Special Feature: Spatial Information Technology towards Intelligent Vehicle Systems

Research Report

Free-viewpoint Image Reconstruction for ADAS Virtual Assessment

Takashi Machida, Satoru Nakanishi and Tomokazu Sato

Report received on Jan. 31, 2018

ABSTRACT Advanced driving assistance systems (ADASs) are currently being assessed through enormous numbers of field operational tests (FOTs) as part of ongoing efforts to prepare them for meeting the vast combinations of potential traffic situations that they can be expected to encounter in actual driving environments. In this paper, we propose a virtual assessment methodology for ADASs that is designed to reduce the running distance of a conventional FOT. More specifically, we reconstruct images at an arbitrary viewpoint from a three-dimensional (3D) point cloud and an omnidirectional image sequence measured by a 3D scanning system. For this purpose, we focus on a *view-dependent* depth map (VDDM) based on a 3D surface model reconstructed from the 3D point cloud. To reduce artifacts caused by incorrect depth information such as the unmeasured regions in the 3D point cloud and errors in the reconstructed 3D surfaces, we employ a *view-dependent* depth testing (VDDT) technique that compares depth information in the VDDM with that in the omnidirectional depth maps. In experiments conducted to evaluate our proposed method, we compared the differences between free-viewpoint and real images by using an image processing algorithm that detects lines on a road. The results show that the free-viewpoint images were similar to real images with respect to the outputs of the line detection algorithm.

KEYWORDS ADAS Assessment, Virtual Assessment, Field Operational Test, Free-viewpoint Image Rendering, *View-dependent* Depth Testing, *View-dependent* Texture Mapping

1. Introduction

Advanced driving assistance systems (ADASs) such as auto driving systems are designed to improve driving safety in order to reduce the risk of traffic accidents by detecting dangerous traffic situations. In creating ADAS assessments, vast amounts of field operational tests (FOTs) modeling such traffic scenarios are being conducted(1) because such systems must be capable of performing well under any traffic situation they can be expected to encounter in actual driving environments. However, an obvious problem related to FOTs is the time requirements. Especially in camera-based auto driving systems, the demand for ADAS virtual assessments is dramatically increasing because there are huge pixel patterns in the imagery captured under a variety of traffic situations.

A number of studies have attempted to tackle this issue by using virtual images created with computer graphics (CG) techniques.(2-4) These generated images are expected to be input images for the image processing algorithms, such as vehicle/pedestrian detection and line detection scenarios. However, when images are rendered using conventional methods, there is a high possibility that inaccurate evaluations will result due to the low quality of the generated images. More specifically, there are often significant variations in the results provided by the image processing algorithms used for the real and virtual images.(5) To resolve this problem, free-viewpoint image reconstruction techniques based on real images are being researched as a way to generate high-quality images of real world traffic situations. However, these techniques have not yet achieved a sufficient level of accuracy to permit their use in ADAS virtual assessments, primarily due to the inadequate quality of generated images, which include three-dimensional (3D) geometry errors and image information shortfalls.

In this paper, we propose a free-viewpoint image reconstruction method that can produce images at arbitrary virtual camera positions, as shown in **Fig. 1**. To reconstruct accurate image appearance for virtual viewpoints, we first measure the surrounding areas, in which the positions and directions of the sensors are known, as a 3D point cloud and an omnidirectional image sequence by means of a vehicle-mounted 3D scanning system. Then, we generate depth maps according to virtual viewpoints using a surface reconstruction technique.(6) To compensate for the parts that have incorrect depth information in the generated depth maps due to hidden surfaces or points, we estimate correct depth information by comparing two types of depth map data. One is omnidirectional depth data at measured camera positions, while the other is depth data generated according to virtual viewpoints. Ultimately, we can generate free-viewpoint images by properly pasting the pixels from the omnidirectional images into the depth maps.

Our work has the following two main contributions: (1) achievement of a high-quality free-viewpoint image reconstruction methodology that is sufficiently detailed for application to ADAS virtual assessments, and (2) verification of the performance of the virtual images used instead of actual images for evaluating ADAS image processing algorithms. The remainder of this paper is structured as follows. In Sec. 2, we review ADAS virtual assessment and free-viewpoint image rendering techniques. In Sec. 3, we detail our proposed framework. In Sec. 4, we show the evaluation results of our proposed framework. Finally, in Sec. 5, we summarize the present work.

2. Related Work

2. 1 3D CG Model-based Approach for Virtual Environment Representation

To achieve ADAS virtual assessments, one of the more well-known methods is to present virtual environments in which any traffic situations can be reconstructed using 3D CG models.(2-4) TASS International has developed such virtual representation methods for ADAS virtual assessments, and image processing algorithms can be virtually evaluated.(3) In the same way, Weinberg et al.(4) developed a driving simulator by using 3D CG models.

It is also expected that CG images will eventually be applied to evaluating image processing algorithms, even though the quality of generated images is currently inadequate for ADAS virtual assessments due to significant differences in the results produced by image processing algorithms for virtual and real images.(5) Among the methods developed to present the virtual world accurately are 3D scanning systems such as the GeoMaster NEO(7) and the IP-S2(8) mobile mapping systems. These systems can easily acquire 3D geometries and real world imagery, which enables them to produce high-quality generated images. However, another process-related problem with this method is that it is very costly to create 3D surface models and align images from such scanned data. Thus, when using a 3D CG model-based approach, there are two primary problems that must be addressed: (1) the extreme difficulty of generating high-quality CG images using only 3D CG models, and (2) the huge expense of creating high-precision 3D models and rendering those models with photo-realistic image appearances.

For the first problem as mentioned above, Ono et al.(9) reconstructed virtual images that consisted of two types of image data in order to increase the quality of the generated images. One is a real image used as background (long-distance view) scenery, which is acquired by a 3D scanning system, while the other is 3D CG image data that is input as foreground (short-distance view) scenery. Although this method does enable a comprehensive increase in the quality of generated imagery, it is not sufficiently advanced for use in image generation for ADAS virtual assessments because the foreground scenery, such as poles, road

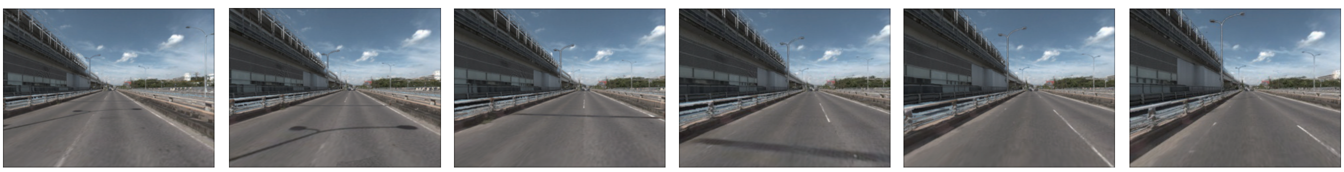


Fig. 1 Examples of generated images for virtual camera path. These images are generated from data captured from a vehicle traveling in the right lane.

surfaces, and guardrails, must still be included using manually designed 3D CG models.

For the second abovementioned problem, it is of primary importance to generate 3D surfaces from the 3D point cloud in order to represent 3D CG models more precisely. One of the 3D surface modeling techniques used to accomplish this is to approximately fit 3D primitives to the 3D point cloud.(10) However, there is a limit to the simple 3D shapes that can be adequately fit using such primitives. In another 3D surface modeling technique, a number of small planes are placed properly on the 3D point cloud.(11) To generate smoother surfaces than those seen in that technique, Liang et al.(12) proposed a more accurate method that determines whether the generated surfaces in the 3D point cloud are within the inside or outside regions based on image segmentation. However, sensor calibration errors and sensor resolution limitations result in two severe problems. One is an artifact that causes the object to look more distorted than the actual geometry, and the other is an artifact that makes it appear as if the object exists in a missing region within the 3D point cloud.

2. 2 *View-dependent* Based Approach for High-quality Virtual Environment Representation

As discussed above, these conventional methods employ *view-independent* 3D models and texture images that can directly present the virtual environment at an arbitrary camera position. However, in such cases, there is a high probability that artifacts caused by errors in the reconstructed 3D geometry will appear in the rendered images. Additionally, although virtual image representation techniques have potential abilities to make image processing algorithm testing more efficient, current techniques cannot realize an adequate quality of generated imagery for use in ADAS virtual assessments.

To solve these issues, the use of *view-dependent* methodologies that incorporate *view-dependent* depth testing (VDDT) and *view-dependent* texture mapping (VDTM) is expected to provide a more effective approach. VDTM, which was first proposed by Debevec et al.,(13) is an artifact-less method that works by reprojecting images according to a virtual viewpoint. In contrast, VDDT, which was proposed by Sato et al.,(14) produces more correct depth information judgements in order to reduce the artifacts in generated images.

In this paper, we extend the existing free-viewpoint image reconstruction techniques by integrating these two *view-dependent* methods (VDTM and VDDT), in order to generate photo-realistic images that are suitable for use in ADAS virtual assessments.

3. Free-viewpoint Image Reconstruction from 3D Point Cloud and Omnidirectional Images

3. 1 Overview

Figure 2 shows an overview of our proposed method, which consists of the two following processes: (a) the data preparation process, and (b) the image rendering process.

The data preparation process (a) deals primarily with increasing the quality of the virtual image, as mentioned in Sec. 2. 1. In this process, we acquire 3D point cloud and omnidirectional images in real world by using GeoMaster NEO(7) as the 3D scanning system. Then, the 3D surface models are presented properly by fitting 3D patches to the large part of the 3D points based on probabilistic sampling method. The 3D surface models and the rest of the 3D points are used for generating omnidirectional depth maps in this process and *view-dependent* depth maps (VDDMs) in the next process.

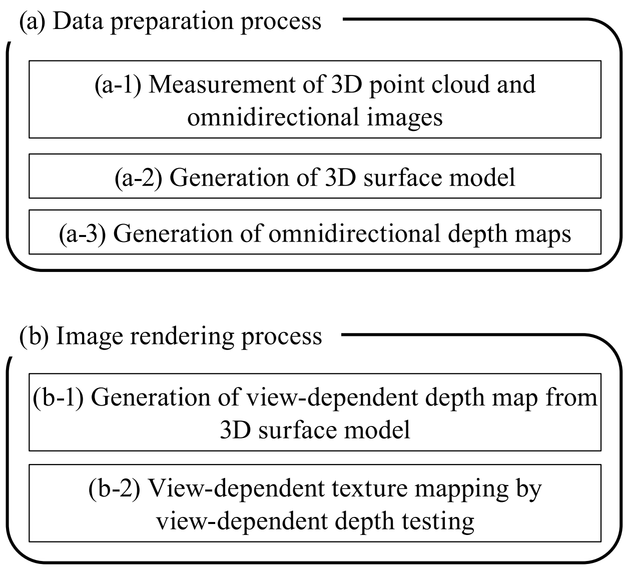


Fig. 2 Overview of the proposed framework.

In contrast, the rendering process (b) deals primarily with reducing the image artifacts in the *view-dependent* methodology mentioned in previous Sec. 2. 2. In this process, we calculate a VDDM according to a virtual viewpoint by sampling the surface models prepared in process (a). The depth map is then refined by interpolation with surrounding depth information and by comparisons with the omnidirectional depth maps generated using the VDDM in process (a). Finally, a free-viewpoint image is reconstructed with the refined depth map using a VDTM technique. Each process is more fully detailed in the following subsections.

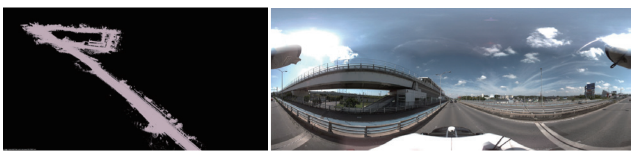
3. 2 Data Preparing Process

(a-1) Measurement of 3D Point Cloud and Omnidirectional Images

In our proposed method, as explained above, 3D point cloud data and omnidirectional images (as shown in **Fig. 3**) are acquired using GeoMaster NEO,[7] which is a 3D scanning system that produces maps with precisely known positions and directions. In this system, RIEGL laser scanner, Point Grey Ladybug 3, RTK-GPS and an odometry sensor are mounted on the mapping vehicle that can be easily acquire data by being driven through the target area.

(a-2) Generation of 3D Surface Model

To create a VDDM in the next process, we craft a 3D surface model estimation from a 3D point cloud. In this case, we employ a plane fitting approach based on RANSAC that is similar to that used in Ref. (6). First, a number of 3D points are randomly selected from the scanned 3D point cloud. The likelihood of a plane existence is then judged using the sampled 3D points. Then, 3D points that exist on the identified plane are collected to make parts for the surface patches. These processes are repeated until almost all 3D points are incorporated into one of the identified planes.



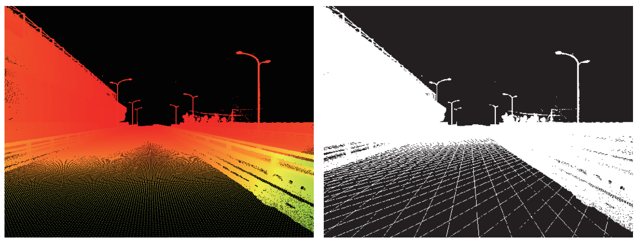
(a) 3D point cloud (b) Omnidirectional image

Fig. 3 Data acquired by GeoMaster NEO.[7]

Figure 4 shows the surface models of the scanned road environment. Although there are some 3D points that are not categorized as belonging to the planes in the virtual environment, it is confirmed that a large percentage of the 3D points is presented as a set of simple square polygons. In the next process, the VDDMs for free-viewpoint image reconstruction are efficiently created based on these planes.

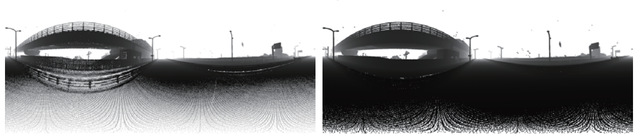
(a-3) Generation of Omnidirectional Depth Maps

To reduce artifacts in generated images, it is important to determine whether the visibilities of points on the 3D surface models constructed in (a-2) are correct. Hence, we compare the depth information obtained from the measured camera position to the depth information obtained from the virtual camera position. For this purpose, we calculate omnidirectional dense depth maps by reprojecting the 3D point cloud onto each omnidirectional image, as shown in **Fig. 5**. Sparse depth data are first acquired by rendering the 3D point cloud (Fig. 5(a)). Then, dense depth maps are created (Fig. 5(b)) by expanding and interpolating the sparse depth information. Note that the infinite depth region is dealt with as the sky region.



(a) 3D point cloud (b) 3D surface model

Fig. 4 3D surface model generated by.[6]



(a) 3D point cloud (b) Depth map

Fig. 5 Omnidirectional depth map.

3. 3 Image Rendering Process

(b-1) Generation of VDDM from 3D Surface Models

In the image rendering process, we use a VDDM created using the VDTM method. More specifically, we first estimate an initial depth map from 3D surface models in a *view-dependent* manner. Since a virtual camera position and direction can be set freely in the virtual environment using 3D surface models, the depth map used in our method is simply created by using OpenGL to produce a depth test for the virtual camera position and direction. However, depth information will still be incomplete due to missing regions on the 3D surface models (such as the black part on a guardrail), as shown in **Fig. 6**(a). Next, the depth map is refined to reduce the lack of the depth information by interpolating it with the surrounding depth information, as shown in Fig. 6(b).

(b-2) VDTM by VDDT

In the final process of our proposed method, the color of each pixel in the generated image is determined by VDTM according to the virtual viewpoints based on the omnidirectional image sequence that is acquired in (a-1).[(14)]

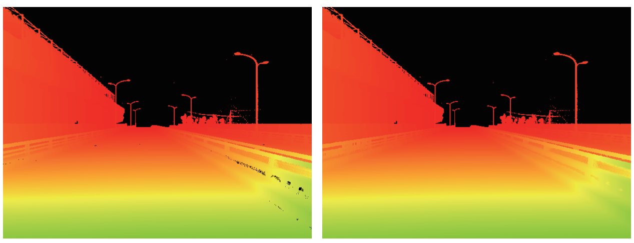
Figure 7 clearly shows that the free-viewpoint image is reconstructed by selecting the *n*-th omnidirectional image for each pixel according to the following equation:

$$n = \arg\min_i \alpha\theta_i + d_i \; (i = 1, 2, ..., m), \qquad (1)$$

where θ_i denotes the angle between $\overline{p_i p_s}$ and $\overline{p_v p_s}$ for each measured camera position p_i at which the i-th omnidirectional image is captured, and d_i denotes the distance $\overline{p_i p_s}$ for each measured camera position p_i. In addition, p_v and p_s denote the virtual camera position and the point on the 3D surface model corresponding to the depth map, respectively. For accurate image reconstruction, the pixel color in the i-th omnidirectional image is selected when both θ_i and d_i are the smallest values among the omnidirectional image set. This is because when θ_i and d_i are the smallest values, the appearance of the 3D point p_s is the closest to the pixel in the i-th omnidirectional image at the measured camera position p_i. Thus, selecting omnidirectional image with Eq. (1) can be expected to increase the accuracy of the generated image. In the case of the situation shown in Fig. 7, the omnidirectional image measured at position p_1 is selected rather than that of the position p_2, as mentioned above. Note that the weight α is determined experimentally.

Since the VDTM method is incapable of pasting the pixels of other objects onto a region of sky that has an infinite depth region, it is necessary to reduce the artifacts resulting from this inability. We consider the VDDT to be an efficient technique for use when compensating for infinite depth region artifacts. However, for extremely noisy inputs, it is important to extend it to a more robust depth test scheme than shown in Ref. (11).

We make our visibility judgments based on the differences between depth information on the omnidirectional depth map reconstructed in (a-3) and the depth information on the VDDM reconstructed in (b-1). In the case of Fig. 7, although the camera position selected by Eq. (1) is, once again p_2, p_1 is selected by comparing the depth $\overline{p_v p_s}$ with the depth $\overline{p_v p_s'}$, where



(a) Initial depth map (b) Refined depth map

Fig. 6 Dense depth map estimation using 3D surface model.

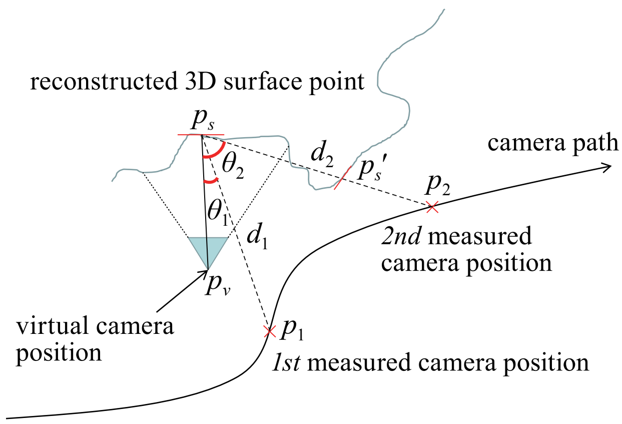


Fig. 7 Parameters for *view-dependent* texture mapping.

p_s' denotes the point that is actually reprojected for the virtual camera position p_v because p_s is occluded by p_s' at the measured camera position p_2.

4. Experiments

In this section, we report on experiments conducted to clarify the performance of the proposed method for evaluating an image processing algorithm as part of an ADAS virtual assessment. More specifically, we compare the outputs of the line detection algorithm(15) for real and virtual images using the 3D point cloud and omnidirectional image sequence shown in Fig. 3, which were acquired by GeoMaster NEO.(7)

We began by generating 34-frame image sequences in which the virtual camera travels in a straight line for 60 m. To verify the differences between the real and the virtual images, the virtual viewpoint in each generated frame was the same as the real viewpoint captured during measurements of the road environment. **Figure 8** shows a comparison of (a) the virtual and (b) real images at two frames. From these results, it was confirmed that the generated images were very similar to the real images, except for insignificant artifacts. In the current image rendering process, the proposed method can be computed in real time (10 fps) for images of 640 × 480 resolution using a personal computer (PC) equipped with an Intel Core i7-3970X CPU and an nVidia GTX1080 GPU. Figure 8 also shows the white lines detected from (c) the virtual and (d) real images respectively by the existing image processing algorithm.(15) From these detected results, it can also be seen that the differences between the virtual and real images are negligible.

Figure 9 quantitatively shows offset positions of white lines detected from the virtual and real images with the method outlined in Ref. (15). The average absolute differences between pairs of offset positions for the real and virtual images are 65 mm (right white line) and 38 mm (left white line), respectively. These results are deemed sufficient for evaluating the stability and the accuracy of the line detection algorithm.

Since we are still in the prototype stage of our research, these experiments were conducted with just a few images. However, it will eventually be necessary to deal with a much larger environment that has a number of image sequences and a huge 3D point cloud. To tackle such a large-scale evaluation, it will be necessary to consider the computational cost for generating images. In the current data preparation process, generating a 3D surface model takes 10 minutes for an area of 50 m squares. In the current

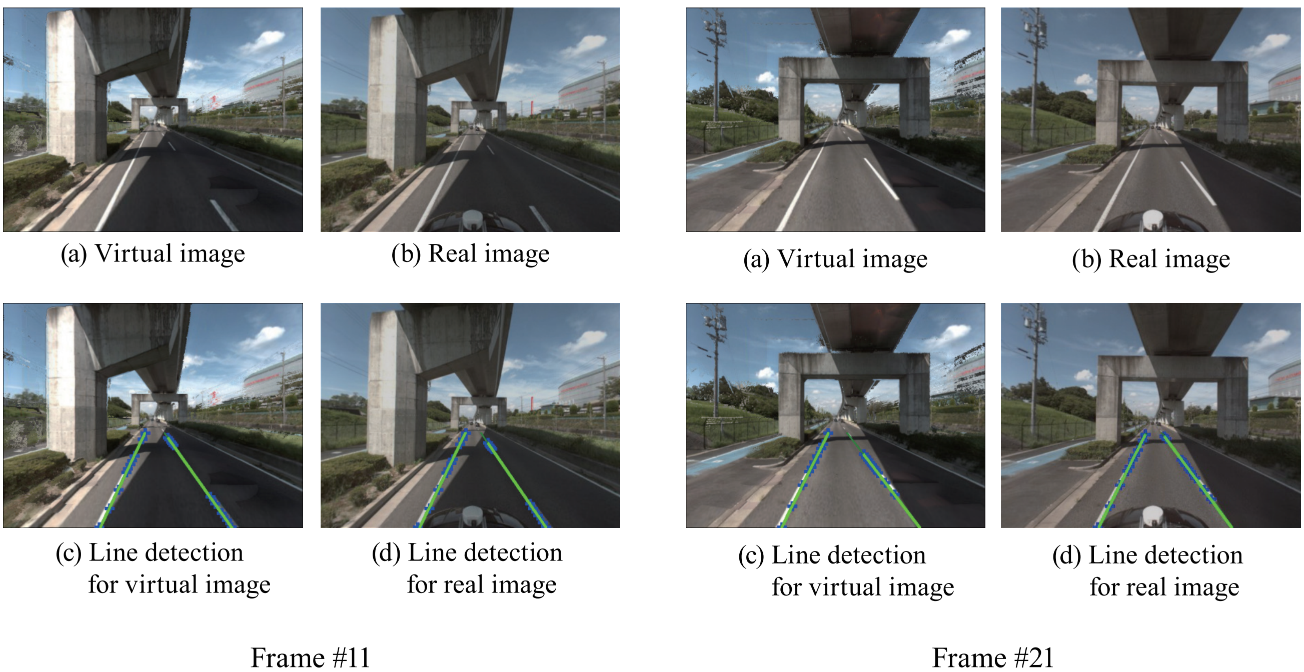


Fig. 8 Results of white line detection for the virtual and real images.

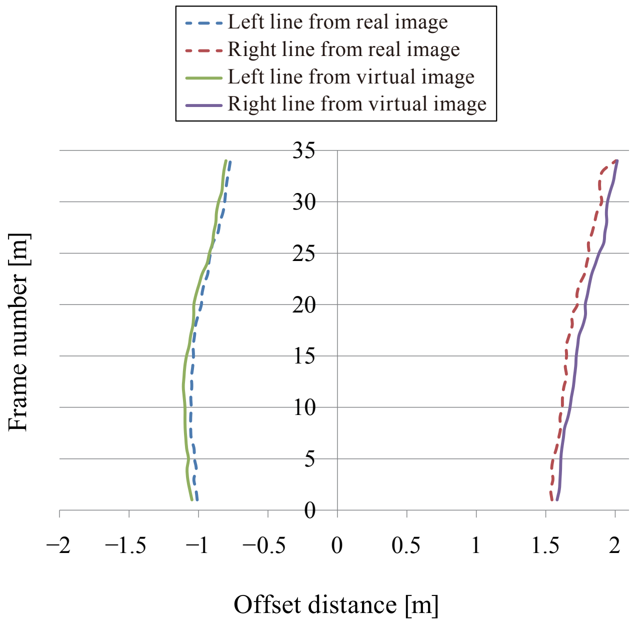


Fig. 9 Offset positions of detected white lines for the virtual and real images.

image rendering process, although free-viewpoint imagery is generated in real-time, the computational time will increase in proportion to the number of 3D points. Hence, it will be necessary to reduce the number of such points necessary for reconstructing a large-scale environment.

5. Conclusion

In this paper, we proposed a free-viewpoint image reconstruction framework that can be used to facilitate ADAS virtual assessments and, via experiments, confirmed that the free-viewpoint images generated by the proposed method were similar to the actual images. We also confirmed that the differences between detected positions extracted from the virtual and real images via the line detection algorithm were very small. These results indicate that our proposed framework has the potential to create ADAS virtual assessments with much higher accuracy levels than can be obtained from conventional FOT efforts. In our future work, we will extend the proposed framework to deal with a much larger environment, various ADAS image processing algorithms, and a wide variety of traffic situations.

Reference

(1) Zhao, D., “Accelerated Evaluation of Automated Vehicles”, Dr. Thesis, University of Michigan (2016), 131p.

(2) FORUM8, “Our Products”, <http://www.forum8.co.jp/product/products.htm>, (accessed 2017-11-24).

(3) TASS International, “PreScan Overview”, <https://www.tassinternational.com/prescan-overview>, (accessed 2017-11-24).

(4) Weinberg, G. and Harsham, B., “Developing a Low-cost Driving Simulator for the Evaluation of In-vehicle Technologies”, *Proc. 1st Int. Conf. Automot. User Interfaces Interact. Veh. Appl.* (2009), pp. 51-54.

(5) Yasuda, H., Yamada, N. and Teramoto, E., “An Evaluation Measure of Virtual Camera Images for Developing Pedestrian Detection Systems”, *Proc. IEICE General Conf.* (in Japanese) (2011), D-11, p. 80.

(6) The Point Cloud Library, “Plane Model Segmentation”, <http://pointclouds.org/documentation/tutorials/planar_segmentation.php>, (accessed 2017-11-24).

(7) Asia Air Survey, “GeoMaster NEO”, <http://www.ajiko.co.jp/product/detail/ID4T0WD5Z0D>, (accessed 2017-11-24).

(8) Topcon, “IP-S2”, <http://www.topcon.co.jp/positioning/atwork/imaging/201008_kobayashi_IP-S2_J.html>, (accessed 2017-11-24).

(9) Ono, S., Sato, R., Kawasaki, H. and Ikeuchi, K., “Image Based Rendering of Urban Road Scene for Real-time Driving Simulation”, *Proc. ASIAGRAPH* (2008), p. 1.

(10) Schnabel, R., Degener, P. and Klein, R., “Completion and Reconstruction with Primitive Shapes”, *Proc. Eurographics*, Vol. 28, No. 2 (2009), pp. 503-512.

(11) Yang, R., Guinnip, D. and Wang, L., “*View-dependent* Textured Splatting”, *The Visual Comput.*, Vol. 22, No. 7 (2006), pp. 456-467.

(12) Liang, J., Park, F. and Zhao, H., “Robust and Efficient Implicit Surface Reconstruction for Point Clouds Based on Convexified Image Segmentation”, *J. Sci. Comput.*, Vol. 54, No. 2-3 (2013), pp. 577-602.

(13) Debevec, P., Yu, Y. and Borshkov, G., “Efficient View-dependent Image based Rendering with Projective Texture-mapping”, *Proc. Eurographics Workshop* (1998), pp. 105-116.

(14) Sato, T., Koshizawa, H. and Yokoya, N., “Omnidirectional Free-viewpoint Rendering Using a Deformable 3-D Mesh Model”, *Int. J. Virtual Reality*, Vol. 9, No. 1 (2010), pp. 37-44.

(15) Watanabe, A. and Nishida, M., “Lane Detection for a Steering Assistance System”, *Proc. IEEE Intell. Veh. Symp.* (2005), pp. 159-164.

Takashi Machida
Research Field:
- Real World Representation Based on Morphologic Approach for ADAS Virtual Assessment

Academic Degree: Ph.D.
Academic Society:
- Information Processing Society of Japan

Awards:
- IPSJ Best Paper Award, 2006

Satoru Nakanishi
Research Field:
- Virtualized Reality Based on Computer Vision and Computer Graphics for ADAS Virtual Assessment

Tomokazu Sato*
Research Fields:
- Computer Vision
- Free Viewpoint Image Generation
- Mixed Reality

Academic Degree: Ph.D.
Academic Societies:
- IEEE
- The Institute of Electronics, Information and Communication Engineers
- Information Processing Society of Japan
- The Institute of Image Information and Television Engineers
- The Virtual Reality Society of Japan

Awards:
- Young Researcher's Award, IEICE, 2001

*Nara Institute of Science and Technology

