2011 IEEE Intelligent Vehicles Symposium (IV)
Baden-Baden, Germany, June 5-9, 2011

3-D Line Segment Reconstruction using an In-Vehicle Camera for Free Space Detection

Hiroyuki Uchiyama, Daisuke Deguchi, Tomokazu Takahashi, Ichiro Ide and Hiroshi Murase
Graduate School of Information Science, Nagoya University
Furo-cho Chikusa-ku Nagoya, Aichi 464–8601, Japan
uchiyama@murase.m.is.nagoya-u.ac.jp

***Abstract*— Free space detection is very important for vehicle navigation and safe driving. 3-D line segment reconstruction of a street is important for the free space detection because a street-view includes many line segments. For the free space detection, we propose a method for reconstructing 3-D line segments in a streetscape using a monocular in-vehicle camera. The 3-D reconstruction of the line segments is achieved by using each three images from an image sequence. Once accurate camera poses of these images are obtained, one of the remaining crucial problems is to match the line segments between the images correctly. A strategy for finding correspondence of the line segments is as follows: First, the correspondences of line segment candidates are searched by using a two-view constraint. However, the two-view constraint has difficulty on determining an unique correspondence geometrically. Therefore, the candidates of the line segment correspondences are reduced using a three-view constraint. In order to improve the accuracy, the proposed method exploits a color feature of the line segment and a preliminary knowledge of the vehicle motion. Finally, the line segments are reconstructed using the correspondences. From an experimental result, we confirmed the effectiveness of the proposed method. Application to the free space detection demonstrated the usefulness of the reconstructed line segments.**

I. Introduction

Detection of a free space [1], [2] is very important for vehicle navigation and safety driving. Here, the free space is a region where a vehicle can pass through, such as a road region. Collision detection and lane keeping can be achieved by using results of the free space detection. Therefore, several methods for free space detection have been proposed. Most of them were based on the idea that the free space is an area without obstacle regions [2]. Therefore, detecting obstacles is important for the free space detection. An example of a street-view in an urban area is shown in Fig. 1. Straight line segment is one of the most important structures for the free space detection. For example, walls, poles, trees and guardrails are mostly composed of line segments. Especially, the vertical line segments are contained in all of the objects. Therefore, estimating the location of the vertical line segments plays an important role in the free space detection.

Fig. 1. Example of a street-view in an urban area. Most of the objects are composed of straight line segments.

There are several methods for estimating structure locations in a streetscape. For example, a laser scanner can be used for reconstructing a 3-D model [3] of the structures. However, it is expensive and difficult to mount on a passenger vehicle. On the other hand, an in-vehicle camera can be used for estimating the structure locations [5], and also for other purposes such as traffic sign recognition or fog density estimation. In addition, the in-vehicle camera has the advantage that it is inexpensive and compact. Several methods use stereo vision with two cameras to reconstruct a 3-D structure [4]. Structure from Motion (SfM) by a monocular vision is another approach for the reconstruction of the 3-D structure [5].

Several methods use point features for reconstructing 3-D structures [6], [7]. However, a thin pole, which is important for collision detection and free space detection, is difficult to detect by the point features due to the lack of textures and corner points around them. On the other hand, there are several methods to reconstruct a line segment from line correspondence [8], [9]. These methods can detect the edge of thin poles. Asai et al. [10] proposed an off-line method to reconstruct a 3-D line segment based on bundle adjustment using a long image sequence. However, the bundle adjustment requires a large computational cost.

In this paper, we propose a method for reconstructing 3-D line segments in a streetscape using a monocular in-vehicle camera aiming for the free space detection. The vertical line segments are reconstructed based on a SfM approach. Matching line segments between images is important for accurate reconstruction. Therefore, to improve the matching accuracy with a small number of images, the method uses multiple-view constraints, the property of line segments, and the preliminary knowledge of the vehicle's motion. In an experiment, the performance of reconstructing 3-D line segments is demonstrated by applying it to the free space detection.

In the following, Section II describes the 3-D reconstruction method of line segments. Section III shows the experimental result and its application to the free space detection. Finally, Section IV gives the conclusion.

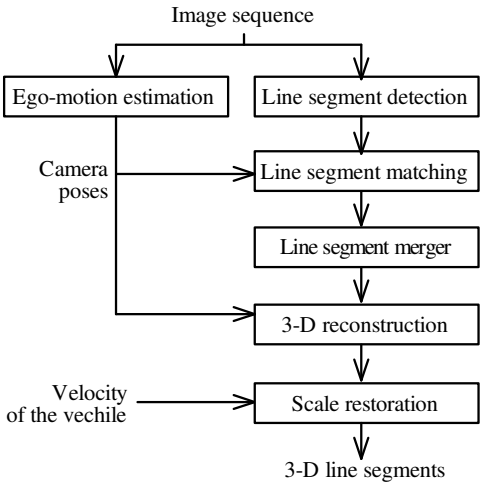


Fig. 2. The flow of the 3-D line reconstruction.

II. 3-D Reconstruction of Line Segments

A. Overview of the Method

This paper uses "line" and "line segment" in different meanings. Here, a line is defined as a straight and thin object with an infinite length. In contrast, a line segment has a finite length with two endpoints.

The 3-D reconstruction of a line segment is achieved by using three images from an image sequence. Once the accurate camera poses of these images are obtained, one of the remaining crucial problems is to match the line segments between the images correctly. Therefore, this paper focuses on the problem of matching line segments. A strategy for finding correspondence of the line segments is as follows: First, the correspondences of line segment candidates are searched by using the two-view constraint. However, by the two-view constraint, it is difficult to determine a unique correspondence geometrically. Therefore, the candidates for the line segment correspondence are reduced using the three-view constraint. In order to improve the accuracy, the proposed method exploits a color feature of the line segment and the preliminary knowledge of the camera motion.

Our method consists of five parts: (1) ego-motion estimation, (2) detection of line segment, (3) matching of the line segments, (4) merger of the line segments, and (5) 3-D reconstruction. The flow of the method is shown in Fig. 2. In the ego-motion estimation part, camera poses are estimated from the image sequence. A conventional method [11] is used for estimating the ego-motion. As the result of the ego-motion estimation, the camera matrices and the fundamental matrices are obtained. In the next part, line segment detection and matching are performed by using the camera poses. However, some line segments may not be connected. Therefore, several line segments are combined in the line segment merger part. In the 3-D reconstruction part, the line segment structure is restored from the camera poses and the results of the line segment matching.

In this paper, images captured at time T_1, T_2 and T_3 ($T_1 < T_2 < T_3$) are denoted as Images 1, 2, and 3, respectively.

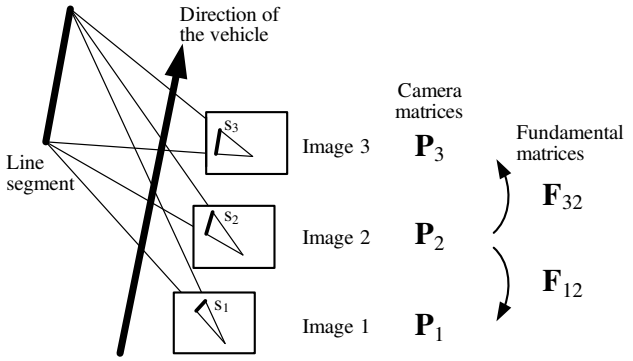
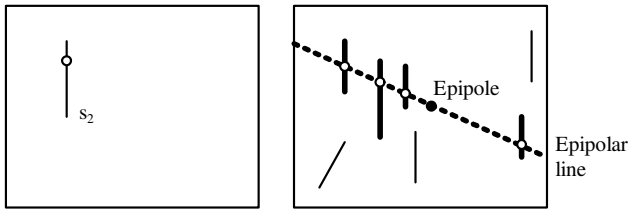


Fig. 3. Geometric parameters for three cameras.



(a) Image 2 (b) Image 1 or 3

Fig. 4. Line segment matching using two views. The line segments intersecting with the epipolar line are the candidates for the correspondence.

Also, the triplet of the corresponding line segments in each image is denoted as $[\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3]$ as shown in Fig. 3. The camera matrices for Images 1, 2, and 3 are denoted as $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$. Here, the camera matrix is a 3×4 matrix representing the camera pose. Matrices $\mathbf{F}_{12}$ and $\mathbf{F}_{32}$ are the fundamental matrices between Image 2 and the other images.

The following sections describe the details of the line segment detection, matching, merger, and the 3-D reconstruction.

B. Detection of Line Segments

In order to the extract line segments from the images, the proposed method uses LSD (Line Segment Detector) [14], which is a gradient-based detector. The LSD can detect the width in addition to the coordinates of the endpoints of a line segment.

C. Matching Line Segments by a Two-camera Constraint

The goal of matching line segments is acquiring a corresponding triplet of the line segments $[\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3]$ from Images 1, 2, and 3. To do this, the proposed method searches the line segments s_1 and s_3 corresponding to a fixed s_2. First, the candidates of corresponding pairs $[\mathrm{s}_1, \mathrm{s}_2]$ and $[\mathrm{s}_2, \mathrm{s}_3]$ for a fixed s_2 are found by using two-view epipolar constraints and preliminary knowledge about the camera motion.

1) Search by Epipolar Constraints: Given a point $\mathbf{x}$ on a line segment s_2 and the fundamental matrix $\mathbf{F}_{12}$, the corresponding point in Image 1 is located on an epipolar line $\mathbf{r}'$ [11] as follows.

$$\mathbf{r}' = \mathbf{F}_{12}\mathbf{x} \tag{1}$$

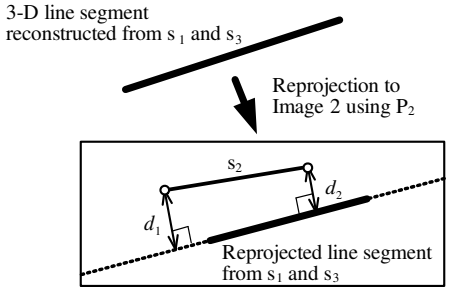


Fig. 5. Computation of the geometric distance between line segments. The distance is measured by the reprojection error.

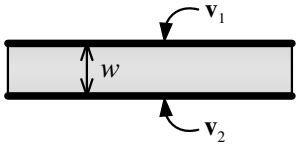


Fig. 6. Color feature of a line segment. The color feature vector $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2]$ is the set of the pixels along the line located at both sides of the line segment region.

As shown in Fig. 4, candidates of the corresponding line segments are found as the line segments intersecting the epipolar line. In Fig. 4(a), the dashed line indicates the epipolar line $\mathbf{r}'$ corresponding to an arbitrary point $\mathbf{x}$ on s_2, while the thick lines indicate the line segments of corresponding candidates. All the points on line segment s_2 are considered to search the candidates. Candidates on Image 3 are also searched in the same way.

2) Search by Camera Motion: To reduce the incorrect candidates, the proposed method uses the preliminary knowledge about the camera motion. It can be assumed that the vehicle moves forward. Furthermore, an epipole on an image indicates the direction of the camera motion. Therefore, when driving a vehicle, line segments should move away from the epipole gradually. The proposed method selects the combinations of the triplet $[\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3]$ satisfying the correct order of distances between the epipole and the centroid of the line segment.

D. Matching Line Segments by a Three-camera Constraint

After matching line segments in two views, three-view geometry is applied for matching. Combination of two kinds of distances is used for this matching:

- Geometric distance d_g
- Color distance d_c

Since there are many vertical line segments in an urban streetscape, several incorrect triplets may satisfy the three-view constraint. Therefore, they are distinguished by using the color distance. If both distances d_g and d_c are less than a threshold θ_g and θ_c, respectively, the triplet of the line segments $[\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3]$ is accepted.

1) Geometric Distance: The reprojection error of the line segment is used for the geometric distance. First, the line segment is reconstructed from s_1 and s_3. Next, the reconstructed line is reprojected to Image 2 using the camera matrix $\mathbf{P}_2$. Then, the geometric distance between the reprojected line and the line segment s_2 is computed as

$$d_\mathrm{g} = d_1 + d_2, \tag{2}$$

where d_1 and d_2 are the Euclidean distances between the reprojected line and the endpoints of s_2 as shown in Fig. 5.

The reconstructed line is obtained as the intersection of the planes π_1 and π_3, where π_1 and π_3 are the planes defined by the camera center and the line segment of Images 1 and 3, respectively, as follows:

$$\pi_1 = \mathbf{P}_1^\mathrm{T}\mathbf{r}_1, \tag{3}$$

$$\pi_3 = \mathbf{P}_3^\mathrm{T}\mathbf{r}_3, \tag{4}$$

where, $\mathbf{r}_1$ and $\mathbf{r}_3$ are the homogeneous representation of lines of s_1 and s_3.

2) Color Distance: To compute the feature of the line segment, the proposed method uses pixel values around the line segment. However, a line segment appears on a junction of the planes or a border between background and foreground. The appearance of the region between background and foreground changes in color due to the point-of-view and an image blur. Therefore, those pixels are avoided for obtaining the color feature.

To compare pixels of the line segments, F-guided matching [12] and the property of the line segment are used. In the F-guided matching, the fundamental matrices $\mathbf{F}_{12}$ and $\mathbf{F}_{32}$ are used for pixel-wise matching. As shown in Fig. 4, the pixel correspondence between line segments can be obtained using $\mathbf{F}_{12}$ and $\mathbf{F}_{32}$. From the property of a line segment, at least one side of the line segments belongs to the same foreground object. Therefore, either side of the line segment is selected when the color distance is computed.

To compute the color distance, a color feature vector $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2]$ of s_2 is obtained. Here, it contains color pixels along both sides of the line segment s_2 (Fig. 6). The width of the line segment w is obtained from the LSD result. Using the F-guided matching, color feature vectors $\mathbf{v}' = [\mathbf{v}'_1, \mathbf{v}'_2]$ and $\mathbf{v}'' = [\mathbf{v}''_1, \mathbf{v}''_2]$ of s_1 and s_3 are obtained. The color distance for $\mathbf{v}$, $\mathbf{v}'$ and $\mathbf{v}''$ is computed as

$$\begin{aligned} d_\mathrm{c}(\mathbf{v}, \mathbf{v}', \mathbf{v}'') = \min\big\{ &|\mathbf{v}_1 - \mathbf{v}'_1|_{\mathrm{L1}} + |\mathbf{v}_1 - \mathbf{v}''_1|_{\mathrm{L1}}, \\ &|\mathbf{v}_2 - \mathbf{v}'_2|_{\mathrm{L1}} + |\mathbf{v}_2 - \mathbf{v}''_2|_{\mathrm{L1}} \big\}, \end{aligned} \tag{5}$$

where, $|\cdot|_{\mathrm{L1}}$ indicates the L1 norm. The value of d_c is normalized by scaling the range to $0 \le d_\mathrm{c} \le 1$.

E. Line Segment Merger

Some line segments may not be connected due to the color similarity between the background and the foreground. Therefore, several line segments are merged using the matching result. A clustering approach is used for the merger.

Several correspondences of triplets $[\mathrm{s}_1^{(1)}, \mathrm{s}_2^{(1)}, \mathrm{s}_3^{(1)}], \ldots, [\mathrm{s}_1^{(N)}, \mathrm{s}_2^{(N)}, \mathrm{s}_3^{(N)}]$ are obtained from the matching result. In the case that one $\mathrm{s}_1^{(n)}$ is assigned to multiple $\mathrm{s}_2^{(n)}$, those $\mathrm{s}_2^{(n)}$ are considered to belong to the same line segment. Thus, if

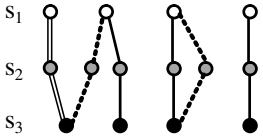


Fig. 7. Clustering for merging line segments. The nodes indicate the line segments in the image, and the edges indicate the correspondence of triplets. In this example, the line segments are combined into three clusters.

a pair of triplets shares a common line segment, the pair is considered to belong to the same cluster. For example, six triplets in Fig. 7 are merged into three clusters. Using the clustering result, the line segments within the same cluster are merged to one line segment.

F. 3-D Reconstruction

From a corresponding triplet $[\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3]$, the line segment is reconstructed. The intersection of the planes defined by the line segment on the image planes and the camera center forms a line segment in the space. In the case that the angle between the planes is small, the estimation accuracy decreases. Therefore, if the angle between planes π_1 and π_3 is less than a threshold θ_a, the line segment is rejected. Since the Structure from Motion approach can reconstruct without scales, the scale parameter is determined by the velocity.

Finally, horizontal line segments are removed. This is because vertical line segments, such as walls and poles, are useful for the free space detection.

III. Experiments and Discussion

To confirm the performance of the 3-D reconstruction method, we conducted an experiment using the in-vehicle camera sequences captured along the street. We evaluated the accuracy of the reconstruction and the line segment matching. Finally, the 3-D reconstruction is applied to the simple free space detection method to demonstrate its performance.

A. Experimental Conditions

To evaluate the line segment matching, the number of matching failures was evaluated. Then, the proposed method was compared with the following three methods: Comparative method 1 did not use the geometric distance. Comparative method 2 did not use the color distance. Comparative method 3 extracted the color feature from the region on a line segment, while the proposed method extracted the color feature from the region around a line segment. The thresholds θ_c and θ_g of the comparative methods 1 and 2 were the same as that of the proposed method, respectively. The threshold for the color feature of the comparative method 3 was adjusted so that approximately the same number of line segments could be extracted.

Ego-motion was estimated using SIFT [13]. From the correspondence of the feature points, the fundamental and camera matrices were obtained.

The images were captured by a HD camera for consumer use. The camera was mounted on the windshield of a vehicle. The original images were $1{,}920 \times 1{,}080$ pixels and captured at 29.97 fps. The images were converted to 960×540 pixels and 6.0 fps for the experiment. The velocity of the vehicle was obtained from a GPS sensor. Successive three frames were chosen for applying the proposed method. We chose ten scenes for the evaluation. The parameters used in the experiment were determined through a preliminary experiment, and $\theta_g = 3.0$ pixels, $\theta_c = 0.06$ and $\theta_a = 0.3$ deg. were used. The intrinsic camera parameters were obtained by a preliminary calibration using [15], and the distortion of the camera images were compensated.

TABLE I
Number of matched line segments in 10 scenes

Method	Comp. 1	Comp. 2	Comp. 3	Proposed
# of matching	503	705	354	379
# of failure	62	89	6	0
Failure rate [%]	12.3	12.6	1.7	0.0

B. 3-D Reconstruction Result

Examples of the results are shown in Fig. 8. Fig. 8 (a) shows the LSD result. 513, 505 and 534 line segments were detected in Images 1, 2, and 3, respectively. Fig. 8 (b) shows the matched line segments. 39 line segments were matched. Fig. 8 (c) shows the reconstructed 3-D line segments. The three dots in the images represent the estimated camera positions. The grid interval is 2.0 m. The results of other scenes are shown in Fig. 9.

The number of matched line segments in the ten scenes is shown in Table I. The number of failures was counted manually. The proposed method achieved the smallest failure rate compared with the other three methods.

We measured the distance between the arrow-shaped indicators in Fig. 8 (b). The distance measured from the reconstruction result was 7.6 m, where the groundtruth was 7.8 m. We considered that the error was sufficiently small for free space detection.

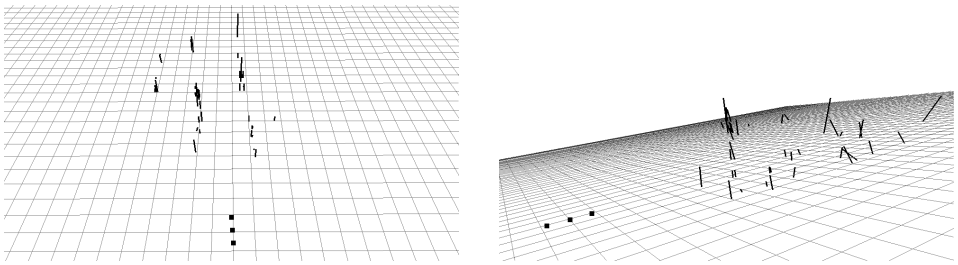
Some vertical line segments at a far distance in Fig. 8 (b) were not reconstructed as vertical line segments in Fig. 8 (c). This is because distant objects cause a small disparity. One of the solutions for this problem is to introduce preliminary knowledge about the street structure such as the Manhattan world assumption [16].

C. Application to Free Space Detection

1) Free Space Detection Method: As mentioned earlier, we define a free space as a region where a vehicle can pass through, such as a road. Several methods for the free space detection are based on the idea that the free space is an area without an obstacle [2]. In our method, following the same idea, the road and the other regions are simply distinguished using the reconstructed line segments. The detail of the method is as follows: First, the structure map is generated. The structure map is a plane which indicates the location of the line segments (Fig. 10 (a)). To generate a structure map, the reconstructed line segments are projected onto the ground. Then, objects at higher places than the vehicle are

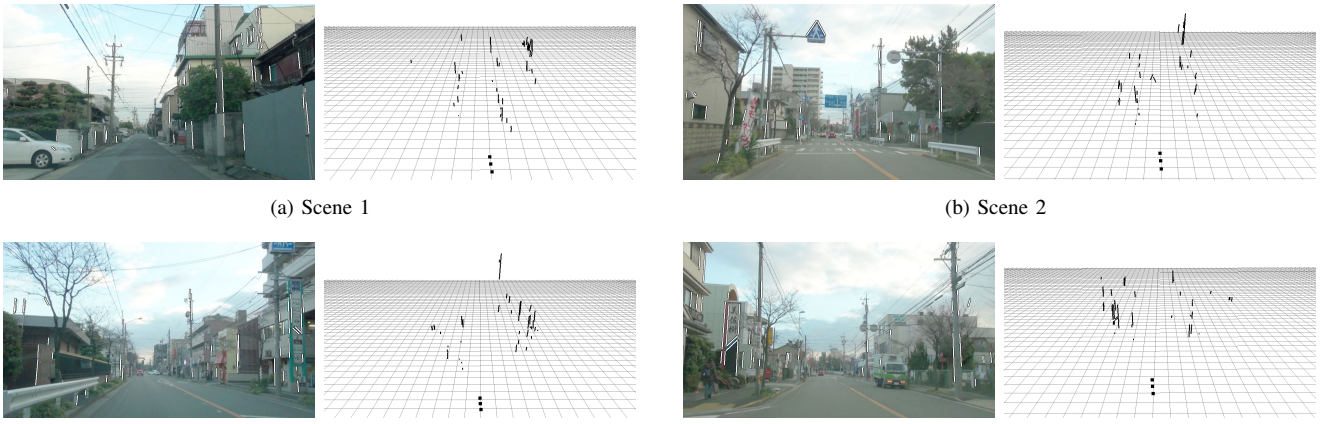
(a) Line segment detection result (Images 1, 2, and 3)

(b) Matching result (Images 1, 2, and 3)



(c) Reconstruction result

Fig. 8. Line segment reconstruction of a scene.



(a) Scene 1

(b) Scene 2

(c) Scene 3

(d) Scene 4

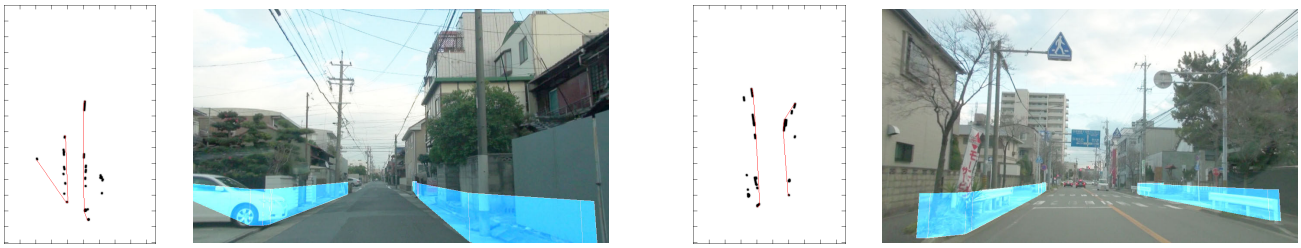
Fig. 9. Result of line segment reconstruction.

removed. Second, to emphasize the free space, the convex hulls of both roadsides are computed (Fig. 10 (b)).

2) Result: We applied the method to the same images with the previous experiment. Examples of the experimental results are shown in Fig. 11. The left images show the structure map, and the right images illustrate the convex hull. The grid interval of the structure map was 5.0 m. The convex hulls in the right images roughly separated the road region from the other regions. It can be observed that the roadside objects, such as poles, supported the convex hulls to separate the road and the other regions correctly.

IV. Conclusion

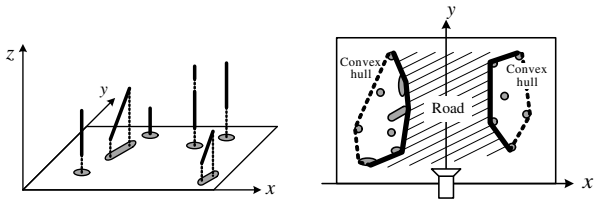
We proposed a method for reconstructing 3-D line segments in a streetscape using a monocular in-vehicle camera. The line segments were matched using a multi-view constraint, the property of the line segment, and the preliminary knowledge of the vehicle motion. To demonstrate the performance of the 3-D reconstruction, the 3-D reconstruction was applied to the free space detection.



(a) Scene 1 (b) Scene 2

(c) Scene 3 (d) Scene 4

Fig. 11. Application to free space detection. The left image shows the structure map, and the right image shows the wall separating road and other regions. The grid interval of the structure map is 5.0 m.



(a) Structure map (b) Convex hulls

Fig. 10. Free space detection method. (a) 3-D line segments are projected to the ground. (b) Convex hulls separate the road and other regions.

Future work includes introduction of preliminary knowledge of the city structure to improve the reconstruction accuracy. We also plan to conduct a larger experiment to evaluate under various environments.

Acknowledgement Parts of this research were supported by a Grant-In-Aid for Scientific Research from MEXT, and a Core Research for Evolutional Science and Technology (CREST) project of JST. MIST library (http://mist.murase.m.is.nagoya-u.ac.jp/) was used for developing the proposed method.

REFERENCES

[1] H. Badino, U. Franke and R. Mester, "Free Space Computation using Stochastic Occupancy Grids and Dynamic Programming," Proc. of ICCV2007 Dynamic Vision Workshop, xp.12, Oct. 2007.

[2] Y. Matsumoto, K. Ikeda, M. Inaba and H. Inoue, "Exploration and Navigation in Corridor Environment Based on Omni-view Sequence," Proc. of Int. Conf. on Intelligent Robots and Systems 2000, vol.2, pp.1505–1510, Nov. 2000.

[3] H. Zhao and R. Shibasaki, "A Vehicle-borne Urban 3-D Acquisition System using Single-row Laser Range Scanners," IEEE Trans. Systems, Man, and Cybernetics, vol.33, issue 4, pp.658–666, Aug. 2003.

[4] W. Mark and D. M. Gavrila, "Real-time Dense Stereo for Intelligent Vehicles," IEEE Trans. Intelligent Transportation Systems, vol.7, issue 1, pp.38–50, Mar. 2006.

[5] K. Yamazaki, M. Tomono, T. Tsubouchi and S. Yuta, " Object Shape Reconstruction and Pose Estimation by a Camera Mounted on a Mobile Robot," Proc. of Int. Conf. on Intelligent Robots and Systems 2004, vol.4 pp.4019–2025, Oct. 2004.

[6] S. Ikeda, T. Sato, K. Yamaguchi and N. Yokoya, "Construction of Feature Landmark Database using Omnidirectional Videos and GPS Positions, " Proc. of Int. Conf. on 3-D Digital Imaging and Modeling 2007, pp.249–256, Aug. 2007.

[7] I. Skrypnyk and D. G. Lowe, "Scene Modelling, Recognition and Tracking with Invariant Image Features," Proc. of Int. Symp. on Mixed and Augmented Reality 2004, pp.110–119, Nov. 2004.

[8] C. J. Taylor and D. J. Kriegman, "Structure and Motion from Line Segments in Multiple Images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.17, issue 11, pp.1021–1032, Nov. 1995.

[9] D. Martinec and T. Pajdla, "Line Reconstruction from Many Perspective Images by Factorization," Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition 2003, vol.1, pp.497–502, June. 2003.

[10] T. Asai, K. Yamaguchi, Y. Kojima, T. Naito and Y. Ninomiya, "3D Line Reconstruction of a Road Environment using an In-vehicle Camera," Proc. Int. Symp. on Visual Computing 2008, pp.897–904, Dec. 2008.

[11] R. I. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," Cambridge University Press, ISBN: 0521623049, 2000.

[12] C. Schmid and A. Zisserman, "Automatic Line Matching Across Views," Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition 1997, pp.666–671, June 1997.

[13] D. G. Lowe, " Distinctive Image Features from Scale-invariant Keypoints," Int. J. of Computer Vision, vol.60, no.2, pp.91–110, Jan. 2004.

[14] R. Gioi, J. Jakubowicz, J. Morel and G. Randall, " LSD: A Fast Line Segment Detector with a False Detection Control," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.32, issue 4, pp.722–732, Apr. 2010.

[15] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.22, issue 11, pp.1330–1334, Nov. 2000.

[16] E. Delage, H. Lee and A. Y. Ng, "Automatic Single-image 3D Reconstructions of Indoor Manhattan World Scenes," Robotics Research Results of the 12th Int. Symp. ISRR, Springer Tracts in Advanced Robotics, vol.28, pp.305–321, ISBN: 3540481109, 2007.