Proceedings of the
IEEE Intelligent Vehicles Symposium 2000
Dearborn (MI), USA • October 3-5, 2000

Stereo for 2D Visual Navigation

Hiroshi Hattori
Research & Development Center, TOSHIBA Corporation
Kawasaki 212-8582, Japan
kan.hattori@toshiba.co.jp

Abstract

We propose a new stereo method for 2D navigation without depth search and metric camera calibration. Although there is such a stereo method using the constraint that a vehicle moves on flat planes, the conventional method has the following drawbacks. First, it lacks of an effective contrivance to deal with camera vibrations and road inclination. Second, there is no measure to assess the danger of detected obstacles. To solve these problems, we develop a domain-specific stereo method which utilizes attributes of roads. We introduce a new camera model that is valid in road scenes in order to simplify the epipolar geometry giving a useful relationship between a pair of images. Furthermore, extraction of two parallel lines marking the driving lane enables us to correctly detect obstacles and even to estimate the degree of the danger. The proposed method is implemented on a Pentium II 450MHz PC and the current operating frequency is 15Hz for 320×240 images without MMX instructions. Through experiments we demonstrate that our method is applicable to a variety of road scenes.

1 Introduction

The ability to detect obstacles and estimate the degree of danger is essential for vehicle navigation. Stereo vision is an effective technique to extract 3D information from 2D images and applicable to visual guidance. While numerous stereo methods have been proposed in recent decades[15], two major issues are inherently involved in most of the conventional methods; *depth search* and *metric camera calibration.* The depth search or stereo matching which is a process to find corresponding points between a pair of images would require an extremely high computational cost and prevent real-time operation. The metric camera calibration on the other hand is a procedure to precisely determine a number of camera parameters including the camera 3D position and orientation with a calibration target whose 3D shape is known. Whereas the procedure yields an accurate determination of the camera parameters, it demands considerable time and labor. Therefore, to avoid those two problems is significant from a practical point of view and possible for some specific vision tasks. The thrust of this paper is to discuss a stereo method in order to detect obstacles and estimate the degree of danger for 2D navigation without depth search and metric calibration.

Techniques for obstacle detection are often based on the assumption that a vehicle or a mobile robot moves on flat planes such as roads and floors[2, 3, 5, 6, 8, 9, 11]. This is called the ground plane obstacle detection and defined as follows. Let one of stereo images be a reference one. Then we transform the other image to the reference view assuming that all image points arise from the ground plane. Comparing the reference image and the transferred one, while the points arising from the ground plane have no disparity, those on obstacle areas have disparities due to the height from the ground plane. Therefore, subtracting these images enables us to detect obstacle regions that have non-zero image difference and it is unnecessary to search for corresponding points. In this paper, we refer the planar transformation between a pair of images, which is important in this method, as the ground plane constraint or *GP constraint.*

Since most of mobile robots or vehicles move on a flat 2D plane, the ground plane obstacle detection is a key issue for many applications. Storjohann[3] introduced the stereo method for the control of an indoor mobile robot. Although the method was sensitive to noise involved in a pair of images, Onoguchi[2] proposed a practical contrivance for removing those areas wrongly detected as obstacle regions. Luong[11] integrated the stereo method with a lane marker detection and showed that the integration resulted in an improved performance. Bertozzi[5] used the obstacle detection method to navigate a vehicle for more than 3,000km. Unfortunately, however, their methods require fully metric calibration of stereo cameras.

0-7803-6363-9/00/$10.00 © 2000 IEEE

The GP constraint, generally represented by a 3×3 homography matrix $\boldsymbol{H}$, can be estimated given a set of corresponding points arising from the ground plane. Therefore, the metric camera calibration is unnecessary to locate obstacles in its field of view. Some researchers adopted the approach, which we also consider in this paper, based on the *weakly calibrated* stereo cameras, i.e., cameras for which only the epipolor geometry is known[6, 8, 9]. The weak calibration is much simpler than is the metric calibration since only point correspondences are necessary. Lourakis[6] and Okutomi[8] presented methods to automatically recover the GP constraint using a stereo matching technique. However, the conventional methods are subject to the following drawbacks. First, their methods utilize the general epipolor geometry summarized in the 3×3 *fundamental matrix* $\boldsymbol{F}$ [10, 12] rather difficult to determine. Second, there is no effective contrivance to deal with camera vibrations and road inclination. Third, although metric information is unavailable such as a Euclidean distance to the obstacle when stereo cameras are weakly calibrated, their methods lack of a measure to estimate the degree of danger of detected obstacles.

In order to alleviate these problems we present a domain-specific stereo method which utilizes some attributes of road scenes. We introduce the *pseudo-projective camera model* that provides a good approximation to the general camera model in road scenes and define the linear epipolar geometry for a pair of cameras. Furthermore, our method uses traffic lines in order to deal with camera vibrations and inclination, to overcome the noise and even to estimate the danger of obstacles assessed in terms of the time to contact. The proposed method is implemented on a Pentium II 450MHz PC and the current operating frequency is 15Hz for 320×240 images without MMX instructions. Through experiments we demonstrate that our method is applicable to a variety of outdoor scenes.

2 Epipolar geometry for road scene

In order to deal with the problem of camera vibrations and road inclination, the epipolar constraint or the geometric constraint for stereo images is indispensable. In this section, we introduce a new camera model that is valid in road scenes and define a linear epipolar geometry for a pair of cameras.

2.1 Pseudo-projective camera model

We propose the *pseudo-projective camera model* based on several proper assumptions for road scenes. That is, we define a road as rectangular area which

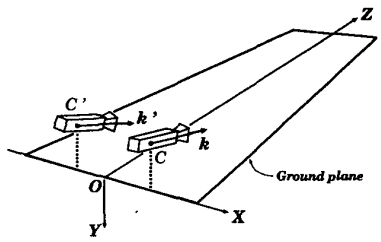


Figure 1: The stereo geometry.

is long in the depth direction and relatively narrow in the width. Also, we consider a restricted variation in the height direction since it is sufficient to define a 3D space that contains a fluctuating reference plane.

Considering a 3D point (X, Y, Z) projected to a 2D image point (u, v), the projective camera model[1] or the most general one defines this projection as

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (1)$$

where w is a non-zero scale factor. The world coordinate is defined as shown in Figure 1. We align XZ-plane with the ground plane where Y=0 and define X-axis as the projection of the line which goes through two camera origins, C and C'. Since the camera origin of each camera is represented as $(X_0, Y_0, 0)$, we derive

$$w = P_{31}(X - X_0) + P_{32}(Y - Y_0) + P_{33}Z, \quad (2)$$

where $\boldsymbol{k}$=$(P_{31}, P_{32}, P_{33})^\top$ represents the camera orientation as shown in Figure 1.

When the scene viewed by the camera is limited to a narrow range so that the approximation of $X - X_0 \ll Z$ and $Y - Y_0 \ll Z$ holds and the camera orientation $\boldsymbol{k}$ is closer to the Z direction than the others so that the relationship of $P_{31} < P_{33}$ and $P_{32} < P_{33}$ holds, the scale factor can be approximated as $w \simeq P_{33}Z$. Since the parameters P_{ji}(i=1∼4, j=1∼3) are defined up to the scale factor, we set P_{33} to 1 without loss of generality. Consequently, the pseudo-projective camera model is described as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad (3)$$

In this formulation we assume that images are distorted by variations in the Z direction, not by the horizontal or perpendicular direction. The pseudo-projective camera model therefore permits a linear

projection with respect to (X, Y). As shown in Figure 2, it implies that the parallelism is preserved in the horizontal(X) and vertical(Y) directions although not in the depth direction Z. The projective camera model treats the scale factor w as a function of (X, Y, Z) so that $w=f(X, Y, Z)$ and the affine camera model[4] assumes w to be a constant. On the other hand, the pseudo-projective camera model regards w as a function of only Z so that $w=f(Z)$. In this respect, the character of the pseudo-projective camera model is in between those of the affine and the projective camera model. The pseudo-projective camera provides a good approximation to the projective camera in a 3D space which is long only in the depth direction and it dramatically simplifies the formulation of the epipolar geometry as revealed in the next subsection. Moreover, it presents an effective invariant for noise reduction as described in subsection 3.2.

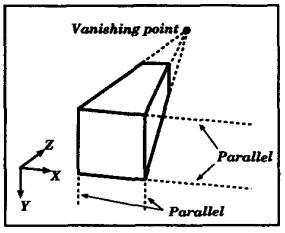


Figure 2: Pseudo-projective projection. The parallelism is preserved in the horizontal and vertical directions whereas not in the depth direction.

2.2 Affine epipolar constraint

Rewriting Equation(3) we have

$$\boldsymbol{u} = \boldsymbol{M}\boldsymbol{X} + Y/Z\boldsymbol{b} + \boldsymbol{c}, \tag{4}$$

where $\boldsymbol{u}=(u, v)^\top$, $\boldsymbol{X}=(X/Z, 1/Z)^\top$, $\boldsymbol{b}=(P_{12}, P_{22})^\top$, $\boldsymbol{c}=(P_{13}, P_{23})^\top$ and

$$\boldsymbol{M} = \begin{bmatrix} P_{11} & P_{14} \\ P_{21} & P_{24} \end{bmatrix}. \tag{5}$$

Analogously, we obtain the following equation for the other camera,

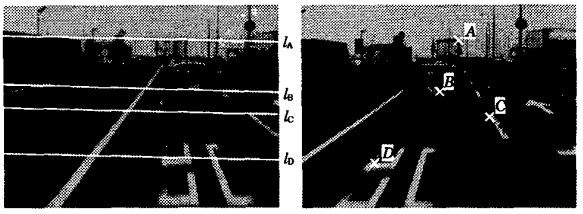
$$\boldsymbol{u}' = \boldsymbol{M}'\boldsymbol{X} + Y/Z\boldsymbol{b}' + \boldsymbol{c}', \tag{6}$$

where the primes indicate the parameters for the second camera. From Equation (4) and (6), we derive

$$\boldsymbol{u}' = \boldsymbol{A}\boldsymbol{u} + Y/Z\boldsymbol{d} + \boldsymbol{t}, \tag{7}$$

where $\boldsymbol{A} = \boldsymbol{M}'\boldsymbol{M}^{-1}$, $\boldsymbol{d} = \boldsymbol{b}' - \boldsymbol{A}\boldsymbol{b}$ and $\boldsymbol{t} = \boldsymbol{c}' - \boldsymbol{A}\boldsymbol{c}$. We then have

$$(\boldsymbol{u}' - \boldsymbol{A}\boldsymbol{u} - \boldsymbol{t}) \cdot \boldsymbol{d}^\perp = 0, \tag{8}$$



(a) Left image. (b) Right image.

Figure 3: The affine epipolar lines.

where $\boldsymbol{d}^\perp$ is a 2×1 vector orthogonal to $\boldsymbol{d}$. This equation can also be written as

$$f_1u' + f_2v' + f_3u + f_4v + f_5 = 0 \tag{9}$$

with $(f_1, f_2)^\top=\boldsymbol{d}^\perp$, $(f_3, f_4)^\top=-\boldsymbol{A}^\top\boldsymbol{d}^\perp$ and $f_5 = -\boldsymbol{t} \cdot \boldsymbol{d}^\perp$. Equation (9) is an approximation to the general epipolar geometry encapsulated by the 3×3 fundamental matrix. Estimation of the affine epipolar geometry is much simpler and more stable than is that of general epipolar geometry since the former has only four degrees of freedom and is linear with respect to the image coordinates $\boldsymbol{u}$ and $\boldsymbol{u}'$.

Figure 3 shows an example of the epipolar lines. The original stereo images are shown in Figure 4(a) and (b). The line l_A, l_B, l_C and l_D in the left view are the epipolar lines for the points A, B, C and D in the right view, respectively. The distance from the camera to the point A is about 24m and to the point D it is about 6m. Despite such a large depth variation the epipolar lines are parallel to each other and each epipolar line obviously lies close to the corresponding point. Actually, the average distance between each point and its epipolar line is 0.5 pixels. We use 29 point matches that are manually established. In the proposed method, only the parameters $f_i(i=1\sim5)$ are computed in advance.

3 Ground plane obstacle detection

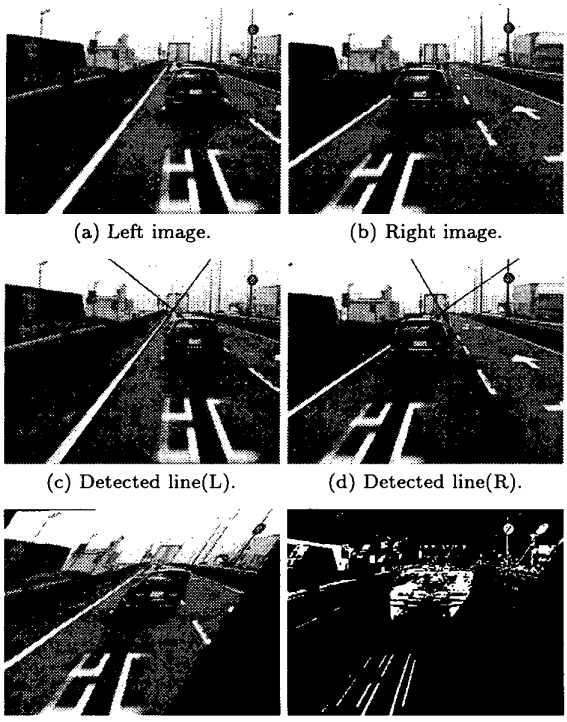
Our method dynamically updates the GP constraint using the two parallel lines on a road. In this section, we describe the GP constraint for a pair of pseudo-projective cameras and further introduce an effective scheme for noise reduction.

3.1 Affine GP constraint

Let $\boldsymbol{u}$ and $\boldsymbol{u}'$ be the corresponding points in a pair of images arising from the ground plane $Y=0$. Equation (7) is then

$$\boldsymbol{u}' = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{t}, \tag{10}$$

where $\boldsymbol{A}$ is a 2×2 matrix and $\boldsymbol{t}$ is a 2×1 vector. This equation represents the GP constraint for two



(a) Left image. (b) Right image.

(c) Detected line(L). (d) Detected line(R).

(e) Transferred image. (f) Detected areas.

Figure 4: The process of obstacle detection.

pseudo-projective cameras. The GP constraint, generally represented by a 3×3 homography matrix $\boldsymbol{H}$, is approximated by the 2D affine transformation under the pseudo-perspective camera model. This simplification contributes a great deal to a real-time operation since it is necessary to transform one of the stereo images using the GP constraint.

When the epipolar constraint is known, minimum of two line matches determine the *affine* GP constraint and we utilize two lines indicating the driving lane. Figure 4 shows the process of our obstacle detection. Figure 4(a) and (b) are input stereo images and two lines on the road are extracted in each image as shown in (c) and (d). Using the detected two lines and the epipolar geometry we estimate the affine GP constraint and transfer the left image as shown in (e). Detected obstacle areas are shown in (f) by subtracting image (e) from (b).

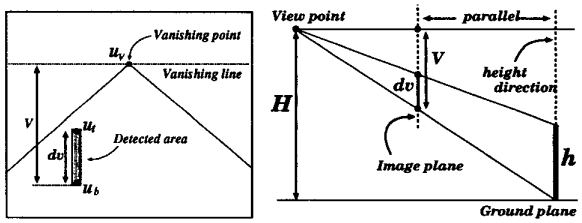
3.2 Noise elimination

In theory it is possible to separate obstacle areas from a free space using the determined GP constraint with the extracted two lines. However, since the lines are often broken or occluded, it is inevitable to have erroneous estimation such as near the boundaries of road markers. For example, in Figure 4(f) some texture areas on the road are wrongly detected as obstacle areas. Although more correspondences found by using a stereo matching technique make the GP constraint estimation more accurate[6, 8], it requires an exhaustive computational cost. Instead, we present a practical and effective scheme for noise reduction.

We determine whether detected areas arise from *real* obstacles or not based on the height information. Since we use weakly calibrated stereo cameras, the height information is basically unavailable. Nevertheless, the relative height information is measurable with the vanishing point of two lines which segment the driving lanes as follows. Consider that the rectangular gray area spanning between $\boldsymbol{u}_t$ and $\boldsymbol{u}_b$ is detected as shown in Figure 5(a). Assuming a roll angle of the camera to be small, the vanishing line for the ground plane becomes a scan-line passing through the vanishing point, $\boldsymbol{u}_V$. This assumption is practical because it is feasible to set up the stereo cameras in such a way that the vanishing line lies almost horizontally. Moreover, the road inclination is generally small in the horizontal direction and the rolling of the camera caused by its own movement is relatively minor compared to the pitching. Let the height of the detected area in the image be dv, and the distance between the bottom of the detected area $\boldsymbol{u}_b$ and the vanishing line be V. Also let the height of the camera and the obstacle with respect to the ground plane be H and h, respectively, as shown in Figure 5(b). Since the pseudo-projective camera model is linear about the height direction, the image plane can be regarded as parallel to the direction. Then we have

$$\eta = h/H = dv/V. \tag{11}$$

For each detected area, we compute the relative height η and reject the area as noise if η is below a threshold.



(a) Image view. (b) Side view.

Figure 5: The relative height η=h/H is an invariant under the pseudo-projective projection.

4 Time to contact

In order to realize safe navigation, it is necessary to estimate the degree of danger due to the detected ob-

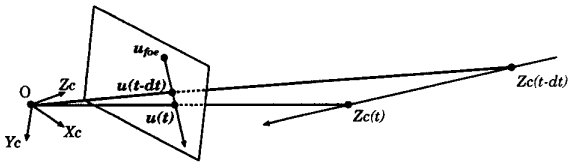


Figure 6: Time to contact for 3D motion.

stacle. Although a *Euclidean distance* is often utilized as such a measure, it is not generally appropriate by itself. For example, the degrees of danger incidental to two obstacles at the same distance are obviously different if one is approaching the observer and the other going away.

The time to contact is useful to estimate the danger more effectively since it is defined taking the relative motion between the obstacle and the observer into consideration. By definition, the time to contact is the amount of the time remaining before an obstacle collides with the camera in case the obstacle and the camera maintain the relative translational velocity. Considering the camera coordinate (X_c,Y_c,Z_c) as shown in Figure 6, the time to contact t_c is defined as

$$t_c = \frac{Z_c(t)}{Z_c(t-dt) - Z_c(t)} dt, \tag{12}$$

where $Z_c(t-dt)$ and $Z_c(t)$ are depths of an obstacle at time $t-dt$ and t, respectively. Since the *cross-ratio* is preserved under projective transformations[1], Equation (12) can also be written as

$$t_c = \frac{|\boldsymbol{u}(t-dt) - \boldsymbol{u}_{foe}|}{|\boldsymbol{u}(t) - \boldsymbol{u}(t-dt)|} dt, \tag{13}$$

where $\boldsymbol{u}(t-dt)$ and $\boldsymbol{u}(t)$ are image points of the obstacle at time t and $t-dt$, respectively, and $\boldsymbol{u}_{foe}$ is the focus of expansion. This equation suggests the convenient aspect of time to contact, i.e., t_c can be estimated by the movement of the obstacle in the image and the focus of expansion and the Euclidean distance is unnecessary. Thus the essential problem is then to establish correspondences between images at different time. Although several researchers have addressed the employment of the time to contact[7, 13, 14], the estimation is considerably difficult in general, requiring a time-consuming dense flow field recovery or point correspondences.

On the other hand, the method in the framework of ground plane obstacle detection makes the estimation of time to contact straightforward. Consider that an obstacle approaches our vehicle as shown in Figure 7. The obstacle detection method provides the boundary

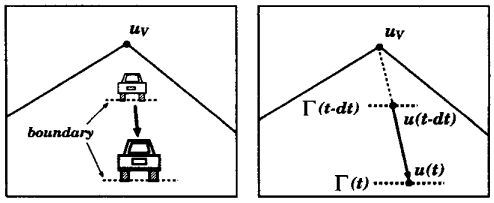


Figure 7: Time to contact for 2D motion.

between the car in front and the road at each time instant. Let $\Gamma(t-dt)$ and $\Gamma(t)$ be the boundaries at time $t-dt$ and t, respectively. Given that both the observer and the obstacle move along the traffic line, since the focus of expansion coincides with the vanishing point $\boldsymbol{u}_V$ where the two parallel lines intersect, the movement of the obstacle is determined. In this paper, we examine the method that treats the simple case although the extension to deal with a more complex movement is also under work.

5 Experiments

We have performed some experiments in which our stereo method was applied to outdoor images under various conditions. Figure 8 shows a sketch of a vehicle used in our experiment. The distance between the two cameras is about 1.15m and each camera is at the height of about 1.5m. We use CCD cameras with 12mm lenses that have image resolution of 320×240 pixels and drive the vehicle at about 100km/h on a highway. Note that these parameters are not utilized in our method and the cameras are not strictly mounted in parallel.

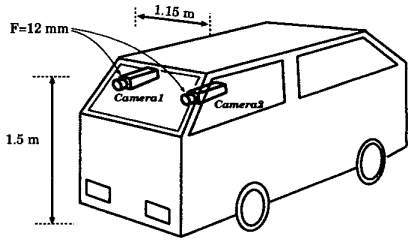
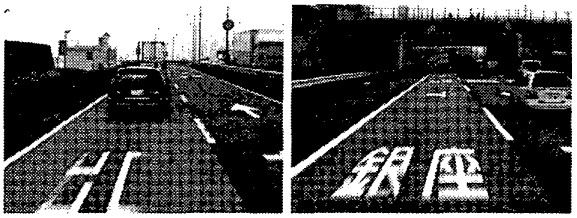


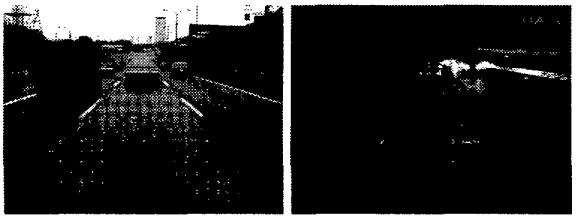
Figure 8: Setup of stereo cameras.

Figure 9 shows examples of output images picked up from sequences under various conditions. Each sequence is 300 frames long. In these experiments we separate the free space from obstacle area in the driving lane as indicated by the white area. Through all these examples, no texture on the road is wrongly detected as an obstacle and a free space is correctly extracted in spite of camera vibrations, road inclinations, shadows, road textures and illumination change.

(a) Motorcycle 1. (b) Car 1.

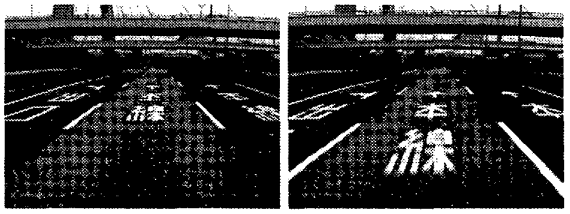


(c) Car 2. (d) Truck 1.

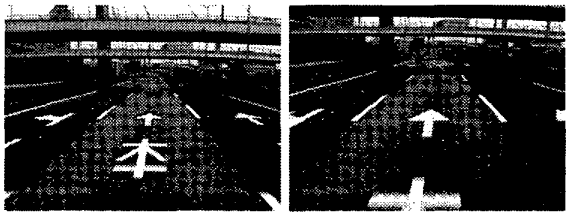


(e) Truck 2. (f) Night.

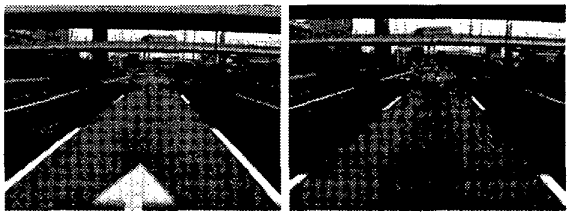
Figure 9: Examples of output images.



(a) Frame 030. (b) Frame 034.



(c) Frame 038. (d) Frame 042.



(e) Frame 046. (f) Frame 050.

Figure 10: Output image sequence of "Vibration".

Figure 10 shows an output image sequence in which strong vibrations due to bumps on the road are involved whereas there is no obstacle. Figure 11 shows the vertical motion of vanishing point computed from the two lines marking the lane between frames 0 and 80. Crucial variations are observed due to vibrations between frames 30 and 50. The maximum and minimum image coordinates are 82 and 56 pixels, respectively. The results shown in Figure 10 are picked up from these frames. In spite of strong vibrations, no texture on the road is mistakenly detected. Figure 12(a) shows a derived image difference using a fixed GP constraint determined in advance where much of the markers in the center of the road are wrongly detected. On the other hand, Figure 12(b) shows a result using the GP constraint updated with the two lines where the noise areas dramatically decrease owing to the compensation. In fact, while there are 5339 image points that have non-zero difference in the road in (a), the number is 2202 in (b). Namely, about 60% of the image points that are wrongly detected as obstacle area is reduced in this example.

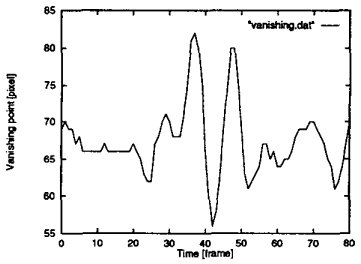
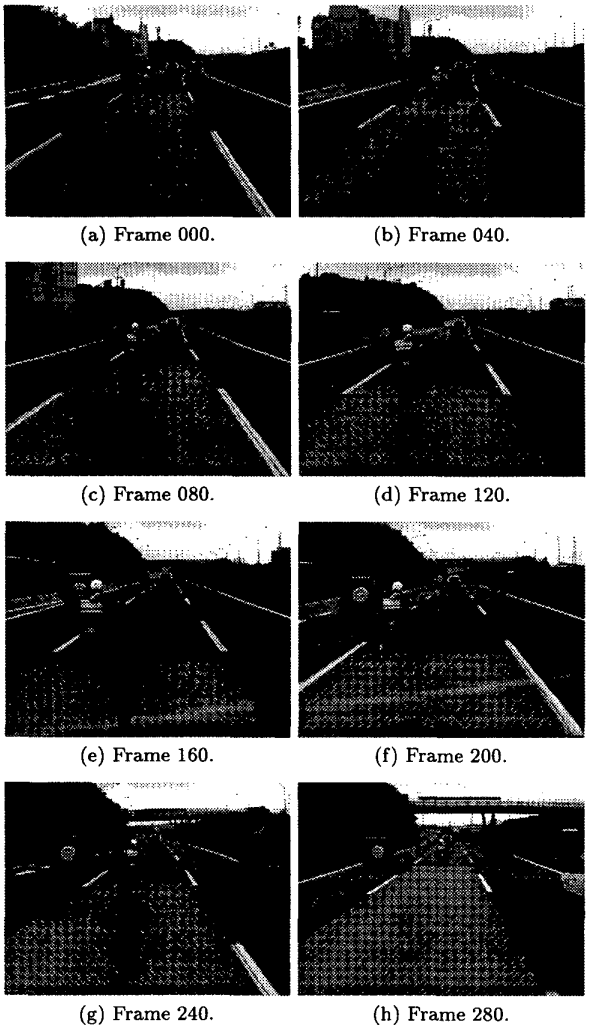


Figure 11: The vertical motion of vanishing point.

(a) Without correction. (b) With compensation.

Figure 12: (a) Markers on the road are falsely detected. (b) Such areas are considerably suppressed.



(a) Frame 000. (b) Frame 040.

(c) Frame 080. (d) Frame 120.

(e) Frame 160. (f) Frame 200.

(g) Frame 240. (h) Frame 280.

Figure 13: Output image sequence of "Motorcycle 2".

Figure 13 indicates an image sequence including a motorcycle as an obstacle. The motorcycle is approaching between roughly frames 0 and 150, relatively stands still to our vehicle between frames 150 and 170, and then goes away. Figure 14 shows the time-variation of the boundary line between the free space and the obstacle (motorcycle). The bold line indicates the estimate by our method and the dotted line is for the ground truth obtained manually. The computed result closely follows the ground truth and the average error is 2.9 pixels. The proposed method not only separates the free space from obstacles but also estimates the time to contact as described in Section 4. Figure 15 shows the time-variation of the time to contact from frame 60 to 140. As the motorcycle approaches our vehicle at a roughly uniform relative velocity in these frames, the time to contact should decrease linearly. Our method results in reasonable estimation since the profile is almost linear as seen in Figure 15.

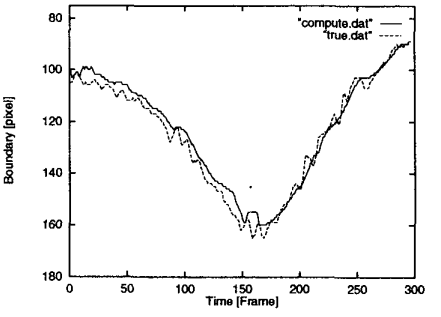


Figure 14: The vertical movement of the boundary. The bold line indicates the estimate by the proposed method, and the dotted line is obtained manually.

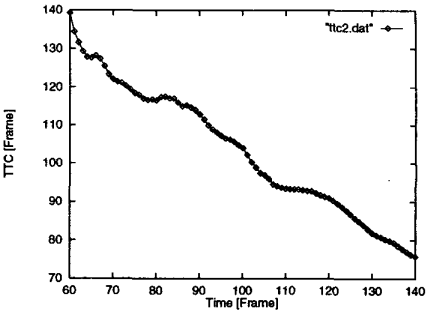
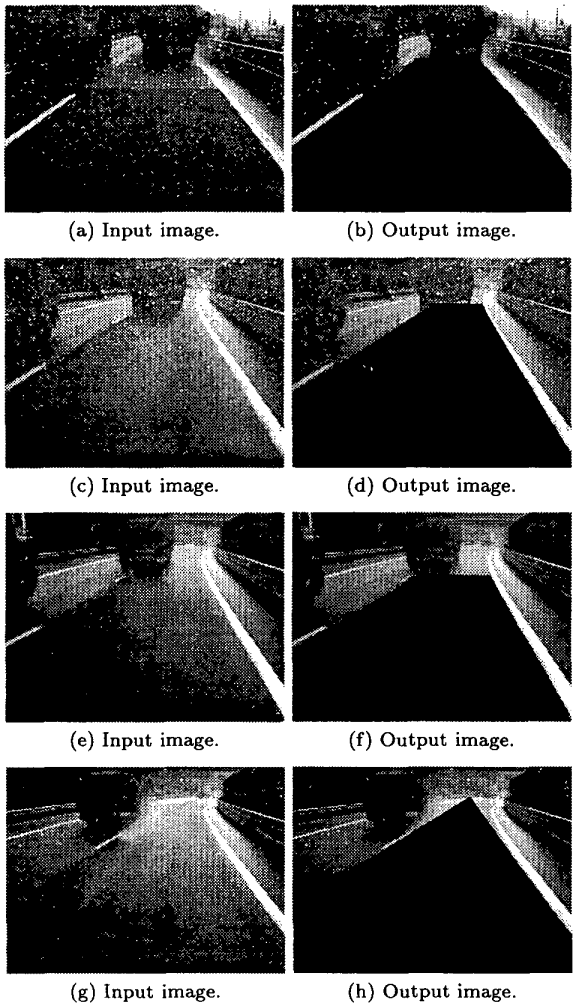


Figure 15: Time to contact versus time. For uniform motion the time to contact should decrease (or increase) linearly. The output profile of the time to contact is close to be linear.

Figure 16 shows input and output image sequences in rain. In this figure, the detected free space is indicated by the black mask to show the results more clearly. Our method accurately identifies the free space in spite of heavy rain whereas such a situation is rather crucial for the stereo matching. This result shows a clear advantage of our method over conventional methods requiring depth search in terms of the robustness as well as the computational cost.



(a) Input image. (b) Output image.

(c) Input image. (d) Output image.

(e) Input image. (f) Output image.

(g) Input image. (h) Output image.

Figure 16: Input and output image sequences of "Rain". The truck in front is in the driving lane and then moves to the left. Free space in the driving lane is indicated by the black mask for visualization.

6 Summary

In this paper we have proposed an application-specific stereo method to realize safe navigation, requiring neither depth search nor metric camera calibration. We have introduced the pseudo-projective camera model that is valid in road scenes and simplified the epipolar geometry for a pair of cameras. Furthermore, we have shown that the extraction of two parallel lines on the road is effective to correctly detect obstacles and even to estimate the degree of the danger. The experimental results have demonstrated our method is applicable to a variety of road scenes. In future work, we are planning to combine our method with the technique to detect a curved line[11] in order to alleviate the limitation to the scene. Also, we will develop the method to estimate the time to contact under a more complex motion.

References

[1] J.L.Mundy and A.Zisserman, editors. *Geometric invariance in computer vision.* The MIT Press, 1992.

[2] K.Onoguchi, N.Takeda, and M.Watanabe. Planar projection stereopsis method for road extraction. In *Proc.IROS*, pages 249–256, 1995.

[3] K.Storjohann, Th.Zielke, H.A.Mallot, and W.von Seelen. Visual obstacle detection for automatically guided vehicles. In *Proc. ICRA*, pages 761–766, 1990.

[4] L.S.Shapiro. *Affine Analysis of Image Sequences.* PhD thesis, Dept. Engineering Science, University of Oxford, 1993.

[5] M.Bertozzi and A.Broggi. GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing*, 1997.

[6] M.I.A.Lourakis and S.C.Orphanoudakis. Visual detection of obstacles assuming a locally planar ground. In *Proc. 3rd ACCV*, volume II, pages 527–534, 1998.

[7] M.I.A.Lourakis and S.C.Orphanoudakis. Using planar parallax to estimate the time-to-contact. In *Proc. CVPR*, volume II, pages 640–645, 1999.

[8] M.Okutomi and S.Noguchi. Extraction of road region using stereo images. In *Proc. ICPR*, pages 853–856, 1998.

[9] M.Xie. Ground plane obstacle detection from stereo pair of images without matching. In *Proc. 2nd ACCV*, volume II, pages 280–284, 1995.

[10] O.D.Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. 2nd ECCV*, pages 563–578, 1992.

[11] Q.-T.Luong, J.Weber, D.Koller, and J.Malik. An integrated stereo-based approach to automatic vehicle guidance. In *Proc. 5th ICCV*, pages 52–57, 1995.

[12] Q.-T.Luong and O.D.Faugeras. The fundamental matrix: theory, algorithms, and stability analysis. *IJCV*, 17(1):43–76, 1996.

[13] R.Cipolla and A.Blake. Surface orientation and time to contact from image divergence and deformation. In *Proc. 2nd ECCV*, pages 187–202, 1992.

[14] R.C.Nelson and J.Aloimonos. Using flow field divergence for obstacle avoidance: towards qualitative vision. In *Proc. 1st ICCV*, pages 188–196, 1988.

[15] U.R.Dhond and J.K.Aggarwal. Structure from stereo - a review. *Trans. Systems, and Man Cybernetics*, 19:1489–1510, February 1989.