presented in Table 2 based on the average recall and precision scores. Note that the presented percentages were computed using similar definitions

as the ones shown in (3) and (4). However, these equations are now evaluated in terms of intersecting area (correct), missing area (false negative), and overestimated area (false positive). In this way, Table 2 reveals the completeness of our approach since the recall is very high and greater than the one given by the object detector. This implies that the blobs retrieved by the foreground detector overlap at a large percentage of the detected area with their ground truth. Conversely, the object detector may return blobs that are missing important parts of the vehicles.

In terms of precision, the object detector performs slightly better than our approach because most of the background subtraction detections are overestimated due to the shadows that could not be completely removed by the kurtosis filter. Moreover, it should be noted that the precision score achieved by the object detector was reached due to the considerable number of underestimated detections, as its recall score suggests. In this way, several incomplete detections biased positively the precision score since their precision was 100%.

From the tables presented, we can conclude that the kurtosis filter and the morphological background subtraction were able to cope with the major illumination issues encountered in tunnels. The performance of our approach can be summed up by stating that the system is able to detect all the vehicles without introducing a significant number of false positives. Moreover, the average detected area is overestimated around 37.7% respect to the ground truth. However, the integrity of the vehicle detection is preserved. Conversely, the object detector¹² ignores 13% of the vehicles and one out of five detections correspond to a false positive. Assessing the detections against their ground truth is shown that approximately 20% of the detected area is out of the ground truth, while the completeness of the vehicle detection is not guaranteed.

Table 1: Vehicle detection recall/precision.

<i>Method</i>	% <i>Percentage Detection</i>	
	Recall	Precision
Object Detector	87.04	79.66
Foreground Detector	100	98.86

Table 2: Blob detection recall/precision.

<i>Method</i>	% <i>Percentage</i>	
	Recall	Precision
Object Detector	75.87	80.04
Foreground Detector	97.82	72.62

5. CONCLUSIONS

A vehicle detector suited for video surveillance in tunnels has been presented in this paper. The proposed method strongly relies on a local higher-order statistic filter, which plays a key role in the overall performance of the system. The advantage of using this filter was twofold. On the one hand, the filter provided illumination invariance due to shadows, reflections and low lighting. On the other hand, it returned a sharp edge response and stood out textured information. In this way, the foreground detection was obtained without compromising the actual vehicle detection with noisy responses because of critical thresholds.

It is also worth to note that the proposed morphological background subtraction and the Bayesian classifier played a relevant role in terms of recall and precision. The fact that the foreground detector ensures the connectivity and compactness of the detections minimizes the risk of false detections. Also, it establishes the completeness and tightness properties of the detections. On the other hand, the Bayesian classifier enforces the rejection of false positive detections overlooked by the foreground detector.

Finally, the performance achieved by the proposed system complies with the precision/recall requirements that an incident detection system may ask for its optimal performance. This fact is important because other IDS components such as vehicle trackers and identification algorithms depend on the input provided by the vehicle detector. Consequently, the propagation of errors in subsequent stages may compromise the performance of the entire incident detection system.

ACKNOWLEDGMENTS

This work has been supported by the Flemish Interdisciplinary Institute for Broadband Technology (IBBT).

REFERENCES

- [1] Naussbaumer, C., “Comparative analysis of safety in tunnels,” in [*Young Reseachers Seminar 2007*], European Conference of Transport Research Institutes (May. 2007).
- [2] Kastrinaki, V., Zervakis, M., and Kalaitzakis, K., “A survey of video processing techniques for traffic applications,” *Image and Vision Computing* **21**, 359–381 (2003).
- [3] Buch, N., Velastin, S. A., and Orwell, J., “A review of computer vision techniques for the analysis of urban traffic,” *IEEE Transactions on Intelligent Transportation Systems* (99), 1–20 (2011).
- [4] Robert, K., “Night-time traffic surveillance: A robust framework for multi-vehicle detection, classification and tracking,” in [*Proc. of the Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*], *AVSS '09*, 1–6 (2009).
- [5] Wu, B.-F., Kao, C.-C., Liu, C.-C., Fan, C.-J., and Chen, C.-J., “The vision-based vehicle detection and incident detection system in hsueh-shan tunnel,” in [*IEEE International Symposium on Industrial Electronics, ISIE 2008*], 1394–1399 (30 2008-july 2 2008).
- [6] Chen, Y.-L., Wu, B.-F., Huang, H.-Y., and Fan, C.-J., “A real-time vision system for nighttime vehicle detection and traffic surveillance,” *IEEE Transactions on Industrial Electronics* **58**, 2030–2044 (may 2011).
- [7] Alt, N., Claus, C., and Stechele, W., “Hardware/software architecture of an algorithm for vision-based real-time vehicle detection in dark environments,” in [*Proc. of the conference on Design, Automation and Test in Europe*], *DATE '08*, 176–181, ACM (2008).
- [8] O’Malley, R., Jones, E., and Glavin, M., “Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions,” *IEEE Transactions on Intelligent Transportation Systems* **11**, 453–462 (june 2010).
- [9] Kim, S., Oh, S.-Y., Kang, J., Ryu, Y., Kim, K., Park, S.-C., and Park, K., “Front and rear vehicle detection and tracking in the day and night times using vision and sonar sensor fusion,” in [*Proc. International Conference on Intelligent Robots and Systems (IROS 2005)*], 2173–2178 (aug. 2005).
- [10] Sidla, O., Paletta, L., Lypetsky, Y., and Janner, C., “Vehicle recognition for highway lane survey,” in [*Proc. of the 7th International IEEE Conference on Intelligent Transportation Systems*], 531–536 (oct. 2004).
- [11] Hastie, T., Tibshirani, R., and Friedman, J., [*The Elements of Statistical Learning: Data Mining, Inference, and Prediction*], Springer (2003).
- [12] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., and Ramanan, D., “Object Detection with Discriminatively Trained Part-Based Models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**, 1627–1645 (Sept. 2010).