

# Driving View Simulation Synthesizing Virtual Geometry and Real Images in an Experimental Mixed-Reality Traffic Space

Shintaro Ono<sup>†</sup> Koichi Ogawara<sup>†</sup> Masataka Kagesawa<sup>†</sup> Hiroshi Kawasaki<sup>††</sup>  
 Masaaki Onuki<sup>‡</sup> Ken Honda<sup>‡‡</sup> Katsushi Ikeuchi<sup>†</sup>

<sup>†</sup>The University of Tokyo  
 4-6-1 Komaba, Meguro-ku,  
 Tokyo 153-8505 JAPAN  
 {onoshin, ogawara, kagesawa,  
 ki}@cvl.iis.u-tokyo.ac.jp

<sup>††</sup>Saitama University  
 255 Shimo-Okubo, Sakura-ku,  
 Saitama 338-8570 JAPAN  
 h-kawa@mm.ics.  
 saitama-u.ac.jp

<sup>‡</sup>Mitsubishi Precision Co., Ltd.  
 345 Kami-Machiya,  
 Kamakura, Kanagawa  
 247-8505 JAPAN  
 monuki@mpcnet.co.jp

<sup>‡‡</sup>Chodai Co., Ltd.  
 1-20-4 Nihombashi-  
 Kakigaracho, Chuo-ku,  
 Tokyo 114-0013 JAPAN  
 honda-k@chodai.co.jp

## Abstract

We propose an efficient and effective image generation system for an experimental mixed-reality traffic space. Our enhanced traffic/driving simulation system represents the view through a hybrid that combines virtual geometry with real images to realize high photo-reality with little human cost. Images for datasets are captured from the real world, and the view for the simulation system is created by synthesizing image datasets with a conventional driving simulator.

## 1. Introduction

We have been developing a novel driving view simulation system that uses mixed-reality techniques to create a realistic traffic space. This system is an extended framework of a conventional driving/traffic simulator<sup>1</sup>. Unlike conventional driving simulators, the view the user sees is produced by synthesizing a geometric model and real video images. In general, geometric models are suited for dealing with versatile operations such as embedding other objects; however, they have relatively poor photo-reality, and they require some manual construction work. Image-based models can produce a photo-realistic view; however, they are poor at versatile operations. We therefore use both models, giving them different roles and synthesizing them into a single view that emphasizes the merit of each and compensates for the defects of each.

Among related works on this topic, the urban scene is reconstructed by a geometry-based method in [6, 7, 8], and by an image-based method in [9]. Both approaches are used for rendering in [10, 11]. Driving simulators are actively developed by private enterprises and academic fields [14, 15]. Some works deal with real images [12, 13]; however, they assume that users drive along a designated curve and make no lane changes.

## 2. System overview

The overall picture of the system is represented in Fig. 1. Traffic Simulator (TS) [2] simulates macroscopic traffic flows, and Driving Simulator (DS) [3] recreates the microscopic behavior of the vehicle as the user operates it. KAKUMO [2, 3] simulates the microscopic position of each vehicle from the macroscopic traffic flow.

<sup>1</sup>Developed in the collaborative research project ‘‘Sustainable ITS’’ at The Univ. of Tokyo, since Apr.2003 (<http://www.its.ccr.u-tokyo.ac.jp>).

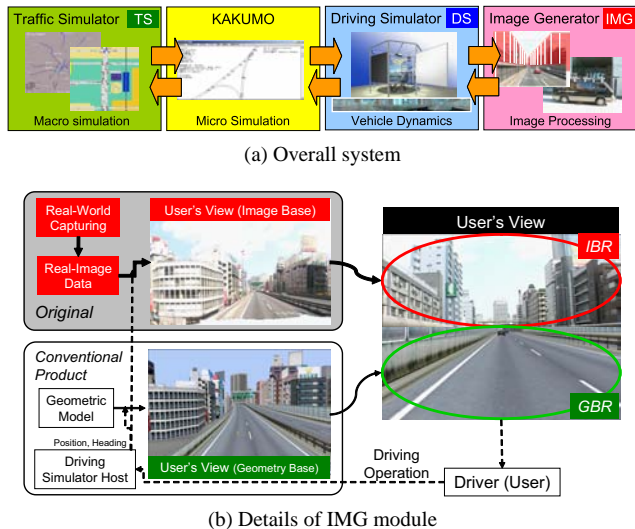


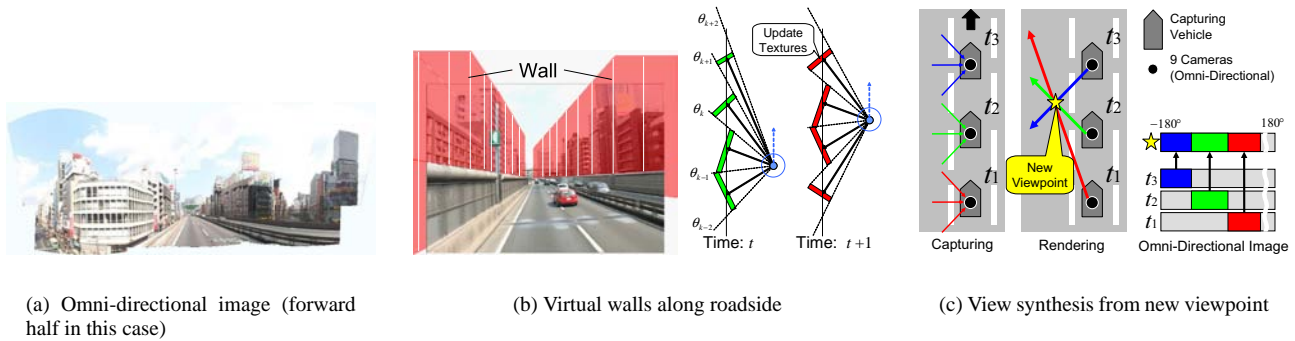
Figure 1. System overview

IMG is the module that we were in charge of developing, and the model we describe in detail in this paper. It generates the view the user sees in real-time (60Hz in frame rate) with high reality according to the position of the user’s vehicle given from DS. We use both geometry- and image-based approaches and synthesize them in the display phase. For the near view, including roads, traffic signs, and other vehicles, we take the geometry-based approach, using a conventional product of Mitsubishi Precision Co. Ltd. For the far view, including buildings and sky, we take the image-based approach using a technique developed to process video images captured along the model course. We describe these techniques in the next section of this paper.

## 3. Reconstruction of view from user viewpoint

For the far view seen by the user, we capture images from the real world. Our capturing vehicle, whose roof is equipped with nine video cameras, runs along the targeted road. The nine camera images are blended to become an omni-directional video image stream [4] along the road (Fig.2(a)). Once these omni-directional images, viewed from the running path, are accumulated, a view from outside the path can be created through image processing. Therefore, the capturing operation has to be carried out only once.

Geometric models assumed in the image-based near



**Figure 2. View synthesis of image part**

view are just two simple walls along the roadside (Fig. 2(b) left). The walls are split into slits. When rendering the image-based part, the slits are rotated toward the user's visual-line direction.

If the user's viewpoint is located within the capturing path, the part of the omni-directional image that extends from  $\theta_k$  to  $\theta_{k+1}$  is mapped to the  $k$ -th slit as textures (Fig. 2(b) right), where  $\theta_n$  is a direction to the edge of the  $n$ -th slit. If the user's viewpoint is located out of the capturing path, we first create *virtual* omni-directional image by improved method of [5], and map it to the slits. In Fig.2(c) for example, left view from a star-signed point can be composed of forward-left, left, and backward-left part of omni-directional image captured at time  $t_1, t_2, t_3$  respectively.

The image-based part and the geometry-based part are independently rendered by separate computers. The geometry-based part uses a conventional Driving Simulator product. Both RGB values and depth value of each pixel are output through DVI terminals. To integrate each view, a hardware product called VizCluster is used. It receives the output, determines the RGB values of each pixel according to the depth value, and passes them to multiple projectors.

#### 4. Rendering result

Fig. 3 shows the synthesized result of image and geometry parts. The dashed line in the right figure represents the boundary of both parts. Sky, buildings, and advertising displays are rendered realistically, and vehicles can dynamically change their location. Distortions seen in the demo video are caused by several conditions. One is a projection to just two flat walls, which becomes especially unrealistic when objects like an elevated bridge intrude. We are planning to add one more wall above the road to avoid the problem. Vibration of the capturing vehicle also causes defects. This vibration can be mitigated to some extent by calculating a self pose of capturing vehicle matching each frame of the omni-image to adjacent frames.

#### 5. Summary

This paper introduced our driving-view generation system, which is composed of a geometry-based and an image-based part and rendered by methods appropriate for each part. This system can offer a realistic view with a freely moving viewpoint and simultaneously can realize versatile

uses such as changing behaviors of other vehicles. For the future, we are planning to improve the quality of the image-based part by the methods described above, and to apply this system to driving simulation in sag a zones<sup>2</sup> of highways.

**Acknowledgement:** This work was in part supported by JSPS Fellow No.16-10876 and ITS division, National Institute for Land and Infrastructure Management.

#### References

- [1] K. Ikeuchi et al., "Mixed Reality Traffic Experiment Space under Interactive Traffic Environment for ITS Research", Proc. 11th World Congress on ITS, 2004
- [2] T. Shiraiishi et al., "Development of a Microscopic Traffic Simulation Model for Interactive Traffic Environment", Proc. 11th World Congress on ITS, 2004
- [3] Y. Suda et al., "Development of driver model using driving simulator with interactive traffic environment", Proc. 11th World Congress on ITS, 2004
- [4] H. Kawasaki et al., "Multiple video camera calibration using EPI for city modeling", Proc. Asian Conference on Computer Vision, 2004
- [5] T. Takahashi et al., "Arbitrary view position and direction rendering for largescale scenes", Proc. Computer Vision and Pattern Recognition, Jun. 2000
- [6] C. Frueh et al., "Constructing 3D City Models by Merging Ground-Based and Airborne Views", Proc. Computer Vision and Pattern Recognition, 2003
- [7] H. Zhao et al., "Reconstruction of Textured Urban 3D Model by Fusing Ground-Based Laser Range and CCD Images", IEICE Trans. Inf. and Syst., Vol.E83.D, No. 7, pp.1429,1440, 2000
- [8] S. Teller et al., "Calibrated, Registered Images of an Extended Urban Area", Proc. IEEE Computer Vision and Pattern Recognition, 2001
- [9] D. Kotake et al., "Cybercity Walker 2001: Walking Through and Looking Around a Realistic Cyberspace Reconstructed from the Physical World", Proc. 2nd Int '1 Symposium on Mixed Reality, pp.205,206, 2001
- [10] D. Aliaga et al., "MMR: An Integrated Massive Model Rendering System Using Geometric and Image-Based Acceleration", Proc. Symposium on Interactive 3D Graphics, pp. 199-206., 1999
- [11] G. U. Carraro et al., "Techniques for Handling Video in Virtual Environments", Proc. SIGGRAPH, 1998
- [12] M. Katakura et al., "Development of Moving Scenery Image Experiment System" (in Japanese), Annual Conf. of Japan Society of Civil Eng., pp.57.4, 2002
- [13] H. Ohta et al., "Effects of Traffic Safety Devices on Driver's Behavior at a Curve Section"(in Japanese), Infrastructure Planning (Japan Society of Civil Eng.), 2002
- [14] T. Shiiba et al., "Virtual Running Tests of Automobile with Driving Simulator" Int '1 Jnl. of ITS Research, Vol. 1, No. 1, pp.33,39, 2003
- [15] J.H. Kim et al., "Acquisition and Modeling of Driving Skills by Using Three Dimensional Driving Simulator", IEICE Trans. Fundamentals Vol. E88.A, No. 3, pp.770,778, 2005



**Figure 3. Synthesized rendering result**

<sup>2</sup>Zones with a series of subtle gradient change causing traffic jam.