

Obstacle Detection Using Projective Invariant and Vanishing Lines

Ryuzo Okada Yasuhiro Taniguchi Kenji Furukawa Kazunori Onoguchi
Multimedia Laboratory, Corporate Research & Development Center, Toshiba Corporation
1 Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, Kanagawa 212-8582, Japan
E-mail: ryuzo.okada@toshiba.co.jp

Abstract

This paper presents a novel method for detecting vehicles as obstacles in various road scenes using a single on-board camera. Vehicles are detected by testing whether the motion of a set of three horizontal line segments, which are always on the vehicles, satisfies the motion constraint of the ground plane or that of the surface plane of the vehicles. The motion constraint of each plane is derived from the projective invariant combined with the vanishing line of the plane that is a prior knowledge of road scenes. The proposed method is implemented into a newly developed on-board LSI. Experimental results for real road scenes under various conditions show the effectiveness of the proposed method.

1. Introduction

Obstacle avoidance is a fundamental requirement for autonomous mobile robots and vehicles, and numerous vision-based obstacle detection methods have been proposed. Some of them segment out obstacles from the ground plane based on differences of geometric properties, such as the motion parallax [2, 3, 5, 10, 14], the projective invariant [16, 18], and the depth information [1, 7, 9]. Others detect known obstacles based on their 2-D image pattern learned beforehand [12, 19]. For mobile robots and vehicles that are capable of working in the real world, the visual cues such as the depth and the 2-D image pattern need to be suitably chosen according to an individual target application.

In this paper, our goal is to detect other vehicles as obstacles in road scenes using an onboard visual sensor system, which is useful for Intelligent Transport Systems (ITS) such as Adaptive Cruise Control System (ACC). Such systems require the following properties: (1) adaptability to a large variety of vehicles, (2) compactness, and (3) robustness against varied conditions in road scenes. Since single camera systems are more compact than stereo camera systems, we chose a single camera system as the visual sensor. Considering the adaptability and robustness requirements,

we avoided adopting the learning strategy, because detection of all types of vehicles under the whole range of road conditions requires a large number of images under various conditions, and these are not readily available. Instead, we chose the motion information as the visual cue.

Typical motion-based obstacle detection methods [5, 14, 16] estimate the motion of the ground plane and detect obstacle regions whose motions differ from that of the ground plane. Zhang et al. [21] adopt an approach without motion estimation of the ground plane. When some of feature points, such as corners, are on the ground plane and the others belong to obstacles and thus such a set of feature points does not have a planar motion, it is determined that an obstacle exists. The basic assumption of these methods is that there are a sufficient number of feature points both on the ground plane and on obstacles. However, there are often few feature points on the ground plane in real road scenes, as seen in Fig. 8 and Fig. 10(b).

One general approach to motion-based obstacle detection might be to reconstruct the 3-D structure of a scene from the motions in an image sequence [4, 8, 15]. The motions are computed by finding correspondences between points on a pair of images captured from different viewpoints. Sinclair et al. [18] extract planar regions that conserve the values of two projective invariants based on the correspondences of five corner points, and estimate the normal of each segmented plane. However, in some real road scenes, such as rainy scenes and scenes with variable lighting conditions, these bottom-up methods pose difficulty in terms of robustness because errors in finding correspondences, which is a low-level procedure, result in poor estimation of the 3-D structure or the normals.

To achieve robustness, prior knowledge about road scenes has been introduced. Carlsson et al. [2] estimate the velocity of the known ground plane, and detect obstacles as regions whose motions differ from the estimated motion of the ground plane. This method also requires feature points on the ground plane for motion estimation of the ground plane. Hu et al. [11] and Sato et al. [17] restrict the ego-motion of an observer based on the prior knowledge of road

scenes. In such methods, only negligible components of the egomotion in real road scenes should be restricted. Thus, making effective use of prior knowledge for robust obstacle detection in various scenes is a significant issue.

In this paper, we present an obstacle detection method that effectively combines a projective invariant, i.e. the cross ratio, and prior knowledge of a road scene. The prior knowledge that we employ is the vanishing lines¹ of the ground plane and the surface planes of the obstacles. The proposed method works without feature points on the ground and with pitching and yawing of the observer. Although we assume that some textures are on the surface of the obstacles and rolling of the observer is negligible, the assumptions are valid and the proposed method robustly works in real road scenes.

2. Obstacle Detection with Projective Invariant

2.1. Projective Invariant for Horizontal Lines

Projective invariants are quantities that do not change under projective transformations. The cross ratio, which is known to be a projective invariant, is defined for a group of four colinear points, A, B, C, and D, as:

$$I_{ABCD} \equiv \frac{\overline{AC}}{\overline{AD}} \cdot \frac{\overline{BD}}{\overline{BC}}, \quad (1)$$

where I_{ABCD} denotes the cross ratio of the four colinear points, A, B, C, and D, and \overline{AC} represents the distance between A and C, etc (see Fig. 1).

In this paper we introduce the following two assumptions on road scenes: (1) artifacts in road scenes, such as other vehicles and road signs painted on the ground plane, contain some parallel horizontal line segments, that we call horizontal segments, and (2) the surface of an artifact is approximated to a plane. Let L_A , L_B , L_C , and L_D be coplanar horizontal segments on the surface plane of an artifact. They pass through A, B, C, and D, respectively, and are perpendicular to the line passing through A, B, C, and D. The distance between two colinear points, e.g. \overline{AC} , is replaced by the distance, Δ_{AC} , between coplanar horizontal segments, L_A and L_C .

We assume a pinhole camera model. Although, for simplicity, we describe the case with the optical axis of the onboard camera parallel to the road (see Fig. 1), the following discussion is easily extended to a general camera configuration with a depression angle. For this camera configuration, the surface of the obstacles is approximately represented by a fronto-parallel plane, which we call the obstacle plane, and the ground plane is perpendicular to the obstacle plane and parallel to the optical axis.

¹It should be noted that the projective invariant is employed together with vanishing points for purposes other than obstacle detection, such as 3-D motion analysis [13] and estimation of time-to-contact.

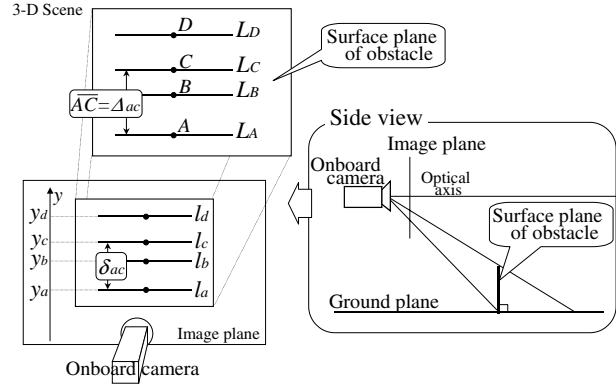


Figure 1. Four colinear points and four coplanar parallel horizontal segments on the surface plane of an obstacle.

A horizontal segment, L_A , on the ground plane or on the obstacle plane is projected to a parallel horizontal segment in an image, l_a , which we also call a horizontal segment. The distance, δ_{ac} , between two horizontal segments, l_a and l_c , is represented by the difference of their vertical positions, y_a and y_c . The value of the cross ratio for a group of four coplanar horizontal segments, L_A , L_B , L_C , and L_D , is constant for the projective transformation:

$$I_{ABCD} = \frac{\Delta_{ac}}{\Delta_{ad}} \cdot \frac{\Delta_{bd}}{\Delta_{bc}} = \frac{\delta_{ac}}{\delta_{ad}} \cdot \frac{\delta_{bd}}{\delta_{bc}}, \quad (2)$$

where $\delta_{ac} = y_a - y_c$ etc.

While a plane containing a group of four coplanar horizontal segments, L_A , L_B , L_C , and L_D , moves between the time t_1 and t_2 , their cross ratio is conserved:

$$\frac{\delta_{ac}(t_1)}{\delta_{ad}(t_1)} \cdot \frac{\delta_{bd}(t_1)}{\delta_{bc}(t_1)} = \frac{\delta_{ac}(t_2)}{\delta_{ad}(t_2)} \cdot \frac{\delta_{bd}(t_2)}{\delta_{bc}(t_2)} = I_{ABCD}. \quad (3)$$

Eq. (3) is a motion constraint for the four coplanar horizontal segments. If any of the four horizontal segments belong to an obstacle plane and the others are on the ground plane, such a group of four horizontal segments does not satisfy the motion constraint because the segments are not coplanar. An obstacle detection method based on this coplanarity condition, which tests the planarity of a set of motions [21], requires texture, including corners and horizontal segments, both on the ground plane and on obstacles. Such an assumption is often invalid in real road scenes.

2.2. Introducing Vanishing Lines

We introduce the orientation of the ground plane and that of the obstacle plane, which are the prior knowledge of road scenes, into estimation of the cross ratio. Given the orientation of a plane, the vanishing line of the plane is easily

computed. By substituting the vanishing line for one of the four horizontal segments in Eq. (3), we combine the projective invariant and prior knowledge of road scenes.

The vanishing line of the ground plane and that of the obstacle plane are respectively $y = 0$ and $y = \infty$ in our camera configuration as long as both pitching and rolling of the onboard camera are zero.

We replace l_d in Eq. (3) by the vanishing line of the ground plane, $y = 0$, and obtain the motion constraint for the ground plane, which we call the *GP motion constraint*:

$$\frac{\frac{1}{y_a(t_2)} - \frac{1}{y_c(t_2)}}{\frac{1}{y_a(t_1)} - \frac{1}{y_c(t_1)}} = \frac{\frac{1}{y_b(t_2)} - \frac{1}{y_c(t_2)}}{\frac{1}{y_b(t_1)} - \frac{1}{y_c(t_1)}} = M_h, \quad (4)$$

where $M_h = 1$, and $y_i(t)$ denotes the vertical position of l_i at time t .

Similarly, the motion constraint for obstacles, which we call the *obstacle motion constraint*, is obtained by replacing l_d by the vanishing line of the obstacle plane, $y = \infty$:

$$\frac{y_a(t_2) - y_c(t_2)}{y_a(t_1) - y_c(t_1)} = \frac{y_b(t_2) - y_c(t_2)}{y_b(t_1) - y_c(t_1)} = M_v, \quad (5)$$

where $M_v = \frac{y_a(t_2) - y_b(t_2)}{y_a(t_1) - y_b(t_1)}$.

2.3. Obstacle Detection

Basically, we determine that a group of three horizontal segments belongs to an obstacle if the motion of the group satisfies the obstacle motion constraint better than the GP motion constraint. As described in Section 2.1, the obstacle detection method based on four horizontal segments tests the coplanarity of the segments, which is a variant of the existing method proposed by Zhang et al. [21]. In contrast, our method based on three horizontal segments tests whether the orientation of a plane that they belong to is close to that of the ground plane or to that of the obstacle plane. Examining the orientation makes it possible to detect obstacles in a scene without feature points on the ground plane.

We predict the positions of $y_b(t_2)$ and $y_c(t_2)$ based on the GP motion constraint and the obstacle motion constraint, respectively. The predicted position based on the GP motion constraint, $y_b^h(t_2)$ and $y_c^h(t_2)$, and the predicted position based on the obstacle motion constraint, $y_b^v(t_2)$ and $y_c^v(t_2)$, are obtained as follows:

$$\left\{ \begin{array}{l} \frac{1}{y_b^h(t_2)} = \frac{1}{y_c(t_2)} + M_h \left(\frac{1}{y_b(t_1)} - \frac{1}{y_c(t_1)} \right) \\ \frac{1}{y_c^h(t_2)} = \frac{1}{y_a(t_2)} + M_h \left(\frac{1}{y_c(t_1)} - \frac{1}{y_a(t_1)} \right) \end{array} \right., \quad (6)$$

$$\left\{ \begin{array}{l} y_b^v(t_2) = y_c(t_2) + M_v(y_b(t_1) - y_c(t_1)) \\ y_c^v(t_2) = y_a(t_2) + M_v(y_c(t_1) - y_a(t_1)) \end{array} \right. . \quad (7)$$

Small differences between the observations², $y_b(t_2)$ and

²The position of a horizontal segment at times t_2 is obtained by tracking the horizontal segment (see Section 3.1).

$y_c(t_2)$, and the predictions based on the obstacle motion constraint, $y_b^v(t_2)$ and $y_c^v(t_2)$, signify that the three horizontal segments belong to obstacles. Thus, we define the validity of three horizontal segments, l_a , l_b , and l_c , for the obstacle motion constraint as:

$$s_{abc} = D_{abc}^h - D_{abc}^v, \quad (8)$$

$$D_{abc}^v = |y_b^v(t_2) - y_b(t_2)| + |y_c^v(t_2) - y_c(t_2)|, \quad (9)$$

$$D_{abc}^h = |y_b^h(t_2) - y_b(t_2)| + |y_c^h(t_2) - y_c(t_2)|. \quad (10)$$

When s_{abc} has a large positive value, the three horizontal segments belong to obstacles.

Predicted positions based on the GP motion constraint, however, include errors caused by pitching and rolling of the onboard camera because pitching and rolling vary the vanishing line of the ground plane. Though rolling is negligible because rolling of the observer is relatively minor in road scenes, pitching should be taken into consideration. We redefine the difference, D_{abc}^h , to be the minimum difference between the observations and the predictions for various shifts, d , of the vanishing line:

$$D_{abc}^h = \min_{|d| < d_{max}} D_{abc}^h(d), \quad (11)$$

$$D_{abc}^h(d) = |y_b^h(t_2, d) - (y_b(t_2) + d)| + |y_c^h(t_2, d) - (y_c(t_2) + d)|, \quad (12)$$

where

$$\frac{1}{y_b^h(t_2, d)} = \frac{1}{y_c(t_2) + d} + M_h(d) \left(\frac{1}{y_b(t_1)} - \frac{1}{y_c(t_1)} \right), \quad (13)$$

$$\frac{1}{y_c^h(t_2, d)} = \frac{1}{y_a(t_2) + d} + M_h(d) \left(\frac{1}{y_c(t_1)} - \frac{1}{y_a(t_1)} \right), \quad (14)$$

$$M_h(d) = 1 + \left(\frac{d}{f} \right)^2. \quad (15)$$

Note that f denotes the focal length and d_{max} is the maximum shift of the vanishing line.

The error of prediction based on the obstacle motion constraint is negligible, because pitching can be approximated to vertical shifts of an image, and the obstacle motion constraint (Eq. (5)) is independent of the vertical shifts, i.e. $(y_a(t_2) - d) - (y_c(t_2) - d) = y_a(t_2) - y_c(t_2)$ etc.

2.4. Discussion

We have presented an obstacle detection method based on three horizontal segments. In this section we discuss algorithms based on four and two horizontal segments in addition to the proposed method, and show the conditions under which each algorithm is applicable. The conditions will clearly illustrate the advantage of the proposed method.

As described in Section 2.1, obstacles can be detected by testing the coplanarity of four horizontal segments without prior knowledge about road scenes. In contrast, the proposed method is capable of detecting obstacles in a scene

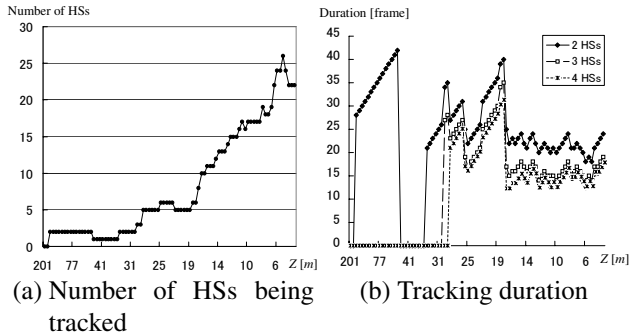


Figure 2. Comparison between algorithms with two, three, and four horizontal segments (HSs).

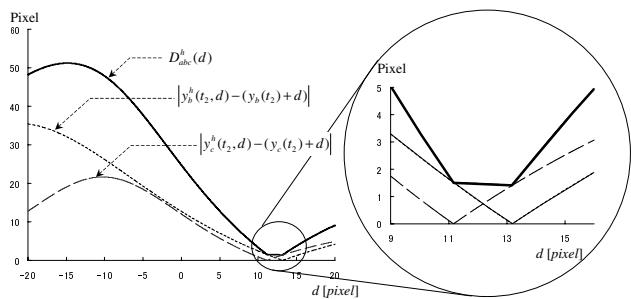


Figure 3. $D_{abc}^h(d)$ and its components, $|y_b^h(t_2, d) - (y_b(t_2) + d)|$ and $|y_c^h(t_2, d) - (y_c(t_2) + d)|$, in a typical scene.

without feature points on the ground plane by introducing the vanishing lines as a prior knowledge.

If the image shift caused by pitching of the onboard camera is negligibly small, the pitching compensation described by Eq. (11)–(15) is unnecessary. In this case, the validity for the obstacle motion constraint can be defined based on two horizontal segments as:

$$s_{bc} = |y_b^h(t_2) - y_b(t_2)| - |y_b^v(t_2) - y_b(t_2)| \quad (16)$$

$$\frac{1}{y_b^h(t_2)} = \frac{1}{y_c(t_2)} + M_h \left(\frac{1}{y_b(t_1)} - \frac{1}{y_c(t_1)} \right) \quad (17)$$

$$y_b^v(t_2) = y_c(t_2) + M_v (y_b(t_1) - y_c(t_1)) \quad (18)$$

Note that $M_v = y_c(t_2)/y_c(t_1)$.

Fig. 2(a) shows the number of detected horizontal segments belonging to a vehicle approaching an onboard camera. The distances, Z , where two, three, and four horizontal segments are first observed are 170, 30, and 28 [m], respectively. This means that a method based on few horizontal segments is capable of computing the validity for remote vehicles, and is thus advantageous for the detection of such remote vehicles.

Motion-based obstacle detection methods require large motions in an image sequence for reliable obstacle detection. As we observe the motions by tracking each horizontal segment (see Section 3.1), a method that tracks a group of horizontal segments for a long period is capable of detecting obstacles early. Fig. 2(b) shows the average duration (number of frames) of tracking for all combinations of two, three, or four horizontal segments. Since the method based on two horizontal segments has the longest tracking duration, this method is capable of finding obstacles earlier than the other two. This is because tracking failure makes the tracking duration for three or four horizontal segments shorter than that for two horizontal segments.

However, the method based on two horizontal segments is applicable only to scenes where the image shifts due to pitching of the onboard camera are negligibly small. On the other hand, our method is capable of canceling pitching through Eqs. (11)–(15). When three horizontal segments, L_a , L_b , and L_c , belonging to an obstacle approach the onboard camera with pitching, the difference, D_{abc}^h , for the GP motion constraint does not become zero, because no shifts, d , simultaneously makes the two differences, $|y_b^h(t_2, d) - (y_b(t_2) + d)|$ and $|y_c^h(t_2, d) - (y_c(t_2) + d)|$, zero (see Fig. 3).

The method based on four horizontal segments requires texture, including corners and horizontal segments, on the ground plane, and the method based on two horizontal segments places a restriction on pitching. In contrast with these methods, the proposed method is applicable to various real road scenes because it does not have such restrictions.

3. Preceding Vehicle Detection System

3.1. Algorithm

Fig. 4 shows the flow of procedures in each frame for a preceding vehicle detection system based on the proposed method described in Section 2. We first set a detection area and extract horizontal segments in this area. Our system employs a fixed detection area covering the driving lane with typical width in front of the observer (see Fig. 5). Next, we add the detected horizontal segments to a list, which we call a tracking list, and track the horizontal segments in the tracking list. Since tracking failure causes a spatio-temporally inconsistent motion of horizontal segments, we eliminate such horizontal segments from the tracking list. Finally, we detect obstacles based on the proposed method. The details of the procedures are as follows.

Detection of Horizontal Segments Horizontal segments in an image are observed as horizontal edges. We introduce the separability [6] of image intensities for robust detection of horizontal edges in real road scenes. The separability is a quantity which represents the statistical difference between

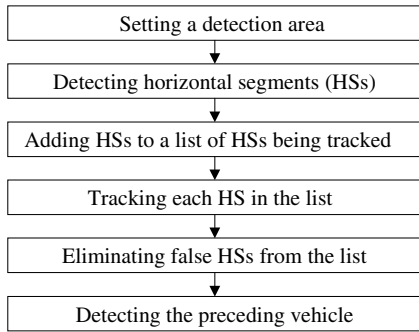


Figure 4. Flow of procedures for a preceding vehicle detection system.

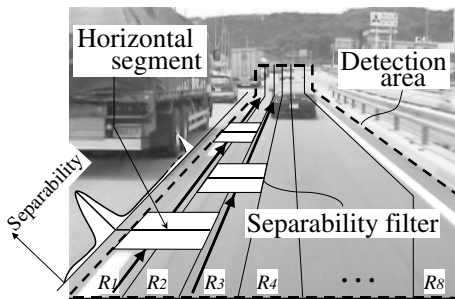


Figure 5. Detection of horizontal segments.

the image intensities of two arbitrary regions, with large values of separability signifying the existence of edges between the two regions. In order to extract horizontal edges, two vertically aligned regions are appropriate.

The detection area is first horizontally divided into eight even regions, R_1, \dots, R_8 (see Fig. 5), and a “separability filter” is set, which consists of two vertically aligned regions covering the two adjacent even regions, e.g. R_1 and R_2 . Next, the separability is computed while the position of the separability filter is shifted upward. Finally, separability filters whose separabilities are local maximums are selected and the border line segments between their two vertically aligned regions are determined to be horizontal segments.

Tracking of Horizontal Segments We track horizontal segments by finding correspondences between horizontal segments in the tracking list in a frame, t , and those detected in the next frame, $t + 1$, using a template matching technique. We chose the orientation code matching [20] (OCM) proposed by Ullah et al. as one of the robust template matching techniques. The similarity measure between any two images is defined based on the quantized orientation of the edges, which is referred to as the “orientation code” (OC). The similarity measure is the sum of the differences in OCs.

Considering that the border line between the ground

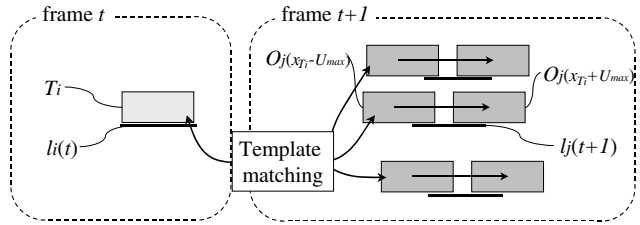


Figure 6. Tracking of a horizontal segment.

plane and the preceding vehicles, which we call the “vehicle bottom,” is detected as a horizontal segment, the template, T_i , is set to be the upper rectangular region for each horizontal segment in the tracking list in a frame, t (see Fig. 6). The size of the template is adaptively determined based on the size of the typical preceding vehicle in an image. For each horizontal segment, $l_i(t)$, in a frame, t , candidates for the corresponding horizontal segment are selected in the next frame, $t + 1$, around the position of $l_i(t)$. For each candidate, $l_j(t + 1)$, the sum of the differences in OCs is computed between the template, T_i , and the upper rectangular regions, $O_j(x)$, of $l_j(t + 1)$ for various horizontal positions, $\{x | x_{T_i} - U_{max} < x < x_{T_i} + U_{max}\}$, where x_{T_i} is the horizontal position of the template, T_i , and U_{max} is the maximum horizontal velocity between two successive frames. This one-dimensional search is applied to all the candidates, and the position with the smallest sum of the differences in OCs is determined to be the next position in the frame, $t + 1$.

Elimination of Horizontal Segments In order to reduce false detection, it is important to eliminate horizontal segments whose motions are incorrectly estimated.

We eliminate horizontal segments satisfying any of the following conditions:

1. The template, T_i , contains few edges.
2. The template, T_i , contains few corners.
3. The trajectory departs significantly from the constant motion model.
4. The motion is inconsistent with the motions of adjacent horizontal segments.

Template matching is unstable for conditions 1 and 2. Conditions 3 and 4 are for preserving the temporal consistency and the spatial consistency of motions, respectively.

Detection of Obstacles From the tracking list we select three horizontal segments, l_a , l_b , and l_c , whose horizontal positions are close to each other. The validity, s_{abc} , is computed based on Eqs. (8)-(15). We determine that t_2 and t_1 are, respectively, the current frame and the earliest frame where all of l_a , l_b , and l_c are in the tracking list.

If the horizontal shifts of l_a , l_b , and l_c between t_1 and t_2 are exceptionally different from each other or the vertical position order changes, they may not belong to a single

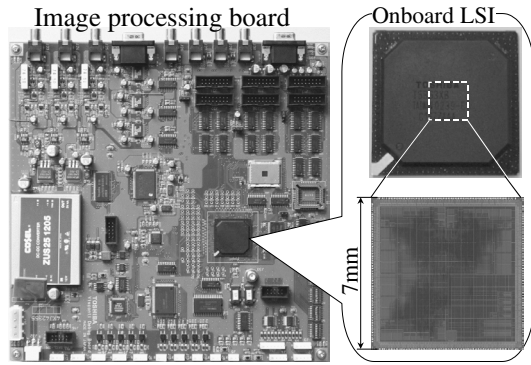


Figure 7. Onboard image processing system.

rigid object. Furthermore, if both D_{abc}^v and D_{abc}^h have large values, l_a , l_b , and l_c belong neither to the ground plane nor to the obstacle plane. We do not compute s_{abc} in such cases.

In order to give the system higher immunity against measurement errors in each frame, we accumulate the time series of the validity:

$$S_{abc}(t_2) = \sum_{t_1 < t \leq t_2} s_{abc}(t), \quad (19)$$

where $s_{abc}(t)$ denotes the validity in the frame, t . If $S_{abc}(t_2)$ is larger than a positive threshold, the three horizontal segments, l_a , l_b , and l_c , are candidate obstacles. Note that l_a belongs to several groups of three horizontal segments other than the group l_a , l_b , and l_c . If l_a is determined to be a candidate obstacle a sufficient number of times, we determine that l_a belongs to an obstacle.

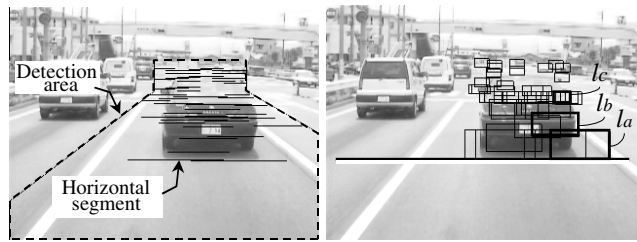
The vehicle bottom is estimated as the horizontal segment that has been determined to belong to an obstacle and is closest to the bottom of the image.

3.2. Implementation on the Onboard LSI

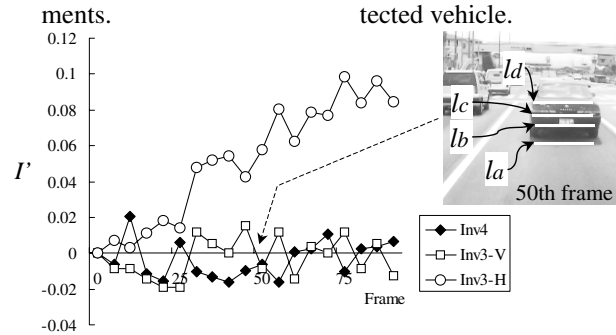
In order to achieve onboard image processing, we have developed an LSI (see Fig. 7) that has high image-processing performance and fulfills the specifications for onboard devices³. This LSI is a general-purpose image processor that is capable of executing any type of image processing by replacing the software.

Our LSI has three processing units and a data conversion unit. Each processing unit consists of a RISC processor core and a VLIW coprocessor that is capable of executing several types of SIMD operations. The data conversion unit converts a data set such as an image according to a conversion table that defines the conversion of each datum. Typical

³Our LSI works at temperatures between -40°C and 85°C with a power consumption of less than 1 W at an operation frequency of 150 MHz.



(a) Detected horizontal segments. (b) Tracking result and detected vehicle.



(c) Projective invariants.

Figure 8. Result of obstacle detection.

types of conversion include affine transformations of an image. Our preceding vehicle detection system employs this unit for computing the OC.

Fig. 7 shows a prototype image processing board using our LSI. The image processing board has three video input channels to provide video sequences for the LSI, and has a VGA video output channel to display the results of image processing.

Our LSI is capable of executing the proposed method, which is described in Section 3.1, at about 10 – 50 [ms/frame] using one processing unit and the data conversion unit. Thus, it is possible to execute other image-processing applications on the other two processing units. For reference, about 4 – 40 [ms/frame] is required when a standard PC with a Pentium III 933-MHz CPU is used without MMX instructions.

4. Experiments

Fig. 8 shows results of preceding vehicle detection obtained by our system described in Section 3. In Fig. 8(a), horizontal segments detected in the detection area are superimposed on the original image. The rectangles in Fig. 8(b) indicate tracking windows for the horizontal segments, and the horizontal long line is the estimated vehicle bottom. The accumulated validity, $S_{abc}(t)$, for the horizontal segments, l_a , l_b , and l_c , has a large positive value of 26, and the three horizontal segments are determined to belong to the preceding vehicle.

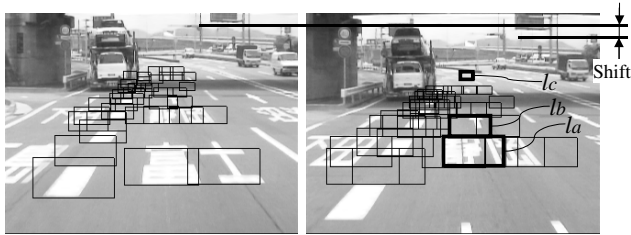


Figure 9. Texture on the ground plane.

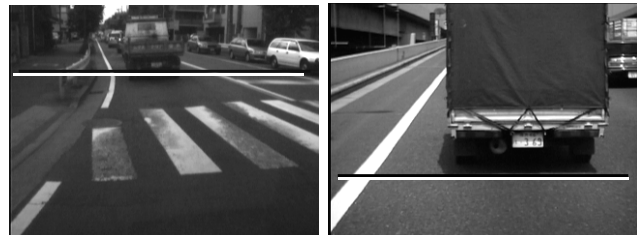
The graphs in Fig. 8(c) plot values of the cross ratio for the four horizontal segments, l_a , l_b , l_c , and l_d (shown as “Inv4”) and those for the three horizontal segments, l_a , l_b , and l_c , combined, respectively, with the vanishing line of the ground plane (shown as “Inv3-H”) and that of the obstacle plane (shown as “Inv3-V”)⁴. Note that we normalize the values as follows: $I'(f) = (I(f) - I(0))/I(0)$, where $I(f)$ is the value of the cross ratio at a frame f and $I'(f)$ is the normalized cross ratio. The Inv4 graph, whose normalized cross ratio stays around zero, indicates that the four horizontal segments are coplanar, and the obstacle detection methods that test coplanarity cannot detect the preceding vehicle when all the four segments belongs to the obstacle plane. Since $I'(f)$ of the Inv3-H graph has large value while that of the Inv3-V graph stays around zero, the proposed method based on three horizontal segments is capable of detecting the preceding vehicle because the proposed method discriminates the orientation of the plane that the three horizontal segments belong to.

Fig. 9 shows a result of obstacle detection for a scene containing road signs on the ground plane. In this scene, pitching of the onboard camera causes vertical shifts of the image. Nevertheless, the accumulated validity, $S_{abc}(t)$, for the horizontal segments, l_a , l_b , and l_c , has a small negative value of -21, which means that the group is correctly classified as the ground plane.

Figs. 10(a)-(f) show the results of preceding vehicle detection in various scenes. Our system is capable of detecting various types of vehicles in various scenes with poor lighting conditions and with disturbances such as raindrops and wipers on the windshield.

We evaluated the performance of our preceding vehicle detection system. Fig. 11(a) shows results of a simulation that computes the distances at which a vehicle approaching the observer was detected for various relative velocities. To observe sufficiently large motions in an image in the case of slowly moving vehicles, our system needs to track horizontal segments over a large number of frames. When the duration of tracking is short, e.g. $T=10$ [frame], due to poor

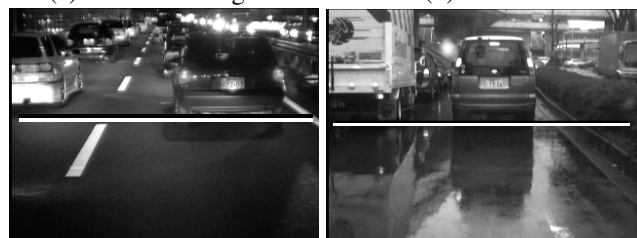
⁴The cross ratios combined with the vanishing line of the ground plane and that of the obstacle plane are $(\frac{1}{y_b} - \frac{1}{y_c}) / (\frac{1}{y_a} - \frac{1}{y_c})$ and $(y_b - y_c) / (y_a - y_c)$, respectively.



(a) Truck and cross walk. (b) Pick-up truck with black tarp covering.



(c) Vehicle cutting in. (d) Trailer.

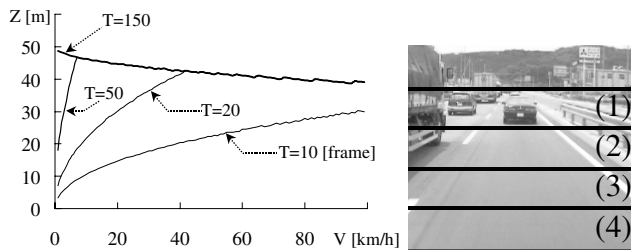


(e) Night scene. (f) Rainy scene.

Figure 10. Detection of various types of vehicles in various conditions.

lighting conditions, image noise, wipers, and so on, tracking of three horizontal segments is terminated before the system can detect obstacles, and the distances at which vehicles are detected therefore become short.

Fig. 11(c) shows results of obstacle detection for scenes of fine weather, rainy weather, and night, consisting of video sequences that are for 105, 40 and 75 minutes long, respectively. The image region below the vanishing line of the ground plane is divided into four vertically even regions, (1), (2), (3), and (4), as shown in Fig. 11(b), and the number of successful detections, the number of failures in detecting a preceding vehicle, and the number of misestimations of the vehicle bottom are shown in Fig. 11(c) for each vertically even region. The number of false detections of the ground plane as an obstacle is also shown. Preceding vehicles are successfully detected in most cases in regions (2), (3), and (4). However, in region (1) where the preceding vehicles are far ahead, it is difficult to stably detect them. Failure of detection and misestimation of the vehicle bottom occur under bad lighting conditions where preceding vehicles are not clearly seen, because the detection and the



(a) Distances (Z) at which an obstacle is detected for various relative velocities (V) and tracking durations (T). (b) Classification of vehicle positions.

Pos.	Fine (105 min.)			Rain (40 min.)			Night (75 min.)		
	○	×	P×	○	×	P×	○	×	P×
(1)	87	20	5	11	10	1	26	17	3
(2)	157	2	19	46	3	5	99	6	7
(3)	152	1	15	61	1	11	90	0	13
(4)	63	0	7	30	0	6	39	0	8
F	2			4			10		

(c) Results of obstacle detection (○: successful detection, ×: failure in finding an obstacle, P×: misestimation of obstacle position, F: false detection of the ground plane as an obstacle).

Figure 11. Performance evaluation.

tracking of horizontal segments are difficult in such scenes. False detection is also caused by failure in tracking horizontal segments.

5. Conclusions

We proposed an obstacle detection method that combines a projective invariant (cross ratio) and prior knowledge of road scenes. The prior knowledge consists of the vanishing lines of the ground plane and the obstacle plane. Our method is capable of detecting obstacles even when there are no feature points on the road and the onboard camera moves with pitching and yawing. Experimental results shows that our obstacle detection system is capable of detecting preceding vehicles, except for vehicles far ahead, in various real road scenes. We implemented the proposed method on a newly developed LSI that satisfies the specifications for onboard devices and achieves on-line image processing at 10 – 50 [ms/frame].

Acknowledgments

The authors would like to thank Prof. T. Kitahashi of Kansei Gakuin Univ. for his useful suggestions, and H. Hattori of Toshiba Corp. for his numerous comments.

References

- [1] M. Bertozzi and A. Broggi. Gold: A parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Trans. Image Processing*, 7(1):62–81, January 1998.
- [2] S. Carlsson and J.-O. Eklundh. Object detection using model-based prediction and motion parallax. *Proc. of ECCV*, pages 297–306, 1990.
- [3] W. Enkelmann. Obstacle detection by evaluation of optical flow fields from image sequences. *Proc. of ECCV*, pages 134–138, 1990.
- [4] A. W. Fitzgibbon and A. Zisserman. Multibody structure and motion: 3-d reconstruction of independently moving objects. *Proc. of ECCV 2000*, 1:891–906, 2000.
- [5] P. Fornland. Direct obstacle detection and motion from spatio-temporal derivatives. *Proc. of Int. Conf. on Computer Analysis of Images and Patterns*, pages 874–879, 1995.
- [6] K. Fukui. Edge extraction method based on separability of image features. *IEICE Trans. Inf. & Syst.*, E-78-D(12):1533–1538, 1995.
- [7] K. Hanawa and Y. Sogawa. Development of stereo image recognition system for ADA. *Proc. IEEE Intelligent Vehicle Symposium*, pages 177–182, 2001.
- [8] R. I. Hartley. In defense of eight-point algorithm. *IEEE Trans. on PAMI*, 19(6):580–593, 1997.
- [9] H. Hattori and A. Maki. Stereo without depth search and metric calibration. *Proc. CVPR*, 1:177–184, 2000.
- [10] T. C. H. Heng, Y. Kuno, and Y. Shirai. Active sensor fusion for collision avoidance. *Proc. of IROS*, 3:1244–1249, 1997.
- [11] Z. Hu and K. Uchimura. Real-time multi moving objects detection and tracking from dynamical road scene using FOE. *Proc. of 6th World Congress on ITS*, 2000.
- [12] T. Kato and Y. Ninomiya. An approach to vehicle recognition using supervised learning. *Proc. of Machine Vision Applications*, pages 77–80, 1998.
- [13] T. Kitahashi and H. Endo. A new method of 3-d motion analysis using a concept of projective geometry. *Proc. of IJCAI*, pages 902–904, 1985.
- [14] M. I. A. Lourakis and S. C. Orphanoudakis. Visual detection of obstacles assuming locally planar ground. *Proceedings of 3rd ACCV*, 2:527–534, 1998.
- [15] T. Morita and T. Kanade. A sequential factorization method for motion analysis. *IEEE Trans. on PAMI*, 19(8):858–867, 1997.
- [16] N. Pears and B. Liang. Ground plane segmentation for mobile robot visual navigation. *Proc. of IROS*, 2001.
- [17] J. Sato and R. Chipolla. Obstacle detection from image divergence and deformation. *Proc. of IV-96*, pages 165–170, 1996.
- [18] D. Sinclair and A. Blake. Quantitative planar region detection. *IJCV*, 18(1):77–91, 1996.
- [19] H. Takizawa and T. Ito. Method of preceding vehicles recognition using self organization map. *Technical Report of IEICE*, 101(302):23–28, 2001 (In Japanese).
- [20] F. Ullah, S. Kaneko, and S. Igarashi. Orientation code matching for robust object search. *IEICE Trans. on Inf. & Syst.*, E84-D(8):999–1006, 2001.
- [21] Z. Zhang, R. Weiss, and A. R. Hanson. Obstacle detection based on qualitative and quantitative 3d reconstruction. *IEEE Trans. on PAMI*, 19(1):15–26, 1997.