Fusion of a Camera and a Laser Range Sensor for Vehicle Recognition

Shirmila Mohottala Shintaro Ono Masataka Kagesawa
Katsushi Ikeuchi
Institute of Industrial Science, University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
{shirmi,onoshin,kagesawa,ki }@cvl.iis.u-tokyo.ac.jp

Abstract

This paper presents a system that fuses data from a vision sensor and a laser sensor for detection and classification. Fusion of a vision sensor and a laser range sensor enables us to obtain 3D information of an object together with its textures, offering high reliability and robustness to outdoor conditions. To evaluate the performance of the system, it is applied to recognition of on-street parked vehicles scanned from a moving probe vehicle. The evaluation experiments show obviously successful results, with a detection rate of 100% and an accuracy over 95% in recognizing four vehicle classes.

1. Introduction

On-street parking has been recognized as one of the major causes of congestion in urban road networks in Japan. In order to determine effective strategies to minimize the problem, the road administrators need various statistics on on-street parking. Currently there are no automated systems that can scan road situations, so the surveys on road conditions are done manually by human operators. Manual counting is time-consuming and costly. Motivated by this strong requirement, we propose a system that can scan the roadside environment and extract information about on-street parked vehicles automatically. Roadside data is acquired using both a laser range sensor and a vision sensor mounted on a probe vehicle that runs along the road, and then the two sets of scanned data are fused for better results.

Vision sensors offer a number of advantages over many other sensors, and they are particularly good at wide and coarse detection. But they have limited robustness in outdoor conditions. Further, dealing with on-board cameras is much more complicated than with stationary cameras, because when the camera is moving, the environment changes significantly from scene to scene, and there is no priori knowledge of the background. On the other hand, recent laser range sensors have become simpler and more compact, and their increasing reliability has attracted more attention in various research applications. They are very efficient in extracting 3D geometric information, but one main drawback of the laser sensors is their lack of capability in retrieving textures of objects.

However, fusion of a vision sensor and a laser range sensor enables the system to obtain 3D information of the scanned objects together with their textures, offering a wide range of other possibilities. Sensor fusion provides several advantages:

- Vision-based algorithms alone are not yet powerful enough to deal with quickly changing or extreme environmental conditions.
- Multi sensors greatly increase robustness and reliability.
- Different sources of information can be used to enhance perception or to highlight the areas of interest.
- Various traffic parameters can be obtained at once.

In addition, the system will be modified in our future work for 3D town modeling to extract buildings or road signs from 3D data and to retrieve their textures as well. These capabilities are extremely important for many ITS applications such as 3D navigation systems, 3D urban modeling, and map construction.

Laser and vision sensor fusion has been widely used in applications in robotics. In [2, 10] laser range data are used to detect moving objects, and the object is extracted based on the position information. [1, 9] use laser and stereoscopic vision data for robot localization and motion planning. In [13], a sensor- fused vehicle detection system is proposed that determines the location and the orientation of vehicles using laser range information, and applies a contour-based pattern recognition algorithm.

The rest of this paper is organized as follows. In section 1.1, the system configuration and the sensors employed for

978-1-4244-3993-5/09/$25.00 ©2009 IEEE

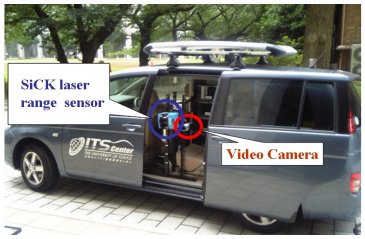
scanning are explained. Next, starting from section 2, the data processing and recognition algorithms are presented in four sections as follows, along with the experimental results of each method.

- Segmentation of vehicles from laser range data
- Calibration of laser sensor and camera
- Refinement of segmentation result using both laser and image data
- Classification of vehicles.

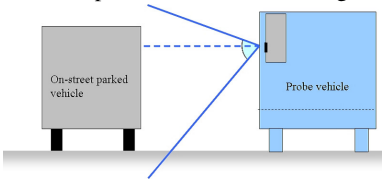
Finally we discuss merits and demerits of the whole system in section 6, and summarize in the last section.

1.1. System Configuration

The laser range sensor and the video camera are mounted on a probe vehicle, close to each other facing same direction, so that both sensors can scan the object at the same time (Figure 1 (a)). The probe vehicle runs in the next lane to the parked vehicles, and performs scans as shown in Figure 1 (b).



(a) The laser range sensor and the video camera mounted on the probe vehicle for scanning.



(b) Scanning

Figure 1. System configuration

- Laser Range Sensor
  Laser range sensors measure the time of flight or the phase shift of the laser pulse to determine the distance from the scanner to each point on the object from which the laser reflects. This depth information enables the acquisition of the 3-dimensional geometric of an object. Laser scanning is accurate and instantaneous. There are various types of laser sensors and different ways of scanning. Considering the safety of use on public roads, and the robustness required to scan from a moving vehicle, we used a SICK LMS200 laser range sensor. In order to fit our need to scan while progressing and extract the vehicle shape, we set the sensor transversely as shown in Figure 1, so that the scanning is done vertically.
- Camera
  To acquire vision data, we use a normal video camera with the resolution set to 848×480 and the frame rate to 30 frames per second. A wide view is required because the targeted vehicles need to appear in full within a frame, even when the probe vehicle is close to the parked vehicle.

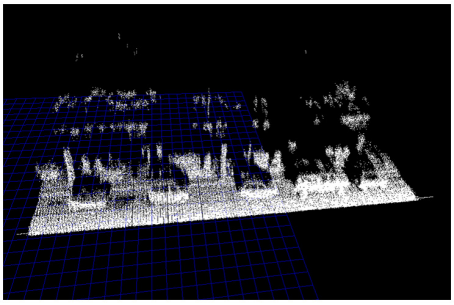


Figure 2. A depth image.

2. Segmenting Vehicles from Laser Range Data

In this section we present the method for segmenting vehicles from laser range data. The laser range sensor reports laser readings as 3D points in a laser coordinate system, for which the origin is the laser sensor. By fusing with an accurate positioning system like Global Positioning System (GPS) to acquire the position of the sensor, these points can be projected to the World Coordinate System (WCS). Assuming the speed of the probe vehicle is constant and runs straight parallel to parked vehicles, we can line up the scanned lines to display the range data in a depth image as shown in Figure 2. Here we set the WCS as follows:

- X-axis : the direction the probe vehicle progresses
- Z-axis : the direction vertical to the road surface
- Y-axis : the direction of the outer product of x and z

The laser scan plane is the plane $X = 0$. It would seem that segmenting the vehicle area from the background could be performed simply by cutting off the background on the basis of depth. But this is impossible in some cases. Considering safety, we chose a sensor with a laser that is not very powerful. Hence the reflection of laser from black vehicles is poor, and sometimes only a few points from the body are reflected. A simple detector based on depth will fail to detect

those black vehicles. Therefore, we apply a novel method, combining two separate vehicle detectors, which we have presented in our previous work [8].

2.1. Side Surface-Based Method

When scanned points are projected on to the Y-Z plane and the total number of points in the Y direction is calculated, we get a histogram as shown in Figure 3. The side surface of on-street parked vehicles appears in a significant peak. We use this feature to extract vehicles. The side surface of a vehicle A is determined as follows:

$$A = \{(x, y, z) \mid z > z_{AB}, y_a < y < y_b\}. \quad (1)$$

By experiment, we set the constants as $Z_{AB} = 0.5(m)$, $y_a = 1.0$ (m), and $y_b = y_{\text{peak}} + 1.0$ (m) and the minimum y that gives a maximum in the histogram is taken as y_{peak}. It is then smoothed with a smoothing filter and only the blocks longer than 2 meters are counted as vehicles.

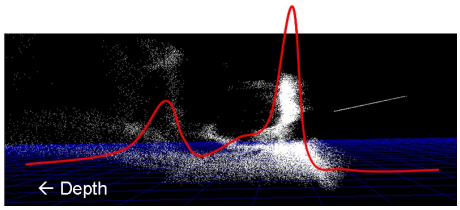


Figure 3. Histogram of the number of range points when the points are projected to the Y-Z plane

2.2. Silhouette-Based Method

The side surface-based method uses the points of vehicles' side surfaces for detection; hence, the result varies depending on the reflectance of the vehicle body. For example, black vehicles that have a low laser reflectance may not give a significant peak; therefore, they may not be detected. In the silhouette-based method. we focus on the road surface in the depth image. The on-street parked vehicles occlude the road surface, so, as shown in Figure 4, they make black shadows on the depth image. Despite of the reflectance of the vehicle body, these shadows appear because the laser does not go through the vehicles.

The silhouette-based method is as follows: First we extract the road surface by $B = \{(x, y, z) \mid z < z_{AB}\}$ where z_{AB} is set to $0.5(m)$. It is then smoothed with a smoothing filter and thresholded to detect vehicle areas. The threshold y_{th} is defined as

$$y_{\text{th}} = \bar{y}_{\text{A}} + 1.5 \text{ (m)}$$

by experiments where $\bar{y}_{\text{A}}$ is the average depth of the segmented region in the side surface-based method.

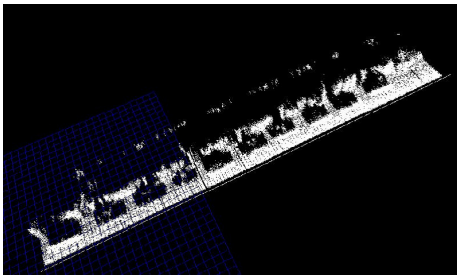


Figure 4. Depth image of road surface

2.3. Detection Results

The on-street parked vehicles are segmented with two detectors explained above, and the results are combined. In experiments all the vehicles were detected 100% accurately. Figure 5 shows an example of segmented results.

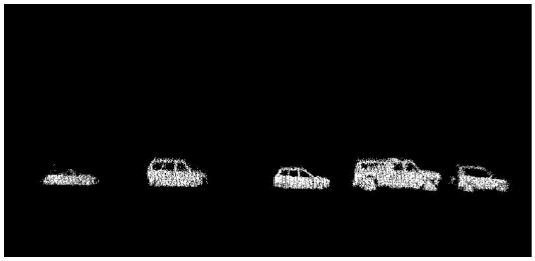


Figure 5. Vehicles detected from laser range data

3. Extrinsic Calibration of Camera and Laser Range Sensor

In a sensor-fused system, the process to determine how to combine the data from each sensor is complicated. There are various sensor fusing methods and calibration methods proposed to use the sensors to their fullest potential. Different approaches have been proposed to calibrate cameras and laser range sensors, and most of them use markers or features like edges and corners that can be visible from both sensors [6][12]. Some methods use visible laser [4]. Zhang and Pless present a method that makes use of a checkerboard calibration target seen from different unknown orientations [15], and a similar approach is proposed in [11]. Our calibration is based on the method in [15], and only the extrinsic calibration method is presented here, assuming that the intrinsic parameters of the camera are known.

Extrinsic calibration of a camera and a laser range sensor consists of finding the rotation Φ and translation Δ between the camera coordinate system and the laser coordinate system. This can be described by

$$P^l = \Phi P^c + \Delta \quad (2)$$

where P^c is a point in the camera coordinate system that is located at a point P^l in the laser coordinate system. Φ and Δ are calculated as follows.

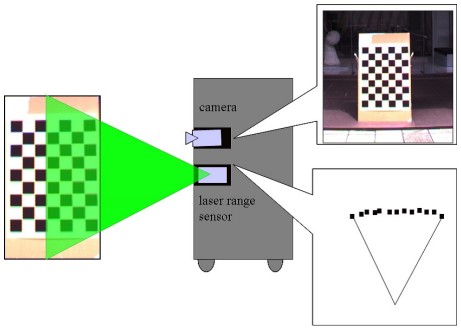


Figure 6. The setup for calibration of camera and laser range sensor

A planar pattern on a checkerboard setup as shown in Figure 6 is used for calibration. We refer to this plane as the "calibration plane." The calibration plane should be visible from both camera and laser sensor. Assuming that the calibration plane is the plane $Z = 0$ in the WCS, it can be parameterized in the camera coordinate system by 3-vector N that is vertical to the calibration plane and $\| N \|$ is the distance from the camera to the calibration plane.

From Equation 2, the camera coordinate P^c can be derived as $P^c = \Phi^{-1}(P^l - \Delta)$ for a given laser point P^l in the laser coordinate system. Since the point P^c is on the calibration plane defined by N, it satisfies $N \cdot P^c = \| N \|^2$.

$$N \cdot \Phi^{-1}(P^l - \Delta) = \| N \|^2 \tag{3}$$

For a measured calibration plane N and laser points P^l, we get a constraint on Φ and Δ.

We extract image data and laser range data in different poses of the checkerboard. Using these data, the camera extrinsic parameters Φ and Δ can refined by minimizing

$$\sum_i \sum_j \left(\frac{N_i}{\| N_i \|} \cdot (\Phi^{-1}(P^l_{ij} - \Delta)) - \| N_i \|\right)^2 \tag{4}$$

where N_i defines the checkerboard plane in the ith pose. Minimizing is done as a nonlinear optimization problem with the Levenberg-Marquardt algorithm in Matlab.

3.1. Calibration Results

Using the above method we calculate the projection matrix of the laser coordinate system to the image coordinate system. In section 2, the vehicles are segmented from their background in the laser range data. Now we can project these laser range data onto the corresponding images. If the sensor and camera are calibrated correctly, the laser scanned points will be projected onto the vehicle area in images. Figure 7 shows an example where one laser scanned line of a vehicle is projected onto the corresponding image.

Figure 7. The points in one laser-scanned line are projected onto the corresponding image. Laser points are marked in green.

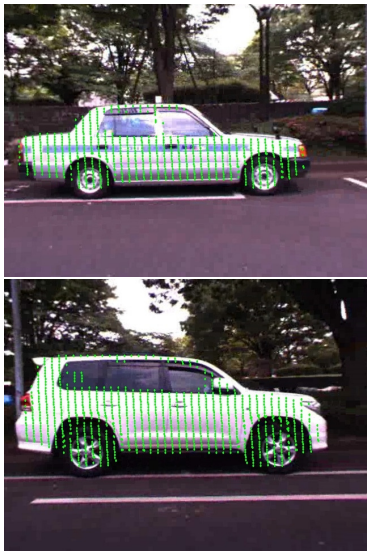


Figure 8. Some examples of laser-scanned lines projected onto the corresponding images. Scanned points are marked in green

If we use an accurate positioning system like GPS along with an INS (Internal Navigation System) to acquire the position of the probe vehicle instantaneously, we can line up all the scanned lines accurately based on the sensor position. This kind of accurate positioning is very useful for applications like 3D town digitalizing. But for our task, a more simple method based on a simple assumption will derive sufficiently accurate results.

Here we assume that the probe vehicle moves straight forward at a constant speed, parallel to on-street parked vehicles. The laser and image data are synchronized manually, by setting the two ends of the data sequences. Next, on the basis of the above assumption, we project all the scanned lines onto the corresponding images, and line them up in constant intervals. Figure 8 presents two examples where the laser-scanned lines are projected onto the vehicles, and

the silhouettes of vehicles appear quite accurately. Even though there is a small error, it does not significantly affect our task.

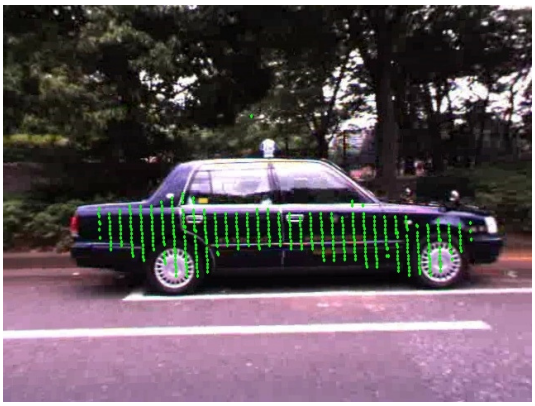


Figure 9. An example of a black vehicle that is not scanned accurately by laser.

4. Refinement of Segmentation

We have already segmented vehicles from laser range data and fused the result with image data. By linking all the maximum points and the minimum points of each line projected onto the images, we can segment the vehicles from images as well. But Figure 9 shows an example in which this will not work. Because the laser sensor we used is not very powerful, some laser points are not reflected well. This was particularly true for black vehicles, in which the total number of scanned points was less than half the number for white vehicles. Therefore, we employ a vision-based segmentation method as well to refine these segmentation results and enhance the robustness.

4.1. Graph Cut Method

Graph-cut method is a powerful optimizing technique that can achieve robust segmentation even when foreground and background color distributions are not well separated. Greig et al. [3] were the first who proposed the ability to apply the graph-based powerful min-cut/max-flow algorithms from combinatorial optimization to minimize certain important energy functions in vision.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph with vertices $\mathcal{V}$, the set of elements to be segmented, and edges $e \in \mathcal{E}$ corresponding to pairs of neighboring vertices. Each edge e is assigned a nonnegative weight (cost) w_e. In the case of image segmentation, the elements in V are pixels, and the weight of an edge is some measure of the dissimilarity between the pixels connected by that edge. There are two special terminals: an "object" and a "background." The source is connected by edges to all nodes identified as object seeds, and the sink is connected to all background seeds. A cut is a subset of edges, which has a cost that is defined as the sum of the weights of the edges that it severs.

$$|C| = \sum_{e \in C} w_e$$

The segmentation boundary between the object and the background is drawn by finding the minimum costs cut on the graph.

4.2. Results of Segmentation Using Laser Range Data and Graph Cut Method

We apply the graph cut method to refine the segmentation results obtained earlier by only projecting segmented laser data. The graph cut method requires a user-initiated indication of background/foreground areas. Foreground is denoted by the earlier segmentation result, based on the points segmented from laser data and projected onto the images. Note that the points closer to the outline on the top and the bottom of a scanned line are ignored, considering the calibration error. Then two lines are drawn on the top and the bottom of the image in red, indicating the background area. These two lines are fixed for all the images and drawn horizontally at $y = 50$ and $y = height - 50$. Figure 10 shows some examples of segmentation results. The background is painted over in purple. Despite the very few number of points extracted from the black vehicle, this method enabled accurate segmentation of the vehicle area.

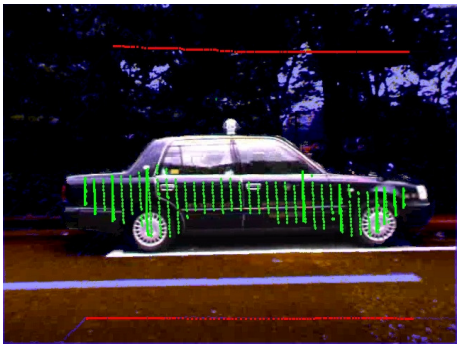


Figure 10. Some examples of vehicle segmentation using laser range data and graph cut method.

5. Classification of On-street Parked Vehicles

This section describes how the system recognizes the classes (categories) of the vehicles segmented above. The classification algorithm is based on our previous work [14], which classified vehicles from an overhead surveillance camera. Compared to the top view, side views of vehicles give more distinguishing features for classification, so the

task seems to be easy. But one main drawback that interferes with the classification process is the reflectance from the vehicle body surface. Important features of some vehicles are barely visible due to reflection of sunlight, while some vehicles appear with very strong noise due to reflection of the surroundings. Moreover, in some vehicles, accessories such as the shape of the doors, lights, mirrors or wheel caps appear with unique features, changing the appearance of the whole vehicle. To deal with these characteristics, we apply some modifications to the previous method, and evaluate the system for recognition of four very similar vehicle classes: sedan, minivan, hatchback, and wagon.

5.1. Classification Method

The classification algorithm is based on the Binary Features method [5], which can be introduced as a modified step of eigen window technique [7]. In this method, Principal Component Analysis (PCA) that uses real valued eigenvectors is replaced with the vector quantization method, to avoid floating point computations and reduce the required computation time and memory. Further, binary edge images are used instead of gray level images, increasing robustness to illumination changes. First, the binary edge images are calculated from the original image, using a Laplacian and Gaussian filter. Then the binary features are selected as described below.

Let $B(e_i(x,y);x_0,y_0;b)$ be a window of size $(2b+1)\times(2b+1)$ pixels around (x_0,y_0) in a binary edge image $e_i(x,y)$. Each window is rated as follows,

$$r_i(x,y) = \min_{\substack{-d\le d_x\le d\\ -d\le d_y\le d}} \Big\{ D_H\Big[\mathbf{\Omega}\{e_i(x',y');x,y,b\}, \mathbf{\Omega}\{e_i(x',y');x+d_x,y+d_y,b\}\Big]\Big\} \quad (5)$$
$$(d_x,d_y)\neq(0,0).$$

where $D_H(\bar{A},\bar{B})$ is the Hamming distance between two binary vectors $\bar{A}$ and $\bar{B}$. The window will rate highly if it is dissimilar to its surrounding. Highly rated windows are taken as features, and then compressed to code features using standard Lloyd's algorithm, based on vector quantization. Code features are computed based on the nearest neighbor clusters of training features. For each coded feature, a record of the group of features it represents and the locations of these features on training images is kept.

To detect an object in an input image, first, the input image is processed the same way as the training images to get a binary edge image. The binary edge input image is then encoded, or in other words, the nearest code corresponding to each pixel $c(x,y)$ is searched using

$$c(x,y) = \arg\min_{a\in[1,2,\ldots,n_c]} \Big\{ D_H\Big[\bar{\Omega}\{e(x',y');x,y,b\},\bar{F}_a\Big]\Big\} \quad (6)$$

where $\bar{F}_a$ denotes the codes computed using Lloyd's algorithm. The encoded input image is then matched with all the training images to find the nearest object through a voting process.

5.1.1 Voting Process

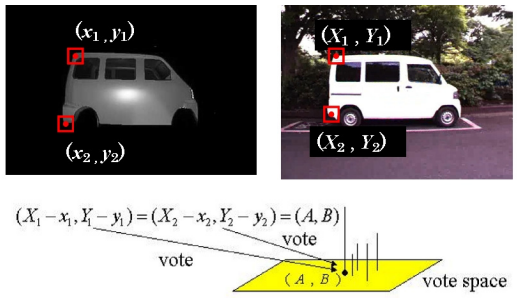
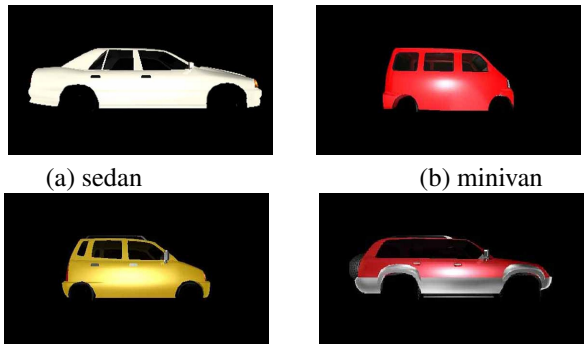


Figure 11. Voting process

We prepare a voting space $V(T)$ for each training image T. Let $w(J,X_i,Y_i)$ be a feature at (X_i,Y_i) on the input image J. First we find the nearest code feature for w, and the corresponding features $w_n(T_k,x_i',y_i')$ represented by that code. w_n is a feature at (x_i',y_i') on the training image T_k. Once the best matched feature is found, we put a vote onto the point (X_i-x_i',Y_i-y_i') of the vote space $V(T_k)$ (Figure 11). Note that one vote will go to the corresponding points of each feature represented by that code.
If the object at (X,Y) in the input image is similar to the object in the training image T, a peak will appear on $V(T)$ at the same location (X,Y). The object can be detected by thresholding the votes.

As our input images are not in a high resolution, the votes may not stack up exactly on one point, but they will concentrate in one area. Therefore, we apply a Gaussian filter to the voting space, improving the robustness of the object scale.



(a) sedan (b) minivan

(c) hatchback (d)wagon

Figure 12. Some of the CG training images used for classification of on-street parked vehicles

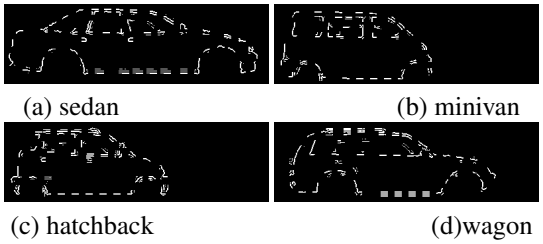


Figure 13. Features extracted from CG training images for classification of on-street parked vehicles

5.2. Classification Results

The system is evaluated through experiments using a set of images with 37 sedans, 28 minivans, 31 hatchbacks, and 43 wagons, a total of 139 images. Classification results are denoted in a confusion matrix in Table 1. Some examples of successfully classified images are presented in Figure 14. The input edge image is shown in blue and the training feature image with the highest match is overlapped and drawn in red. The total classification ratio of the system is 96%. Figure 15 shows an example of the robustness of this method to partial occlusion.

Classified as	Real Class			
	Sedan	Minivan	Hatchback	Wagon
Sedan	37	0	0	1
Minivan	0	25	0	0
Hatchback	0	3	31	1
Wagon	0	0	0	41
Accuracy	100%	89%	100%	95%

Table 1. Classification results of on-street parked vehicles when a few recent minivan models are eliminated from the input image set. Accuracy: 96%

6. Discussion

We presented a recognition system that fuses a laser range sensor and a camera to detect and classify on-street parked vehicles. The processing algorithms were explained in four sections, together with the experimental results. Here we discuss the characteristics and merits of each algorithm, following with issues and future work.

The method to segment the vehicles from laser range data worked very well, detecting the vehicles 100% accurately. The calibration of the two sensors was quite successful, as we can see in experimental results. When projecting laser data onto the images, the sensor position is calculated assuming the probe vehicle runs at a constant speed. But in reality, the speed of the probe vehicle at each point is not constant, and this led to a small error in synchronizing, even though it did not influence the performances of the system.

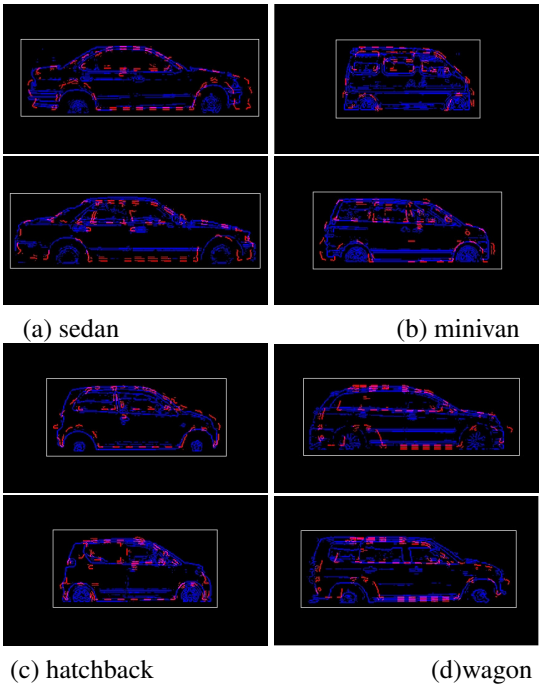


Figure 14. Some examples that are classified successfully. Input vehicle is shown in blue and the training vehicle best matched is indicated in red.

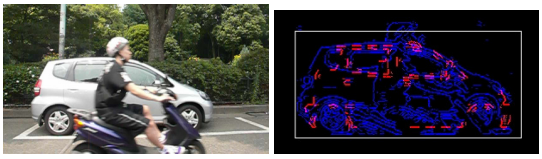


Figure 15. A successful classification of a partially occluded vehicle

In our future work, we plan to fuse a positioning system to acquire the real position of the probe vehicle, so that we will be able to derive more precise results. The graph cut method modified by fusing laser-segmented results to initialize the foreground/ background area could segment the vehicle area from its background accurately. Even the black vehicles where the laser range data are not reflected well were segmented correctly. We believe that our method is robust enough to extract road signs or textures of buildings automatically, which we hope to achieve in our future work.

Classification of on-street parked vehicles showed very good results with an accuracy of 96%. Sedans and hatchbacks achieved a significant classification rate with no failures, while wagons showed a high accuracy of 95%. The classification rate was poor only in minivans. A box-shaped body, larger than sedans or wagons in height and with three rows of seats were some features of the class “Minivan” we set. The real class of each vehicle model was determined

(a) a conventional minivan

(b) recent minivans

Figure 16. How the model of recent minivans differs from a conventional model. Top: a conventional minivan with box-shaped body. Bottom: recent models of minivans

according to automobile manufacturers. But the shape of the recent minivans have changed considerably; they are not box-shaped any more. Their front ends look similar to wagons, and height does not differ much from wagons. Figure 16 shows an example of a conventional minivan and some recent models of minivans that appeared in our experimental data. In future work, we hope to reconsider the basis of these classes to fit today's vehicle models.

7. Conclusion

We presented a recognition system that fuses a laser range sensor and a camera for scanning, and we used the resulting types of information to detect and recognize the target objects. The system was evaluated on its success in detecting and classifying on-street parked vehicles, and it achieved a total classification accuracy of 96%. The sensor-fused system we proposed here can model the 3D geometric of objects and segment their texture from the background automatically, so that the texture can be mapped onto the 3D model of the object. In our future work, we hope to fuse this system with a positioning system like GPS and INS and apply this technique for digitalizing and 3D modeling of urban environments, as a part of our ongoing project.

References

[1] H. Baltzkis, A. Argyros, and P. Trahanias. Fusion of laser and visual data for robot motion planning and collision avoidance. *Machine Vision and Applications*, 15:92–100, 2003.

[2] J. Blanco, W. Burgard, R. Sanz, and J. L. Fernandez. Fast face detection for mobile robots by integrating laser range data with vision. *in Proc. IEEE Int. Conf. Robotics and Automation*, pages 625–631, 2003.

[3] D. Greig, B. Porteous, and A. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society*, 51(2):271–279, 1989.

[4] O. Jokinen. Self-calibration of a light striping system by matching multi-ple 3-d profile maps. *in Proc. the 2nd International Conference on 3D Digital Imaging and Modeling*, pages 180–190, 1999.

[5] J. Krumm. Object detection with vector quantized binary features. *in Proc. of Computer Vision and Pattern Recognition , San Juan, Puerto Rico*, pages 179–185, 1997.

[6] C. Mei and P. Rives. Calibration between a central catadioptric camera and a laser range finder for robotic applications. *in Proceedings of ICRA06, Orlando, May*, 2006.

[7] K. Ohba and K. Ikeuchi. Detectabillity, uniquness, and reliability of eigen windows for stable verification of partially occluded objects. *IEEE Trans. on Pattern Anarysis and Machine Intelligence*, 19(9):1043–1048, 1997.

[8] S. Ono, M. Kagesawa, and K. Ikeuchi. Recognizing vehicles in a panoramic range image. *Meeting on Image Recognition and Understanding (MIRU)*, pages 183–188, 2002.

[9] S. Pagnottelli, S. Taraglio, P. Valigi, and A. Zanela. Visual and laser sensory data fusion for outdoor robot localisation and navigation. *in Proc. 12th International Conference on Advanced Robotics, July 18-20, 2005*, pages 171–177, 2005.

[10] X. Song, J. Cui, H. Zhao, and H. Zha. Bayesian fusion of laser and vision for multiple people detection and tracking. *in Proc. SICE Annual Conf. 2008 (SICE08)*, pages 3014–3019, 2008.

[11] R. Unnikrishnan and M. Hebert. Fast extrinsic calibration of a laser rangefinder to a camera. technical report. *CMU-RI-TR-05-09, Robotics Institute, Carnegie Mellon University, July*, 2005.

[12] S. Wasielewski and O. Strauss. Calibration of a multi-sensor system laser rangefinder/camera. *Proc. of the Intelligent Vehicles '95 Symposium*, pages 472–477, 1995.

[13] S. Wender, S. Clemen, N. Kaempchen, and K. C. J. Dietmayer. Vehicle detection with three dimensional object models. *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, Heidelberg, Germany, September*, 2006.

[14] T. Yoshida, S. Mohottala, M. Kagesawa, T. Tomonaka, and K. Ikeuchi. Vehicle classification system with local-feature based algorithm using cg model images. *The IEICE Transactions on Information and Systems*, (11):1745–1752, 2002.

[15] Q. Zhang and R. Pless. Extrinsic calibration of a camera and laser range finder (improves camera calibration). *Intelligent Robots and Systems (IROS)*, 2004.