SICE Annual Conference in Fukui, August 4-6, 2003
Fukui University, Japan

An Obstacle Detection Method by Fusion of Radar and Motion Stereo

Takeo Kato[1], Yoshiki Ninomiya[2], Ichiro Masaki[3]

1 Toyota Central R&D Labs., Inc. Nagakute, Aichi, 480-1192 JAPAN. takeo@milab.tytlabs.co.jp.
2 Toyota Central R&D Labs., Inc. Nagakute, Aichi, 480-1192 JAPAN. nino@milab.tytlabs.co.jp.
3 Massachusetts Institute of Technology, Cambridge MA 02139 USA. masaki@mit.edu.

Abstract: An object detection method based on the motion stereo technique is proposed, which works by the fusion of millimeter wave radar and a single video camera. The method does not depend on the appearance of the target object, so it is not only capable of detecting automobiles, but also other objects that come within its range.
Keywords: Motion stereo, millimeter wave radar, object detection, sensor fusion

1. Introduction

Adaptive Cruise Control (ACC) Systems, which maintain the distance at which a car follows the preceding vehicle, are available from automobile manufactures. The development of automotive driver assistance systems, such as stop & go ACC systems and obstacle avoidance systems, is now in progress. These systems have to take some action when an object lies in the path of the subject vehicle, such as warning the driver, reducing the speed or changing the course of the vehicle. In order to achieve these functions, the system needs to be able to detect lane markings in order to estimate the course of the vehicle and to detect objects that may be obstructing the road in the area ahead of the vehicle. This paper deals with the detection of objects in the path of the vehicle.

To find the exact area that an object occupies it is necessary to measure both an accurate distance from the subject vehicle to the object and to detect the position of the lateral boundaries of the object, i.e. the left- and right-hand sides of the target object (Fig. 1). Currently available sensing methods, such as radar [1)] and stereo vision [2)] are insufficient for this purpose.

Radar can measure an accurate distance to an object and has robustness against bad weather, e.g. rain and fog; however, it does not have enough lateral resolution to find its boundaries. Expensive laser scanning radar is capable of detecting the area occupied by of an object [3)]. However, the type of laser radar that is affordable for automotive purposes can only reliably detect reflectors, and therefore it does not have the capability to quantify the target area that the object occupies. In addition, even if an object can be detected by radar, a camera is also necessary for the detection of lane markings [4)] to establish whether the detected object lies on the course of the subject vehicle or not.

On the other hand, vision sensors have sufficient lateral resolution to find the boundaries of an object, but distance measurement by vision sensors, i.e. stereo vision, is much less accurate than radar.

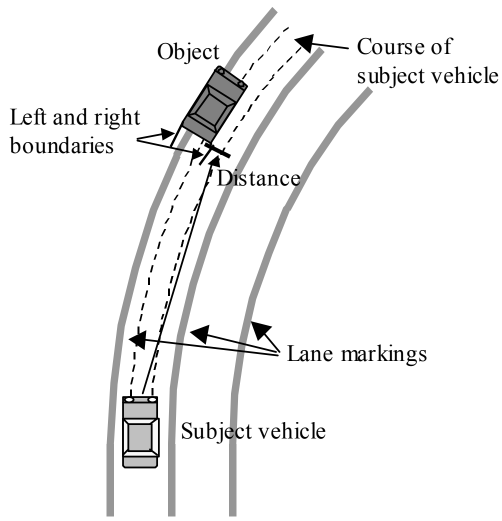


Fig. 1. Object and lane markings.

If we take these complementary characteristics into consideration, a combination of radar and stereo vision might solve the problem, but such a system would be expensive due to the two expensive sensors that it incorporates. For this combination, the role of the vision sensor is to detect the boundaries of the object, while the distance to it is measured accurately by the radar. Therefore, as a less expensive option, a combination of radar- and vision-based sensing using a single camera is promising.

Pattern-based methods have been already been reported for object detection using a vision sensor. These use features that are characteristic of the appearance of the rear of an automobile, e.g. horizontal edges [5)], a combination of vertical edges and horizontal edges [6)] and template matching techniques [7)]. These features are significant when the object is a passenger car, but they are less applicable when the object is another type of vehicle, such as a bus, a truck or a trailer. Furthermore, a motorcyclist or a pedestrian has very few features in common with a passenger car at all.

Therefore, we attempted to develop a method that does not use features based on the appearance of an object [8), 9)]. In our method, detection is performed based on the motion of the object. A brief description is as follows. Objects are initially detected with the radar, and feature points whose motion is easy to track are selected and tracked frame by frame in an image sequence. After this, regions are located within the image for which the distances match the distances measured by the radar. The distance estimation is derived from the motion in the image using a motion stereo technique.

Normally, motion within an image is used to detect objects in its outer regions, e.g. to detect vehicles that are cutting in [10)], because the relative motion in these parts is bigger than it is in the center of the image. In addition, the further away the object is, the less its apparent motion in the image becomes, and so it is more difficult to detect. In this paper, we will show that objects up to 50m ahead of the subject vehicle can be detected by our method, even if they are in the center region of an image.

The remainder of this paper is organized as follows. In Section 2, an explanation is given relating to motion stereo. The details of our proposed method are described in Section 3. Experimental results are presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Motion Stereo

Motion stereo is a technique to recover a 3-D structure from parallaxes as well as binocular stereo vision. Motion stereo uses parallaxes due to camera motion instead of using plural cameras. In order to determine the parallax of each point in an image it is necessary to find a correspondence between a sequence of images. One of the drawbacks of binocular stereo is that mis-correspondences might occur if a repetitive pattern exists in the search area. In the case of motion stereo, the chance of a mis-correspondence occurring is less than it is for binocular stereo because the necessary search area is narrow and the correspondence can be tracked from frame to frame, since the view-point moves continuously.

Object detection methods that make use of motion in an image sequence captured by a forward-viewing camera have already been reported. In 10), passing cars that appear nearby are detected by changes in the image. In 11) and 12), the road surface is assumed to be planar and regions in the image where the motion differs from the motion of the planar road surface are detected as objects. However, although the assumption might hold for highways, urban roads are bumpier than highways and cannot be treated as planar. In addition, these methods may be able to detect objects, but they cannot estimate the distance from them to the camera .

Assuming that most objects consist of planar surfaces, another motion stereo method has been reported, which attempts to measure the distance to a vertical plane orthogonal to the optical axis [13)]. In 14), measurement of the distance to a pole at the roadside is presented. In these methods, the distance to the plane or the pole is estimated from the motion of two points on the object in the image and the camera velocity along the optical axis, as:

$$\hat{Z} = \frac{-b \cdot \Delta Z}{\Delta b}, \tag{1}$$

where b is the distance between the two points in the image plane, the projection of a length B in 3-D coordinates; Δb is the change in b in time interval Δt, and $\Delta Z = v\Delta t$, where v is the camera velocity along the optical axis. v can be computed from the velocity of the vehicle on which the camera is mounted (Fig. 2).

To estimate the distance to an object using this method, we first need to choose two points that belong to same object. However, the appropriate method for choosing these points is not described at all in either 13) or 14). In addition, to estimate the distance $\hat{Z}$ with (1), the change in the distance ΔZ is necessary. ΔZ can be computed from the velocity of the vehicle if the object is stationary, but it is impossible to compute ΔZ in the same way if the object is moving.

3. Proposed Method

The method described in 13) or 14) can estimate the distance to an object if two points on the object appear in an

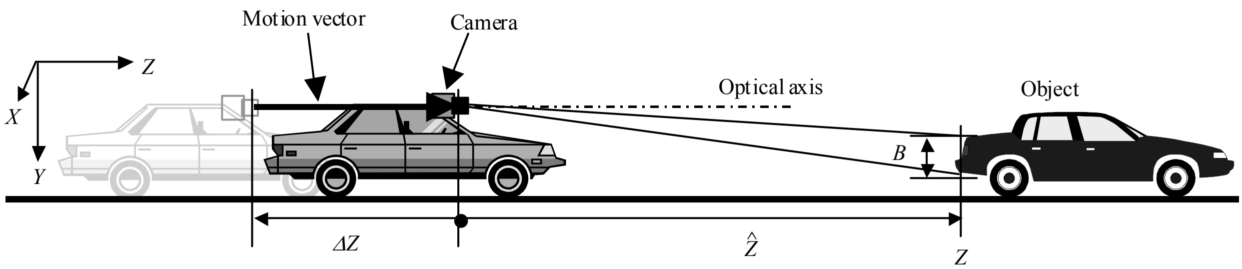
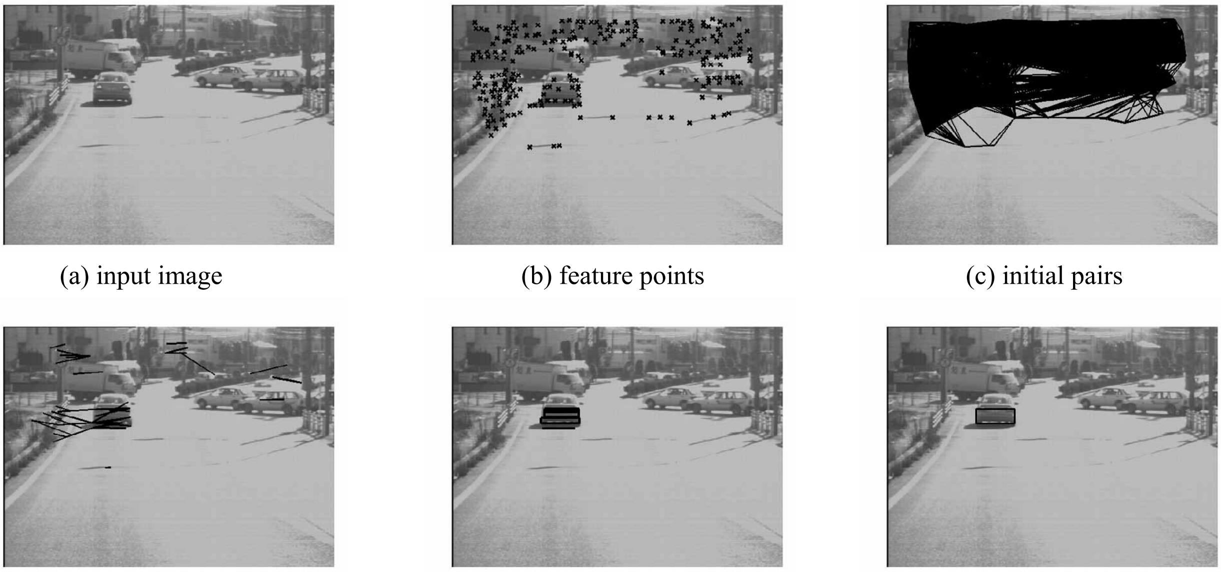


Fig. 2. Motion stereo for automotive.



(a) input image (b) feature points (c) initial pairs

(d) selected pairs (e) rectangles (f) detected region

Fig. 3. Example of detection process.

image. However, there are two problems: how to find the two points on the object and how to compute the change in the distance ΔZ if the object is moving. In our method, these problems are solved by fusion with radar.

The object is initially detected by the radar and then the region corresponding to it is detected in the image. In the process of detecting this region, a virtual vertical plane is assumed at the distance measured by the radar. The feature points, whose motion is easy to track, are selected and tracked frame-by-frame in the image sequence. Instead of estimating the distances from them, their motion in the image are estimated on the assumption that they were in the virtual plane. Then, by comparing the actual motion with the estimated one, the region in which all the actual and estimated changes are equal is detected in the image.

Additionally, the change in the distance to the object ΔZ can be ascertained even if it is moving, because the radar measures the distance to the object at each time step.

The detail of the method is explained below (Fig. 3):

Step 1. Selection and tracking of feature points (Fig.3(b)) in an image sequence captured by an onboard camera [15]. A brief description of this is as follows. Firstly, feature points that are easy to track are chosen in the initial image frame by evaluating the second-order derivatives of the intensities in the windows whose centers are the feature points. After choosing the points, template matching to the next frame is carried out on all the windows. Tracking of the points is achieved by repeating the template matching frame by frame.

Step 2. Produce pairs of feature points (Fig.3(c)) that satisfy

$$W_{\min} \le b_x < W_{\max}\,,\; H_{\min} \le b_y < H_{\max} \tag{2}$$

where b_x and b_y are the x and y components of b, respectively and $W_{\min}$, $W_{\max}$, $H_{\min}$ and $H_{\max}$, are appropriate constants that are assigned in accordance with the size of the targets.

Step 3. In order to select those pairs (Fig.3(d)) that may exist with high probability at the distance measured by the radar, select pairs (P_i, P_j) that satisfy

$$\varepsilon^2(P_i, P_j) = \left|\Delta b(P_i, P_j) - \Delta\hat{b}(P_i, P_j)\right|^2 < \theta_1 \quad , \tag{3}$$

where $b(P_i, P_j)$ is the length of the pair (P_i, P_j), $\Delta b(P_i, P_j)$ is the actual change in the length in a period Δt, and $\Delta\hat{b}(P_i, P_j)$ is the estimation of it

$$\Delta\hat{b}(P_i, P_j) = \frac{-b(P_i, P_j)\cdot \Delta Z}{Z}, \tag{4}$$

where Z is the distance to the virtual plane and ΔZ is the change in Z in Δt (Fig.2).

Step 4. Screen out the pairs that do not satisfy

$$\sum_{i=2}^{N}\sum_{j=1}^{i-1}\varepsilon^2(P_i, P_j) < \theta_2, \tag{5}$$

where $\{P_i \mid i = 1,...,N\}$ are in a rectangle (Fig.3(e)) whose diagonal is a pair, because points close to the pair should also be on the object if the pair is on the object.

Step 5. Compute the means $(\tilde{x}, \tilde{y})$ and the standard deviations (σ_x, σ_y) of the top-left and bottom-right corners of the rectangles and choose the largest rectangle (Fig.3(f)), where each corner is within

$$(\tilde{x} \pm \alpha\sigma_x, \tilde{y} \pm \alpha\sigma_y)\,, \tag{6}$$

as the final object region; here α is an appropriate constant.

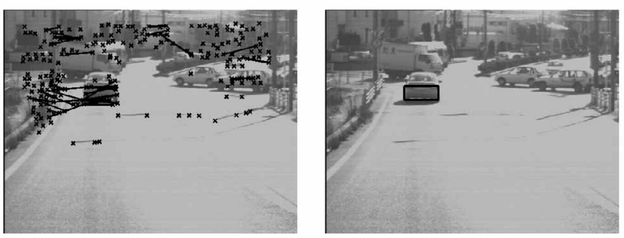
4. Experiment

The method was evaluated using scenes in which the targets were approximately 50m ahead of the subject vehicle. The detection of a stationary vehicle, a moving vehicle, a stationary pedestrian and a 350-ml can on a pedestrian crossing were examined in order to confirm that: 1) the method was independent of the appearance of the objects, 2) that it could detect both stationary and moving objects, and 3) that it gave no false detections. For each target, a 3 second sequence was evaluated.

The evaluation data contained range information from millimeter wave radar and image sequences captured by a video camera. The radar was a Fujitsu-Ten mechanical-scan millimeter wave radar system, which has a detection range of 100 m and a horizontal scanning angle of 14 degrees. Its range and lateral resolution are 0.1m and 2 degrees, respectively. The camera was a XC-7500 with a 25mm focal length lens; the viewing angle was about 15 degrees horizontally and 11 degrees vertically. The sampling resolution of the image was 640 × 480 pixels, the depth was 256-level grayscale and the frame rate was 30 frames/s.

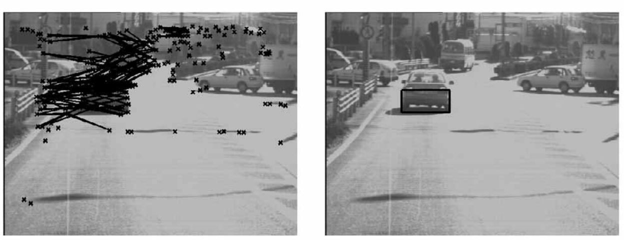
The parameters that were used in the detection process described in Section 3, were as follows. The time interval Δt to produce a parallax was 20 frames, i.e. 0.66 seconds. The values of $W_{\min}$, $W_{\max}$, $H_{\min}$ and $H_{\max,}$ as described in (2), were defined as being between 0.2 m (W) × 0.5 m (H) and 2.2 m (W) × 2.7 m (H), since the range of target sizes that we considered ranged from a human to a large vehicle like a truck. θ_1 in (3) and θ_2 in (5) were selected as 1.0, and α in (6) was set to 2. These variables were determined empirically.

Examples of the detection results are shown in Fig. 4 to Fig. 7. In the figures, the results of each 20-frame interval are shown from top to bottom. In each row, marks (×) and lines in the left image denote the tracked feature points and the pairs that have remained after Step 3, respectively. The rectangle in the right-hand image shows the region that was detected finally; the number below the images indicates the



55.1m

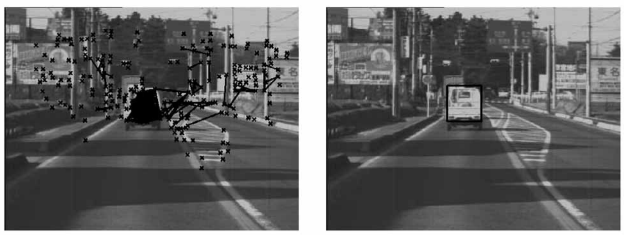
47.2m



39.7m

Fig. 4. Stationary vehicle.

44.9m



46.0m

47.4m

Fig. 5. Running vehicle.

distances measured by the radar. In Fig. 4, Fig. 5 and Fig. 6, some pairs that do not belong to the targets are detected in the intermediate process. However, the final detected regions do not include them and they approximately match with the targets. As for the sequence showing the can (Fig. 7), the radar detected the can and the feature points were tracked at the pedestrian crossing, and therefore some pairs were maintained in the intermediate process. However, through the sequence, no false detections were observed in the final result.

Table 1. Errors in boundary detection

Target	Left	Right
Stationary vehicle	0.01 ± 0.03m	-0.03 ± 0.05m
Moving vehicle	0.12 ± 0.02m	0.04 ± 0.06m
Pedestrian	0.11 ± 0.08m	-0.04 ± 0.08m

Fig. 8, Fig. 9 and Fig. 10 show the detected boundary positions in the images and the true boundary positions observed by eye. Table 1 shows the average error and standard deviation of each detected boundary on a real scale, which were computed from the distance measured by the radar and the difference between the detected boundary and the true boundary in an image. Table 1 indicates that the proposed method can perform a boundary detection process that is accurate to within 0.2m.

5. Conclusions

In order to avoid colliding with an object by using just a driver assistance system, the detection of lane markings to estimate the course of the subject vehicle and detection of an occupied area where an object obstructs the road are both necessary. In this paper, a method for detecting areas of the road that are occupied by an object is presented. This is achieved by the fusion of millimeter wave radar and a single video camera. To find the exact area occupied by an object, accurate measurement of the distance from the subject vehicle to the object and detection of its side boundaries are necessary.

The distance can be measured accurately by radar, so we focused on the development of a method for detecting the boundaries of objects. In the proposed method, the boundaries of the object are estimated from an image sequence based on a motion stereo technique with the help of the distance measured by radar.

Since the method does not depend on the appearance of the object, it is not only capable of detecting an automobile but it can also detect other objects. Furthermore, owing to

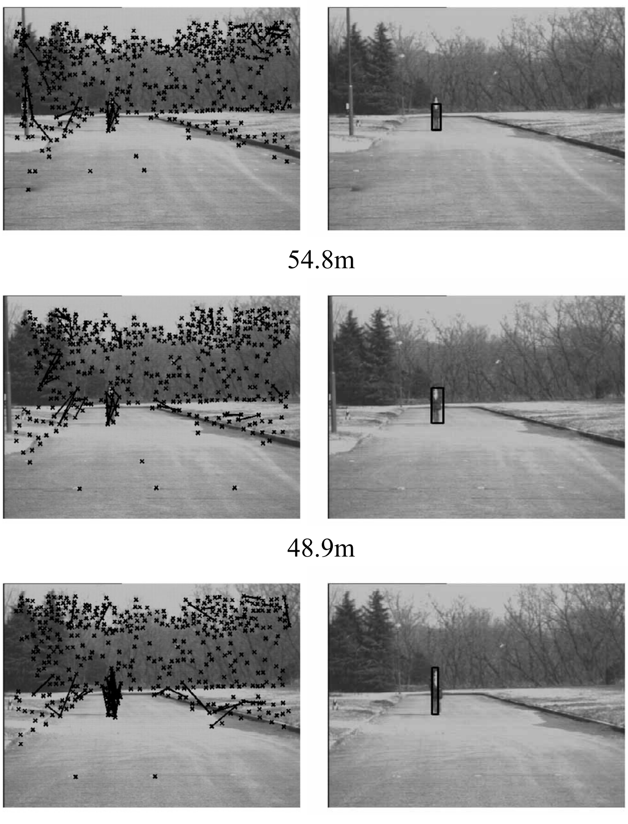


Fig. 6. Pedestrian.

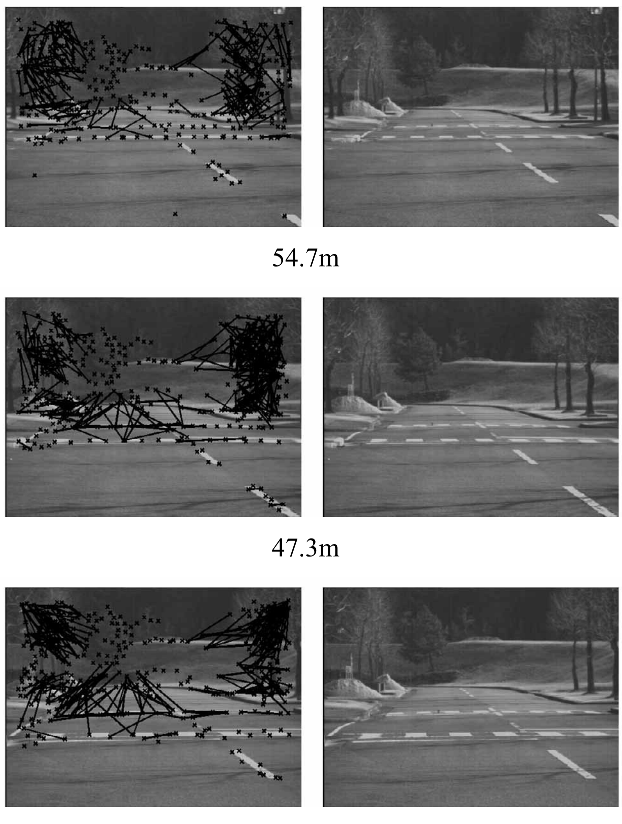


Fig. 7. Can.

the help of the distance information, moving objects as well as stationary objects can be detected by using this method.

The detection of objects by this method was confirmed through an experiment. In the experiment, both stationary and moving objects were detected, as well as a pedestrian and a vehicle. The boundaries of the objects were detected within an accuracy of 0.2m.

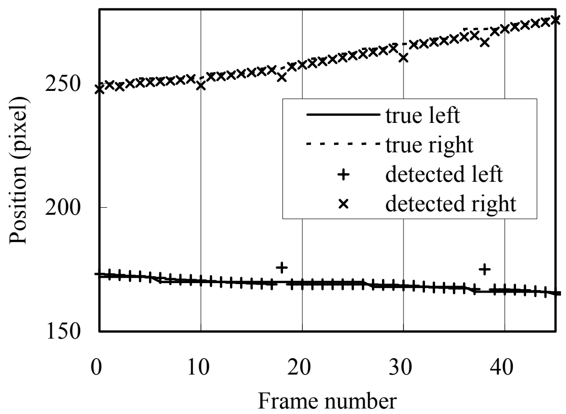


Fig. 8. Stationary vehicle.

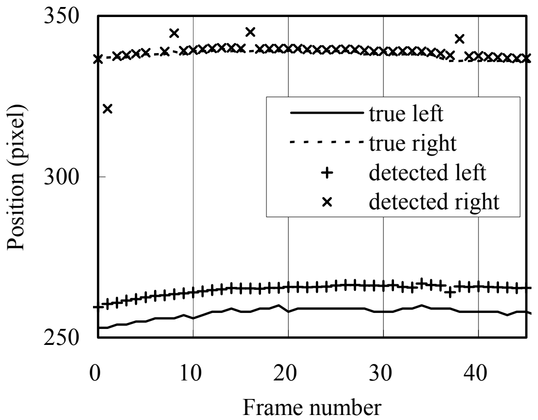


Fig. 9. Running vehicle.

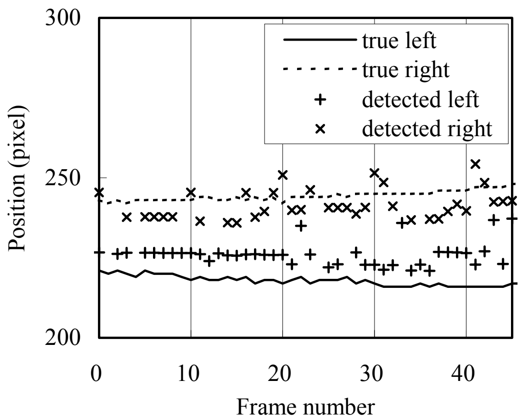


Fig. 10. Pedestrian.

References

1) S. Tokoro, “Automotive Application Systems of a Millimeter-wave Radar,” Proc. IEEE Intelligent Vehicles Symposium, pp.260-265, 1996.

2) K. Saneyoshi, “Drive assist system using stereo image recognition,” Proc. Intelligent Vehicles, pp.230-235, 1996.

3) K. C. J. Dietmayer, J. Sparbert and D. Streller, “Model Based Object Classification and Object Tracking in Traffic Scenes from Range Images,” Proc. Intelligent Vehicles 2001, pp. 25-30, 2001.

4) E. D. Dickmanns and A. Zapp, “A Curvature-based scheme for Improving Road Vehicle Guidance by Computer Vision,” Proc. SPIEE Conf. Mobile Robot, pp161-168, 1986.

5) F. Thomanek, E.D. Dickmanns and D. Dickmanns, “Multiple object recognition and scene interpretation for autonomous road vehicle guidance,” Proceedings of the IEEE Intelligent Vehicles Symposium, pp.231-236, 1994.

6) M. Betke, E. Haritaoglu and L. S. Davis, “Multiple vehicle detection and tracking in hard real-time,” Proc. IEEE Intelligent Vehicles Symposium, pp.351-356, 1996.

7) T. Kalinke, C. Tzomakas and W. v. Seelen, “A Texture-based Object Detection and an adaptive Model-based classification,” Proc. 1998 IEEE International Conference on Intelligent Vehicles, pp.143-148, 1998.

8) T. Kato, Y. Ninomiya and I. Masaki, “An Obstacle Detection Method by Fusion of Radar and Motion Stereo,” Proc. Intelligent Vehicles 2001, pp.37-42, 2001.

9) Takeo Kato, Yoshiki Ninomiya and Ichiro Masaki, "An Obstacle Detection Method by Fusion of Radar and Motion Stereo," IEEE Trans. Intelligent Transportation Systems, Vol.3, No.3, pp.182-188, Sep. 2002.

10) M. Betke, E. Haritaoglu and L. S. Davis, “Multiple vehicle detection and tracking in hard real-time,” Proc. IEEE Intelligent Vehicles 1996 Symposium, pp.351-356, 1996.

11) S. Carlsson and J-O. Eklundh, “Object detection using model based prediction and motion parallax,” Proc. ECCV90, pp. 297-306, Apr. 1990.

12) T. Suzuki and T. Kanade, “Measurement of Vehicle Motion and Orientation using Optical Flow,” Proc. IEEE Int. Conf. on ITS, pp. 25-30, Oct. 1999.

13) J. Huber and V. Graefe, “Motion Stereo for Mobile Robots,” IEEE Trans. Industrial Electronics, Vol. 41, No. 4, pp.378-383, Aug. 1994.

14) K.-D. Kuhnert, “Quantitative Motion Stereo for High Speed Car Driving,” Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 2147-2154, 1992.

15) J. Shi and C. Tomasi, “Good Features to Track,” Proc. IEEE CVPR94, pp. 593 – 600, June 1994.