A Photo-Realistic Driving Simulation System for Mixed-Reality Traffic Experiment Space

Shintaro Ono*, Koichi Ogawara†, Masataka Kagesawa‡, Hiroshi Kawasaki§, Masaaki Onuki¶, Junichi Abeki¶, Toru Yano¶, Masami Nerio¶, Ken Honda‖, Katsushi Ikeuchi‡

* Graduate School of Information Science and Technology, The University of Tokyo
Ee–405 Institute of Industrial Science, 4–6–1 Komaba, Meguro-ku, Tokyo 153–8505 JAPAN
Email: onoshin@cvl.iis.u-tokyo.ac.jp
Telephone: +81–3–5452–6242, Fax: +81–3–5452–6244

† Institute of Industrial Science, The University of Tokyo
Email: ogawara@cvl.iis.u-tokyo.ac.jp

‡ Graduate School of Interdisciplinary Information Studies, The University of Tokyo
Email: {kagesawa,ki}@cvl.iis.u-tokyo.ac.jp

§ Faculty of Engineering, Saitama University
Email: kawasaki@mm.ics.saitama-u.ac.jp

¶ Mitsubishi Precision Co., Ltd.
Email: {monuki,j_abeki,tyano,nerio}@mpcnet.co.jp

‖ Chodai Co., Ltd.
Email: honda-k@chodai.co.jp

***Abstract*— In this paper, we propose an efficient and effective image generation system for "Mixed Reality Traffic Experiment Space", an enhanced driving/traffic simulation system which we have been developing for Sustainable ITS project at the University of Tokyo. Conventional driving simulators represent ther view by a set of polygon-based objects, which leads to less photo-reality and huge human costs for dataset construction. We introduce our image/geometry-based hybrid method to realize more photo-realistic view with less human cost at the same time. Images for datesets are captured from real world by multiple video cameras mounted on a data acquisition vehicle. And the view for the system is created by synthesizing the image dataset. Following contents mainly describe details on data acquisition and view rendering.**

I. Introduction

An endeavor to reconstruct three-dimensional urban models on a virtual space in a computer has become highly interested research topics in the field of computer vision and graphics, virtual- and mixed- reality, remote sensing, architectonics, etc. Such models are expected to benefit various kinds of applications such as city planning, disaster prevention, intelligent transport systems, etc.

Since April 2003, we have been developing a novel mixed-reality simulation system called "Mixed-Reality Traffic Experiment Space" as one part of Sustainable ITS Project[6], a collaborative research project established in Center for Collaborative Research, The University of Tokyo. This simulator is an extended framework of conventional driving/traffic simulator. A macroscopic change of traffic flow and microscopic behaviors of each vehicle based on vehicle dynamics are integrated and aimed to recreate realistic driving situation. Moreover, a view from a driver is produced by synthesizing real video images in real time with high photo-reality. The system currently targets Tokyo Metropolitan Expressway as a model scene.

A view which should be provided to the user in this simulator is nothing less than a view of virtual urban model from ground level. Generally, approaches to reconstruct or represent such spatial model are divided into two types: One is *geometry-(polygon-) based* approach where the view is created with three-dimensional geometric information and surface reflectance attribute of the objects inside. The other is *image-based* approach where the view is created only by processing and synthesizing real video images acquired and accumulated in advance.

Conventional driving simulators often seen in driving schools or railway companies provide driver's view by geometry-based rendering as shown in Fig. 1 . Geometry-based models have relatively less data size, however, their view is poor at photo-reality. Additionally, the development of geometry models such as buildings and traffic signals needs a great deal of human work, which leads to one of the main cause of vast development cost.

Image-based model, on the other hand, is able to produce highly photo-realistic view. A quality of view is essential for our future attempt to collect acknowledgment and decision parameters in human driving operation by using this simulation system. However, image-based approach is not appropriate for

Fig. 1. Geometry-based view of driving environment

interactive use such as dynamically superposing other objects as other vehicles and pedestrians.

In this paper, we propose an novel method to offer useful and valuable view to the user in real-time. Geometry-based model and image-based model are properly used according to their roles and synthesized to a single view compensating each defects each other. To be concrete, near view including roads, guardrails, other vehicles is represented by geometry-based model and far view including buildings and sky is represented by image-based model.

This paper is composed of five sections. In the next section the overview of whole system is described. The third and forth section describes a method to acquire the source video data and method to synthesize views. And final section summarizes the paper.

II. System Overview

A. Whole System

Mixed-Reality Traffic Experiment Spaceis composed by extending conventional frameworks of traffic/driving simulators and by integrating several modules as listed below and shown in Fig. 2.

- TS: Macroscopic traffic simulator [2]
- KAKUMO: Microscopic traffic simulator [2], [3]
- DS: Driving Simulator [3]
- IMG: Image Generator

TS is a module to simulate macroscopic traffic flows with traffic volume parameters and road network model composed of node- and link-based graph structure.

DS is a module to recreate microscopic behavior of self vehicle from user's handling, acceleration, braking operation and vehicle-dynamics model. The behavior of the vehicle is transmitted to the user through the seat.

KAKUMO is a module to simulate microscopic position of each vehicles on a road with macroscopic traffic flow provided as an output of TS. Each vehicle changes its position(lane), heading and velocity according to relative position and velocity between surrounding vehicles.

IMG is a module to produce driver's surrounding view in real-time from the position and pose of self/surrounding vehicles. The detail is described in the following.

With this system configuration, user can experience a driving situation inside the network of TS.

This Mixed-Reality Traffic Experiment Spacecan virtually deal with existing public roads as a simulation scene. Our current prototype system targets a part of Tokyo Metropolitan Expressway[7] No.3 and Loop Line No.1, from Shibuya to Miyakezaka, about seven kilometers long.

B. IMG: The image generating module

This section describes the detail of IMG module. The view which IMG offers to the user is nothing less than a view of virtual urban model from ground level, and approaches to represent such models are classified into geometry-based one and image-base one as described in the first section.

Considering the aptitude of each method in the whole system, geometry-based rendering is superior from the perspective of computing cost and interactions with other objects such as other vehicles, and image-based rendering is superior from the perspective of photo-reality provided to the users. Therefore we propose a novel approach to use each approach according to each objective and synthesize each views at displaying stage. In concrete, each approach handles near part and far part of view respectively as listed below and shown in Fig.3.

- Near-view part: Geometry-based
- Far-view part : Image-based

Near-view part includes roads, guardrails, soundproof walls, traffic signs, signals, other vehicles, pedestrians, etc. This part is rendered by using conventional techniques implemented on DS module, a product of Mitsubishi Precision Co. Ltd.[8] This module can represent behaviors of each vehicles at the rate of 60Hz.

Far-view part includes surrounding buildings and a sky. This part is rendered by processing an video image database,

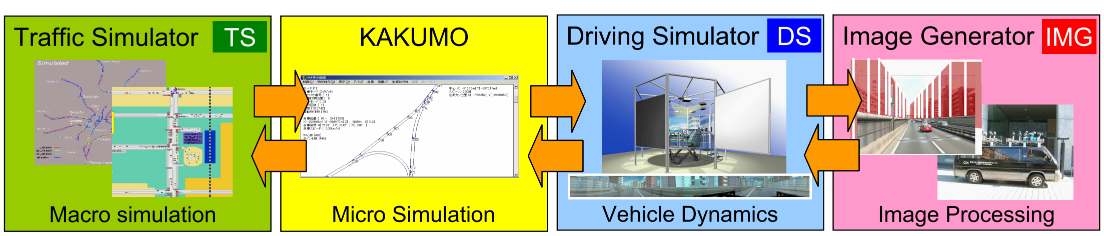

Fig. 2. Mixed-Reality Traffic Experiment Space

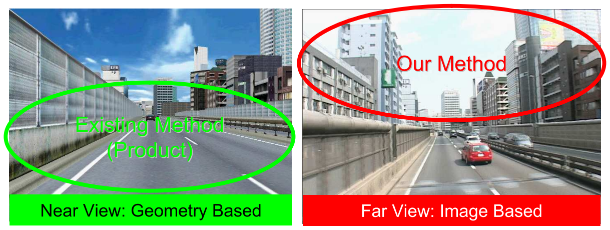

Fig. 3. Hybrid model expression: Near-view is represented by geometry-based model, and far-view is represented by image-based model.

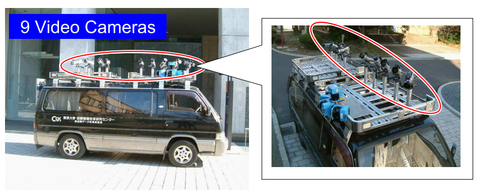

Fig. 4. Data acquisition vehicle

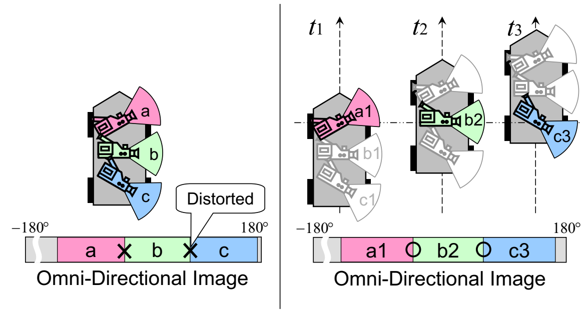

Fig. 5. Spatio-temporal agreement of optical centers

(a) Before mosaicing

(b) After mosaicing

Fig. 6. Omni-directional image (Half-directional in this case)

which is constructed by running along the model course by capturing vehicle in advance. By capturing video images in an omni-directional format, a view from outside the trajectory of capturing vehicle can be reconstructed by image processing. This process is dynamically carried out in real-time according to the position and pose of self vehicle given from DS module.

By the proposed hybrid method, a view exploiting each advantage can be created. Additionally, construction process of geometric model usually with huge manpower is required only in near-view part, leading to less developing cost.

III. Real-world Capturing for View Source

Surrounding view for the user is produced by processing real video images captured by data acquisition vehicle, running along the targeting road, Tokyo Metropolitan Expressway in the prototype system. Fig.4 shows our data acquisition vehicle. Nine video cameras are equipped on the roof and omni-directional video image is created by mosaicing each images captured by these cameras. As the following part describes in detail, once omni-directional images viewed from running path is accumulated, a view from outside the path can be created through image processing, therefore capturing travel is carried out only once.

It is well known that mosaiced image synthesized from multiple camera image includes distortions at joint parts if optical centers of each camera are not coincided into one point as this case (Fig.5 left). We use Kawasaki's spatio-temporal optical synchronization method[4] for solving this problem. By arranging each cameras parallel to moving direction, optical centers are coincided into one point at different timing each (Fig.5 right): As for camera n, at time t_n. Fig.6 shows an example of omni-directional image created by mosaicing multiple video camera images.

IV. Reconstruction of View from User Viewpoint

A. Basic concept of synthesizing arbitrary viewpoint image

A set of omni-directional images captured along running path of data acquisition vehicle enables to create a view from outside the path by stitching parts of omni-directional images[5]. In Fig.7 for example, a left-part view from a star signed point is composed of forward-left, left, and backward-left part of omni-directional image captured at time t_1, t_2, t_3 respectively. Also, right-part view can be created by copying and stitching right-directional rays(counter direction of three arrows in Fig. 7–Rendering) captured at each time.

Since the road can not be regarded as straight in practice, we divide the road into a series of line segment and copied the nearest ray from omni-directional image captured on the segment.

This process is actually implemented as a kind of texture mapping onto virtual walls assumed along the roadside, a boundary zone of near-view and far-view part. The face of wall

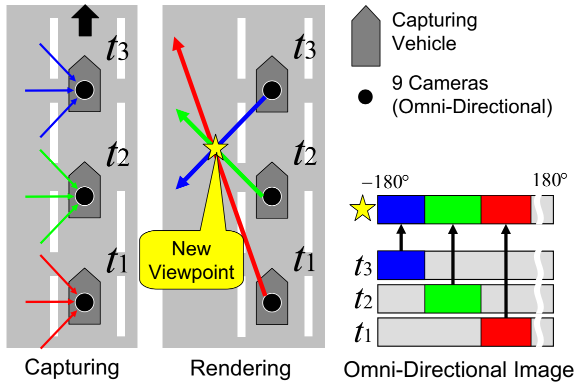

Fig. 7. View synthesis from new viewpoint by using omni-directional image

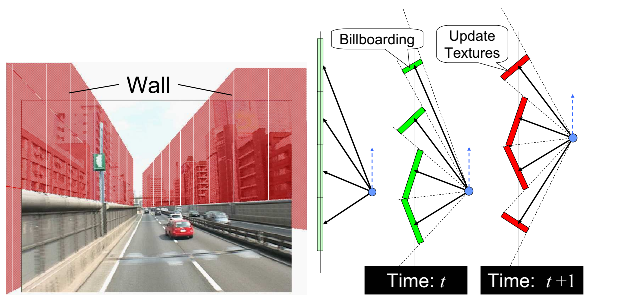

Fig. 8. Virtual walls along roadside. Textures on wall slits are updated according to the position of viewpoint. Slits are rotated to visual line direction of the user.

is divided into some vertical slits and a part of omni-directional images captured from appropriate points are mapped per each slit. Textures are dynamically updated according the position of self vehicle given from DS module. The slits are rotated to the visual line direction of the user (Fig.8).

The method to synthesize arbitrary-viewpoint image proposed in [5] can not deal with front view. As shown in Fig. 9 left, the front-directional ray can be copied from nowhere, i.e., ray parallel to the required ray have never been captured. In such case, we just compensate front-directional ray from the nearest capturing point. In our case practically, we applied this way not for completely front-directional ray but also the ray whose directional difference is within $\pm 30^\circ$ from front.

This compensation process inevitably brings error compared with appropriate view. This error is just same as a parallax in the field of stereo vision, i.e., farther part will include smaller error though, nearer part will include larger error in rendering phase. However, as described in Sec. II, such nearer part is not rendered by image base but by geometry base. This will lead the view more natural even in the case of front view.

B. Speeding up and quality improvement

Textures are to be updated every time the position of self vehicle changes inside DS module. The refreshment rate of vehicle position is 60Hz at DS of conventional geometry-based rendering and 20Hz at KAKUMO output, a simulation result of surrounding vehicles. However, the time loss becomes con-

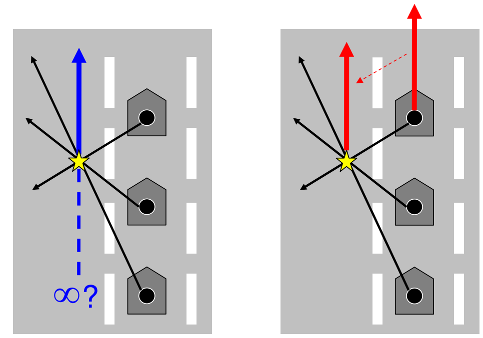

Fig. 9. Exception process for front view

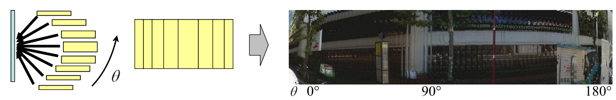

Fig. 10. Texture dataset

siderably large if the process described in previous subsection is carried out and read/write on graphics hardware occurs every time. For speeding up, appearances of a slit from all directions are retained on a graphic memory as a dataset (Fig.10) per each slit and called according to need. Since it is impossible to load all datasets to a graphic memory along targeting road, texture generation task is allocated to multiple machines per some short running regions. A machine which finished rendering its allocated region is allocated a new region and prefetches the dataset of new region.

Though datasets in slit surface include appearances of itself from all direction, they can be obtained only discretely. This is because a frame rate of video camera is finite and therefore positions of omni-directional images exist discretely on the capturing path. When an appearance from direction between two directions retained in a dataset is required, each textures are complemented by alpha-blending and the quality of rendering is improved.

C. Synthesis of image-based part and geometry-based part

As a pre-processing of synthesizing image-based part and geometry-based part, correspondence of each part must be clarified. In concrete terms, look-up table between each frame of omni-directional video image and the position where it is captured in the coordinate of geometry-based model is required. It is possible to get a correspondence of capturing points to the real world in some measure, however, geometry model in ready-made product of driving simulator does not strictly reflect the real world. Therefore, we first created a series of geometry-based view with the same FOV as omni-directional image, and manually indexed two views at several characteristic points. And then we interpolated residual points by using cubic spline interpolation (Fig. 11).

Image-based part and geometry-based parts are synthesized into one view by using special hardware device. The hardware configuration of IMG module is shown in

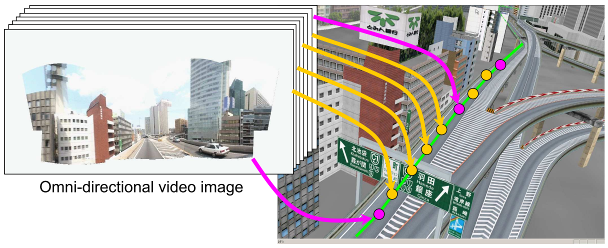

Fig. 11. Indexing of sampling points of omni-directional images and coordinates in geometric model.

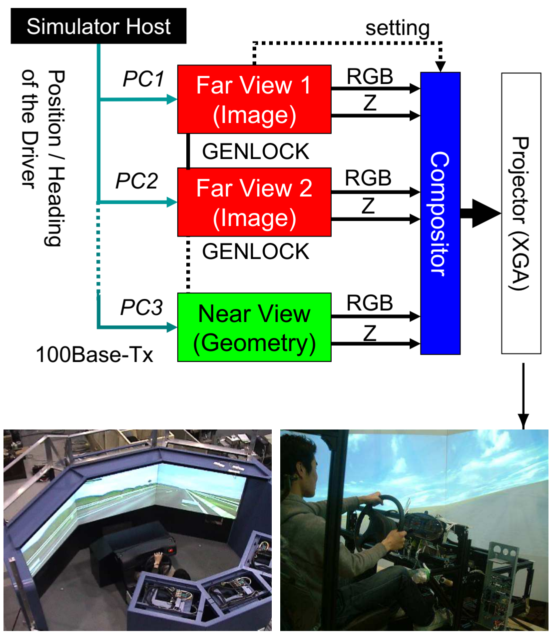

Fig. 12. Hardware configuration of IMG module

Fig.12. Far-view(image-base) rendering machines and near-view(geometry-base) rendering machine outputs both view from same viewpoints with color values (R, G, B) and depth value (Z). These outputs are integrated by a hardware called Compositor (VizCluster), a product of Mitsubishi Precision Co. Ltd., and color values per each pixel are determined according to the depth value of each output. And finally, it is projected to a screen in front of the user through multiple projectors.

V. Rendering Result

Fig. 13(a) shows the example of rendering result of image-based part. Its viewpoint is approximately same as Fig. 1, fully geometry-based one. It can offer relatively highly photo-realistic view to the user. We confirmed that the update of frame worked well in 60Hz.

Fig. 13(b)(c) are the examples of viewpoint changing effect. Since the part near from front direction in each figures are rendered by copying irregular rays as described in Sec. IV, positions of white lines on the road for example are not appropriate. However, the change of depth to side buildings are effectively expressed.

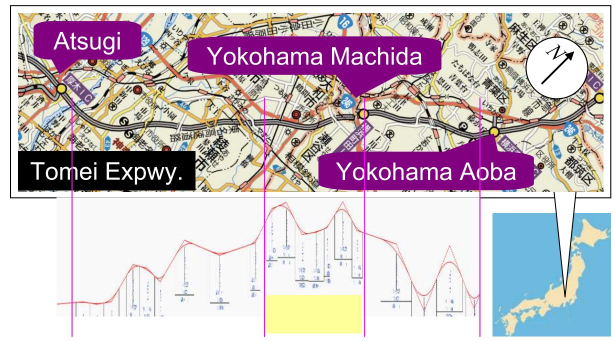

Fig. 15. Sag zone in Tomei Expressway

Fig. 14 shows the synthesizing result of image-based part and geometry-based part through the compositor. Sky, buildings, advertising displays on buildings are rendered in image-base, and roads, soundproof walls, other vehicles, traffic signs, which are located in front of a virtual wall of image-based rendering, are rendered in geometry-base.By such way of right-method-in-the-right-place rendering, we can offer highly photo-realistic view for a driver, and simultaneously can realize interactive use as changing behaviors of other vehicles or changing the contents of traffic signs.

VI. Conclusion

In this paper, we introduced an interactive driving-view generation system, a module of Mixed-Reality Traffic Experiment Space, which we have been developing as one sphere of “Sustainable ITS Project”. Our system synthesize both geometry-based view and image-based view according to their own aptitudes and can offer photo-realistic and utilizable view.

In the image-based part, an arbitrary view can be rendered by processing omni-directional video image which are obtained by capturing and running on the model course in the real world only once. Geometry-based part is composed by existing product which required huge human cost for model construction hitherto, however, in this case, the model construction process is needed only in roadway part.

For the future, we are planning to apply this system for driving simulation in sag zones in highways. A sag zone is composed of a series of subtle gradient changes and is said to psychologically causes traffic congestion. Our photo-realistic system will just be suited for realizing such subtle situation. Currently we are targeting a region of Tomei Expressway, from Yokohama-Aoba I.C. to Atsugi I.C. (Fig. 15), and going to analyze human parameters around driving behavior or judgment, and verify effectiveness of sag-zone traffic signs depending on its contents, dimension, and location.

Acknowledgment

Shintaro Ono is supported by Research Fellowships of Japan Society for the Promotion of Science (JSPS) for Young Scientists.

(a) A view from capturing lane, gazing at front

(b) A view from right side of capturing lane, gazing at left

(c) A view from left side of capturing lane, gazing at right

Fig. 13. Rendering result of image-based part and lane-changing effect

Fig. 14. Synthesized rendering result of image-based part and geometry-based part

REFERENCES

[1] K. Ikeuchi et.al, “Mixed Reality Traffic Experiment Space under Interactive Traffic Environment for ITS Research”, Proc. 11th World Congress on Intelligent Transport Systems and Services (ITSWC 2004), Nagoya, Japan, Oct. 2004

[2] T. Shiraishi et.al, “Development of a Microscopic Traffic Simulation Model for Interactive Traffic Environment”, Proc. 11th World Congress on Intelligent Transport Systems and Services (ITSWC 2004), Nagoya, Japan, Oct. 2004

[3] Y. Suda, M. Onuki, T. Hirasawa, H. Ishikawa, M. Kano, Y. Mashiyama, T. Oda, A. Tagaya, T. Taguchi, Y. Kanki, “Developnemt of driver model using driving simulator with interactive traffic environment”, Proc. 11th World Congress on Intelligent Transport Systems and Services (ITSWC 2004), Nagoya, Japan, Oct. 2004

[4] H. Kawasaki, A. Miyamoto, Y. Ohsawa, S. Ono, K. Ikeuchi, “Multiple video camera calibration using EPI for city modeling”, Proc. Asian Conference on Computer Vision (ACCV 2004), Jeju Island, Korea, Jan. 2004

[5] T. Takahashi, H. Kawasaki, K. Ikeuchi, M. Sakauchi, “Arbitrary view position and direction rendering for large-scale scenes”, Proc. Computer Vision and Pattern Recognition (CVPR 2000), Hilton Head Island, US, Jun. 2000

[6] Sustainable ITS Project, Center for Collaborative Research, The University of Tokyo, `http://www.its.ccr.u-tokyo.ac.jp`

[7] Metropolitan Expressway Public Corporation, Tokyo Metropolitan Expressway, `http://www.mex.go.jp/english`

[8] Mitsubishi Precision Co., Ltd., `http://www.mpcnet.co.jp/e`