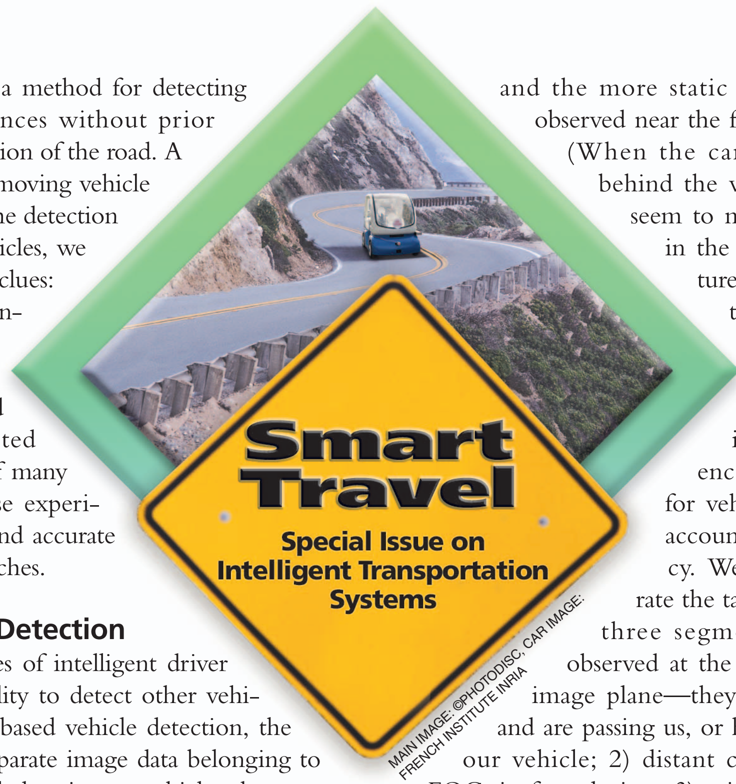
Vehicle Detection with a Mobile Camera

Spotting Midrange, Distant, and Passing Cars



In this article, we present a method for detecting vehicles in image sequences without prior knowledge about the position of the road. A single camera placed in a moving vehicle provides image data. For the detection of midrange and distant vehicles, we use a combination of three clues: shadow, entropy, and horizontal symmetry. To detect passing vehicles, we use temporal differencing and projected motion. We tested our algorithms by means of many different experiments. These experiments illustrate the robust and accurate performance of both approaches.

Vision-Based Vehicle Detection

One of the essential features of intelligent driver assistance systems is the ability to detect other vehicles on the road. In vision-based vehicle detection, the system should be able to separate image data belonging to the background from data belonging to vehicles that potentially limit our motion possibilities. Detection is a two-step process. In the first step, all regions in the image plane that potentially contain a vehicle are identified. In the next step, the selected regions are tracked in time. Based on their observed motion, false detections are removed.

Detection is the process of identifying the presence of clues in the image plane. The characteristics of these clues depend heavily on both the lighting conditions and the position in the image plane where the vehicle is observed. In this article, we focus on daytime operation.

The appearance of vehicles in the image plane changes with its location. One might think of the geometric transformation of a vehicle when observed in the corner of our eye and the more static shape of a vehicle when observed near the focus of contraction (FOC). (When the camera observes the scene behind the vehicle, background points seem to move towards a single point in the image plane. In the literature, this point is referred to as the FOC. Note that for the problem of vehicle detection, the viewing direction of the camera introduces only little differences.) When we select clues for vehicle detection, we have to account for this spatial dependency. We, therefore, roughly separate the task of vehicle detection into three segments: 1) passing cars are observed at the left or right border of the image plane—they are viewed from the side and are passing us, or have just been overtaken by our vehicle; 2) distant cars are observed near the FOC, in frontal view; 3) midrange cars are observed in the area between the regions of passing cars and distant cars; limited geometric deformation might occur.

In the next section, we explain our procedure for detecting midrange and distant vehicles. We then describe the procedure for passing vehicle detection and present the results of two experiments. This article concludes with a discussion.

Detecting Midrange and Distant Vehicles

Shadow, entropy, and horizontal symmetry are used for the detection of midrange and distant vehicles. While these clues have previously been applied to the task of vehicle detection, merging them into a single approach appears to be new. In this section, we therefore concentrate on the

BY MARINUS B. VAN LEEUWEN AND FRANS C.A. GROEN

 1070-9932/05/$20.00©2005 IEEE

One of the essential features of intelligent driver assistance systems is the ability to detect other vehicles on the road.

merging strategy. (For more detailed discussions about each clue and for our arguments against choosing alternatives like vertical line structures or color, please refer to [3].) The discussion of how to compute the clues from the image data is limited to the essentials.

Our detection procedure is illustrated by the example in Figure 1. The procedure starts with identification of image data that potentially belong to the shadow underneath a vehicle. The major advantage of shadow is that "all" vehicles will be detected. This results in a number of subregions that are analyzed further (image with subregions labeled 1–9). In a first step, rows with low entropy are removed from each subregion. If too few rows are left for further analysis, the region is classified as background. Otherwise, the horizontal symmetry of the remaining data is investigated. Asymmetric regions are classified as background (subregions 5, 7, and 9) and symmetrical regions are classified as vehicle (subregions 1, 2, 3, and 4). Besides separating vehicles from background, the boundaries of the region of symmetry provide us with an indication of the position of the vehicle within the subregion. This is illustrated in the images at the bottom of Figure 1.

Shadow is the first clue that is searched for in the image plane. The purpose of shadow searching is to indicate the positions where vehicles are present, with as small an amount of false detections as possible. The work of [4] forms the basis for our method of identifying potential shadow regions underneath vehicles. Following [4], we exploit the fact that the shadow underneath a vehicle will be darker than the intensity with which the surface of the road is observed.

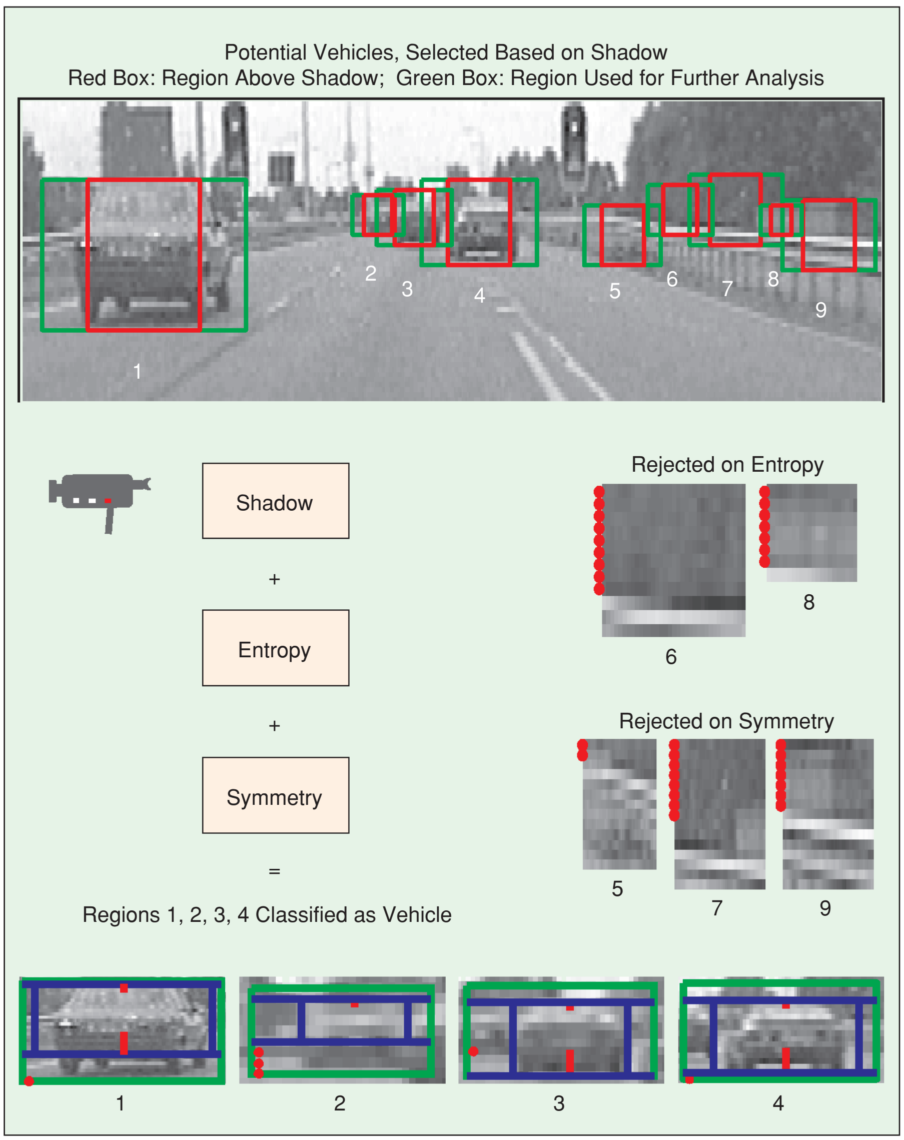


Figure 1. *An example for midrange and distant vehicle detection.*

We assume a normal distribution for the gray values of the road surface. By means of fitting a Gaussian to the histogram of pixels belonging to the road, we estimate the mean value m and variance σ of this distribution. Under good illumination, the upper boundary for shadow can be set to $\mathrm{Thresh}_{sh} < m\text{-}3\sigma$. However, in the case of decreasing illumination, the difference between shadow and road will decrease.

The intensity distribution of the road surface is estimated without the need for preknowledge about the location of the road in the image plane. It is derived by means of a simple algorithm for detecting the free-driving-space (f-d-s) of our vehicle. The f-d-s is defined as the road observed directly in front of the camera. In the case of observing the road through the rear window of the vehicle, this area corresponds with the road surface observed behind the vehicle. An example of the estimated f-d-s is given in Figure 2. First, edges in the image plane are estimated [Figure 2(a)]. Then, we determine the space defined by the lowest central homogeneous region in the image delimited by edges [Figure 2(b)].

Although alternative approaches might lead to more refined estimations of the f-d-s, this coarse procedure is sufficient for our application. Temporal smoothing will be applied to the parameters $\{m, \sigma, \mathrm{Thresh}_{sh}\}$, allowing for accidentally inaccurate estimations of the f-d-s.

Once the upper boundary for shadow pixels has been defined, we select the pixels that exhibit a vertical transition from a brighter gray value to shadow (scanning the image bottom-up). The result is shown in Figure 3. Within this result, we search for horizontally connected points. While scanning the data bottom-up, for each line l_i, 8-connectivity between pixels is determined in lines l_i and l_{i-1}. By thresholding the length of a line segment, we roughly separate lines belonging to potential vehicles from points or small line segments introduced by the background. Above each line connected to a potential vehicle, a region of interest (ROI) is defined as illustrated in Figure 3(c) by the red boxes. Their width corresponds to the length of the horizontal line segments. We use slightly wider regions (green boxes) for analysis of the entropy and symmetry properties of each ROI.

Detection is the process of identifying the presence of clues in the image plane.

Texture is (among others) used by Kalinke et al. in [2]. In their paper, texture is applied to focus their vehicle detection algorithm on sections of the image with high information contents. For our application, we are interested in the entropy in ROIs that possibly contain a vehicle. If a vehicle is observed, its left and right boundaries—as well as other intensity variations in horizontal direction between these boundaries—will cause high entropy (estimated along horizontal lines). If the ROI contains no vehicle, the lack of intensity variation caused by homogeneous areas results in low entropy. We estimate the local entropy along lines of the

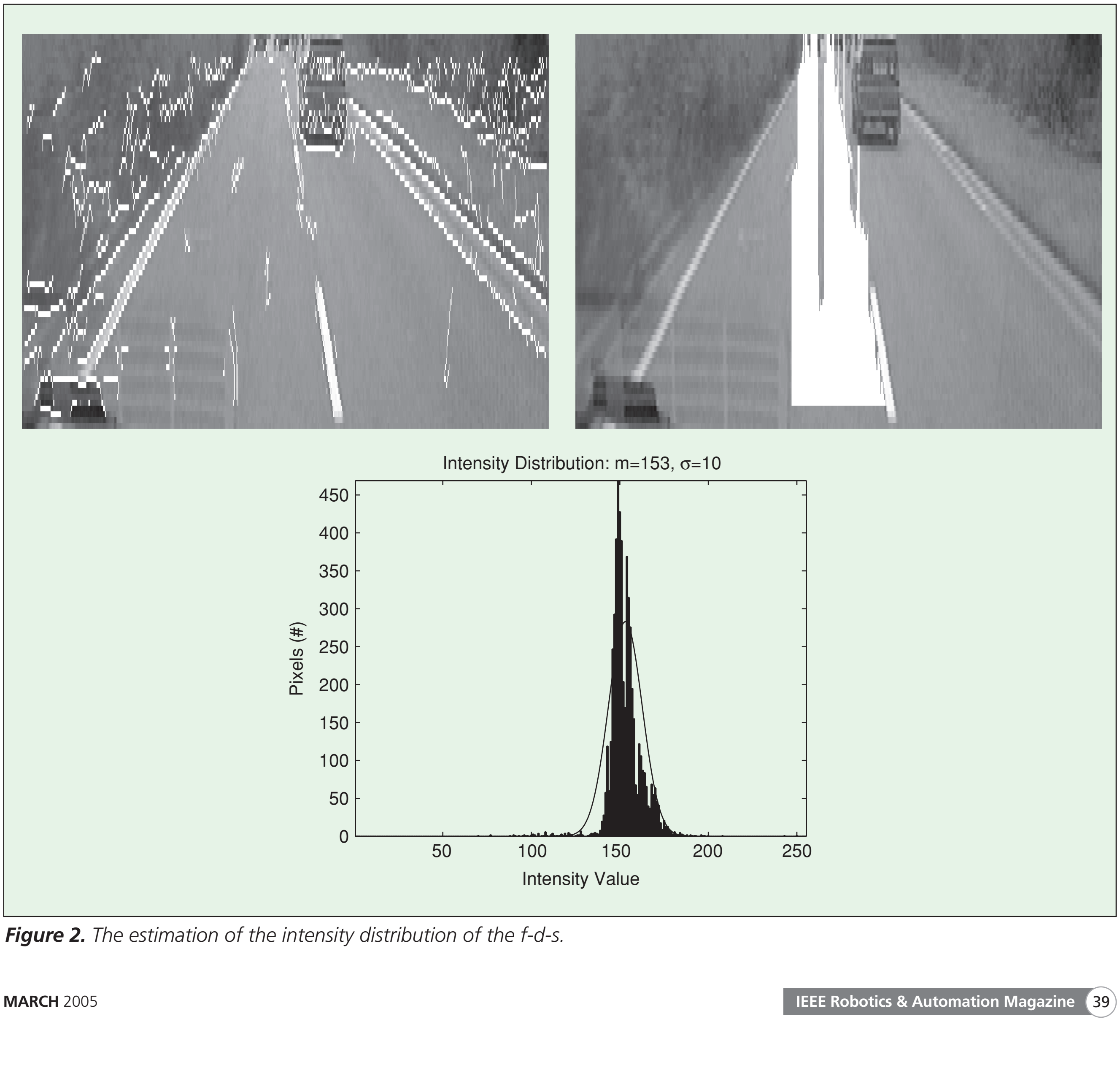


__Figure 2.__ The estimation of the intensity distribution of the f-d-s.

For the detection of midrange and distant vehicles, we use a combination of three clues: shadow, entropy, and horizontal symmetry.

ROI, removing lines that indicate a lack of entropy and, thus, contain little information. An example is given in Figure 1.

The preceding steps of the detection procedure already separate much of the background data from data potentially belonging to a vehicle. They also present the data assigned for further symmetry analysis in such a way that this analysis is logical and can be carried out efficiently. Uniform data has been removed. Good horizontal symmetry can be detected in uniform background regions. Estimating its symmetry would therefore be useless. Moreover, the derivation of the optimal symmetry of uniform areas is very sensitive to small intensity changes. Furthermore, the preceding steps provide us with a limited ROI. These factors bring efficiency to the process of finding the axis of symmetry and the width of the symmetry interval.

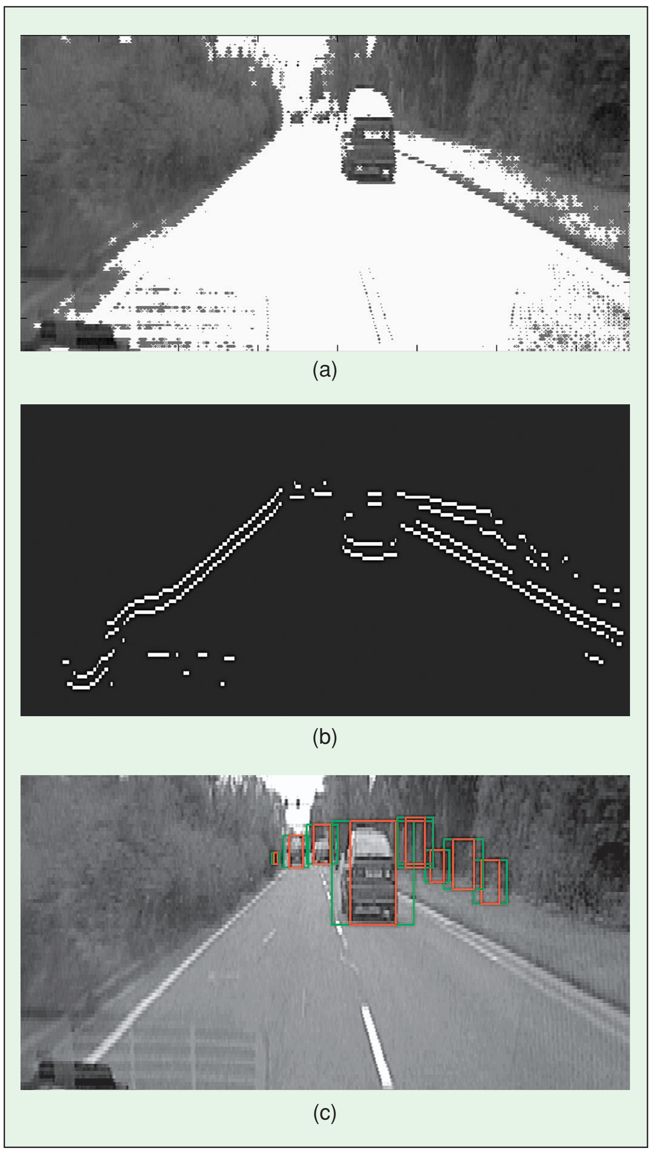


Figure 3. *Hypotheses generated based on shadow.*

For estimating local symmetry, we use the approach described in [5]. The vehicle detection approach used differs from ours in both its preprocessing and postprocessing. They use symmetry detection as a first step in the detection procedure and are mainly concerned with detecting the leading vehicle (midrange). Once a region with "best" horizontal symmetry has been identified, the symmetry of its edge structure is analyzed before deciding whether or not the region contains a vehicle. For a detailed description of how we estimate local horizontal symmetry, refer to [3].

When a vehicle is observed in the lane next to us, asymmetry might be introduced in the lower part of the ROI by the projected geometry of the wheels. Due to inaccurate initialization of the ROI, a few horizontal lines above the vehicle might have been included. In a postprocessing step, we aim to purify the data upon which the analysis is based by removing the lines in the upper and lower quarter that exhibit relatively low symmetry values. The examples at the bottom of Figure 1 illustrate this concept. As a second step, we also refine the estimation of the interval width.

Detecting Passing Vehicles

When a vehicle passes us, it affects the intensity distribution of the area it occludes. Therefore, temporal differences can be applied to identify passing vehicles. Significant intensity variations in the background will lead to false detections. The projected motion component in the horizontal direction of background objects differs from the motion of vehicles. Therefore, motion can be used to suppress false detections. For detecting passing vehicles, we propose two adjustments to the generally applied temporal differences (e.g., [1]). First, we search for passing vehicles in a slightly different region, namely the section where the surface of the road is observed. The intensity distribution of the background observed in this area will usually be very moderate. Second, we analyze the projected motion in this region to suppress false detections.

The region in the corner of the eye where the road surface is observed is referred to as the $\mathrm{ROI}_{\mathrm{coe}}$. The $\mathrm{ROI}_{\mathrm{coe}}$ is subdivided into subregions $\mathrm{ROI}_{\mathrm{coe}}(i)$, consisting of 20 lines. The procedure begins with the calculation of the projected motion. The projected motion for segment i is derived by the summation of intensities in the vertical direction for a certain time interval. The result is an image $\mathrm{Im}_{\mathrm{PM}}(i)$ of size #columns × #frames. Characteristic examples of projected motion images of passing vehicles are given in Figure 4. Next, the mean intensity, $m_i(t)$, is determined. By comparing $m_i(t)$ with the history of the intensity distribution of the f-d-s, we determine significant darkening of $ROI_{coe}(i)$. Darkening is caused by the shadow underneath a passing vehicle, its tires, or its lower part. Increase in the mean intensity, such as that caused by the usually light markers on the road (white or yellow), is ignored. Thus, in some way, we also apply shadow as a clue for passing vehicle detection. As soon as $m_i(t)$ significantly decreases, we investigate the motion characteristics of this darkening. Variation in $m_i(t)$ results in edges in the projected motion image $Im_{PM}(i)$. The orientation of these edges, denoted by the angle θ, is a measure of how quickly the spatial intensity variation propagates through each of the images $Im_{PM}(i)$.

The reason for segmenting ROI_{coe} is twofold. First, darkening observed in each segment will be more severe. A passing vehicle might not appear in the complete ROI_{coe} and will probably not appear in each row at the same time. The darkening of the complete ROI_{coe} will be more moderate and thus more difficult to detect. The second reason for segmenting ROI_{coe} is that the projected motion image will be of better quality.

Experimental Results

In this article, we only discuss two experiments. More experiments can be found in [3].

The sequence of the first experiment consists of 200 frames recorded with a camera focused through the windscreen of the vehicle. In the first half of the sequence (approximately 4 s), the environment contains little structure. This part can be considered typical for highway scenes. In the second half of the sequence, the background contains more structure and, therefore, can be considered more characteristic of urban scenes.

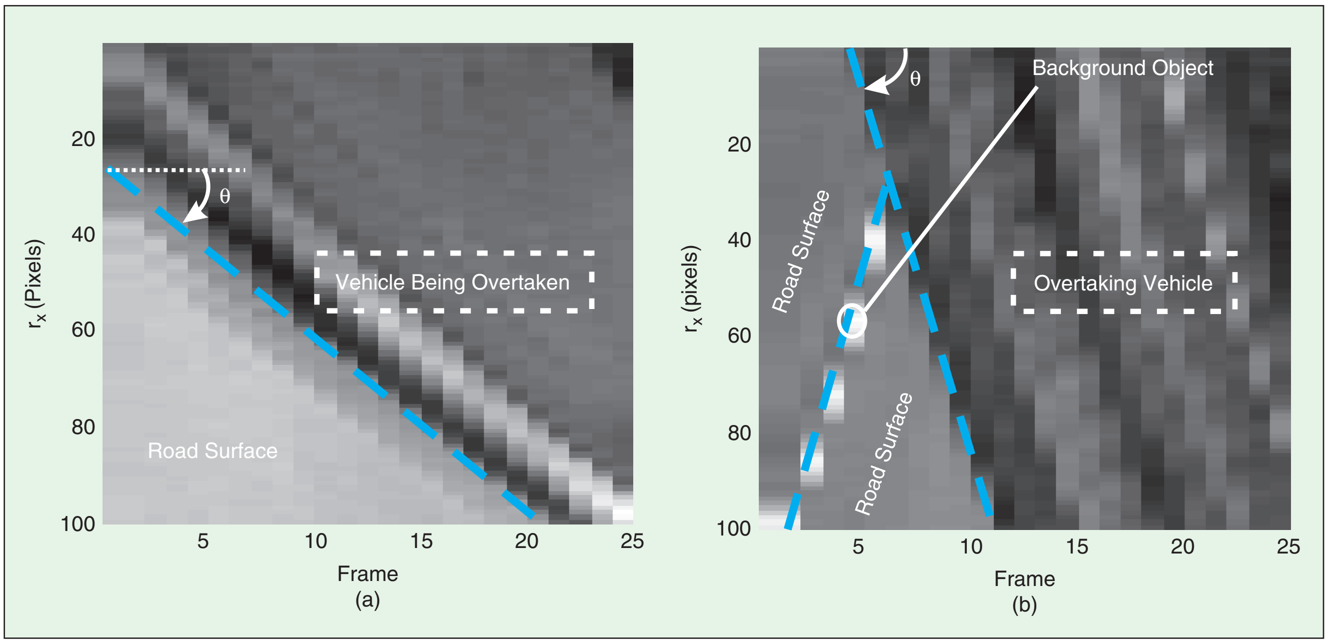


Figure 4. *The projected motion for different viewing directions through (a) rear window and (b) windscreen.*

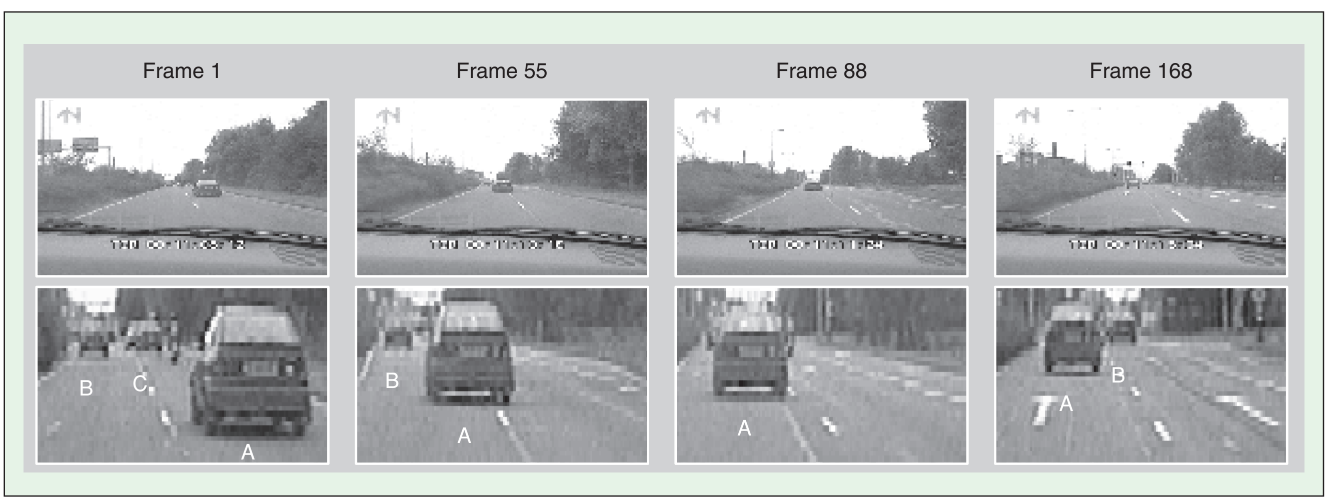


Figure 5. *Four characteristic frames from sequence one (the second row is a zoomed view).*

Vehicle	# Frames	Times Missed	Rate
A (leading vehicle, mid-range)	200	10	0.95
B (before occlusion, distant)	57	2	0.96
B (after occlusion, distant)	33	12	0.64
C (before occlusion, distant)	8	1	0.88
	# Frames	**Times Detected**	**Rate**
False positives, first half	100	4	0.04
False positives, second half	100	94	0.94
False positives	200	98	0.49

Table 1. Detection Results.

underneath vehicle B is not (properly) detected. Finally, vehicle C was also robustly detected.

The number of false positives and the removals based on both entropy and symmetry provide us with more insight on the performance of the algorithm. The amount of false positives is also summarized in Table 1. In the first 100 frames, only four false positives were generated. In the second half, 94 were generated. Rejections based on entropy equaled only 1,151, or about six rejections per frame. No significant difference was observed between the first and second half of the sequence. The number of removals based on entropy plus symmetry equaled 165 in the complete sequence, 25 in the first 100 frames, and 145 in the second half. The difference in the amount of structure present in the first and second half is expressed by the increase in false positives and removals based on entropy plus symmetry.

A visual summary of the sequence is presented in Figure 5. In our discussion, we address the vehicles in the sequence according to the labels shown in this figure. Vehicle A is the leading vehicle and is observed throughout the complete sequence. Vehicle C is only observed for a short time before it becomes occluded by vehicle A. After 57 frames, vehicle B is temporarily occluded by vehicle A; it is detected again after frame 167.

The detection results for each vehicle are presented in Table 1. The results show how robustly the leading vehicle (midrange) was detected by the algorithm. The more distant vehicles were also robustly detected. Prior to occlusion, vehicle B was detected in 55 out of 57 frames (96%). After occlusion, the chance the algorithm detected vehicle B dropped to 64%. Although still better than once every two frames, the performance dropped significantly due to a change in illumination of the road surface. The second time vehicle B is observed, it enters a cross section. The illumination of the road surface directly in front of us stays under the influence of trees alongside the road. The surface of the road observed at the cross section lacks this influence and is therefore of a brighter intensity. As a result, the performance of the shadow detector decreases; several times, the shadow

The most important aspect of our approach for passing vehicle detection is that all vehicles are detected. The number of false detections is also important. When automatically detecting passing vehicles, we need to know when a potential vehicle has completely moved from the corner of the eye (ROI_{coe}). Until the vehicle has passed by, the detection algorithm is paused. The moment to activate the detection algorithm again is usually indicated by some kind of tracking algorithm. In our experiment, we waited until the intensity distribution of ROI_{coe} returned to the distribution of the road surface. This simple strategy appeared to be sufficient for our experiment.

The practical sequence we discuss is about 20 s in length (478 frames). The results are presented in Figure 6. The figure illustrates the intensity variation in ROI_{coe}

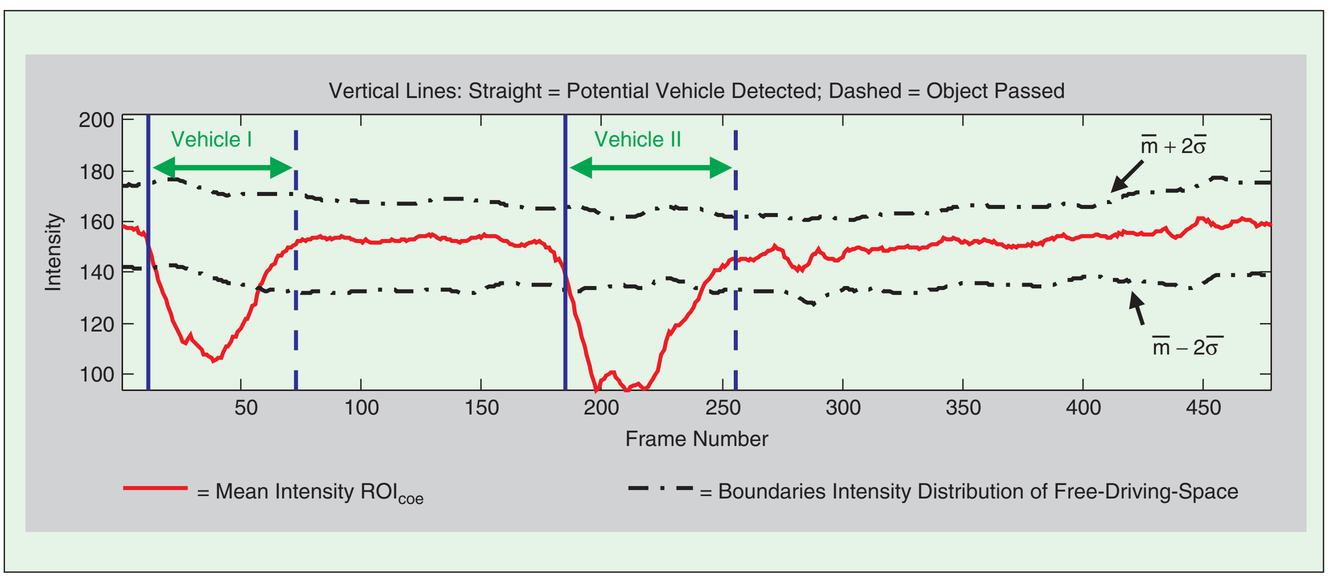


Figure 6. *The experimental result for passing vehicle detection.*

(straight curve). The variation in the intensity distribution of the f-d-s is also indicated (dot-dash curves). Two curves indicate the boundaries $\overline{m} \pm \overline{\sigma}$. These curves are slightly smoothed by averaging over 0.5 s of data. The vertical, blue straight lines indicate the frames where significant darkening was observed, and the vertical, blue dashed lines indicate where the object had moved from the ROI_{coe}. After darkening has been detected, the projected motion is further analyzed. In our experiment, two potential vehicles are identified.

Discussion

In this article, we presented two approaches for vehicle detection: one for detecting midrange and distant vehicles and another for detecting passing vehicles. The approaches are suitable for detecting vehicles observed in frontal or rear view. No preknowledge about the position of the road is required. We only require a rough estimation of the FOC (or FOE).

To detect midrange and distant vehicles, we apply three different clues. An object is only identified as a vehicle if all three clues are found on the object; the likelihood that a background object exhibits all three clues is small. Besides the quality of the interpretation of the observed data (vehicle or background), there are other advantages attached to the combination of these three clues. Shadow detection provides proper initialization of the symmetry analysis. This initialization is crucial for both the efficiency and accuracy of the symmetry estimation step. When a vehicle is detected in a certain ROI, symmetry analysis provides a more refined estimation of the position of the vehicle. Furthermore, the removal of horizontal lines with a lack of entropy ensures that the symmetry analysis makes sense.

The experiment illustrates the performance of our detection approach. Vehicles are observed at high detection rates. Almost no false detections are introduced if the environment contains little structure (e.g., trees alongside the road). In an urban environment, the number of false detections increases. In any future implementation, the number of false potentials generated based on shadow should be reduced. Preknowledge about the location of the road could potentially be used. Or, a more sophisticated relation could be employed between the expected vehicle width at a certain position in the image plane and the shadow potentially belonging to a vehicle. The result will be a more efficient algorithm with less false detections.

Our approach for detecting passing vehicles is based on a combination of temporal differencing and projected motion. We limit the region in the corner of the eye to the section where the road surface is observed. In this region, the background intensity varies very little. Only in special cases will false alarms will be introduced by a discoloring of the road surface. In these cases, false detections are rejected based on their projected motion.

Keywords

Intelligent vehicles, vehicle detection, shadow, texture, symmetry, temporal differencing, projected motion.

References

[1] M. Betke, E. Haritaoglu, and L.S. Davis, "Real-time multiple vehicle detection and tracking from a moving vehicle," *Mach. Vis. Appl.,* vol.12, no. 2, pp. 69–83, Sept. 2000.

[2] T. Kalinke, C. Tzomakas, and W. von Seelen, "A texture-based object detection and an adaptive model-based classification," in *Proc. IEEE Int. Conf. Intelligent Vehicles*, Stuttgart, Germany, Oct. 1998, pp. 143–148.

[3] M.B. van Leeuwen, "Motion estimation and interpretation for in-car systems" Ph.D. dissertation, University of Amsterdam, 2002.

[4] C. Tzomakas and W. von Seelen, "Vehicle detection in traffic scenes using shadows," Institut fur Neuroinformatik, Ruhr-Universitat, Bochum, Germany, Inter. rep. IR-INI 98–06, 1998.

[5] T. Zielke, M. Brauckmann, and W. von Seelen, "Intensity and edge-based symmetry detection with an application to car-following," *Comput. Vis. Graph. Image Process.: Image Understanding*, vol. 58, no. 2, pp. 177–190, 1993.

Marinus B. van Leeuwen received an M.Sc. degree in electrical engineering from the Eindhoven University of Technology, The Netherlands, in 1997. In 2002, he received his Ph.D. in informatics from the University of Amsterdam for research on motion estimation and interpretation for traffic applications. He works at the Philips Industrial Vision Products Department.

Frans C.A. Groen became a scientific staff member of the Pattern Recognition Group at the Applied Physics Department of the TU-Delft in 1970, where he was responsible for robotics research. He was a Fullbright research scientist at the Robotics Institute at CMU in 1984. In 1988, he became a full professor with the Computer Science Department of the University of Amsterdam, heading the research on intelligent autonomous systems. He served part time as a full professor at the Free University of Amsterdam from 1986 to 1996, where he set up a group in Physics Applied Computer Science. In 1996, he was a visiting professor at the University of Utah. He has served on international and national advisory boards, published over 160 papers, and supervised 20 Ph.D. theses. His current research interest is in perception for intelligent autonomous systems and cooperative robotics.

Address for Correspondence: M.B. van van Leeuwen, Philips Applied Technologies, Industrial Vision Products Department, Building SAQ-p342, PO Box 218, 5600 MD Eindhoven, The Netherlands. E-mail: rien.van.leeuwen@philips.com.