IMAGE GENERATION SYSTEM FOR MIXED-REALITY TRAFFIC EXPERIMENT SPACE

Shintaro Ono* Koichi Ogawara*1 Hiroshi Kawasaki*2 Masataka Kagesawa*1
Masaaki Onuki*3 Ken Honda*3 Keiichi Kenmotsu*3 Mayumi Sakai*3
Motomu Tsuji*3 Katsushi Ikeuchi*1

* Ph.D course student; Dept. of Information and Communication Eng.,
Graduate School of Information Science and Technology, The Univ. of Tokyo
Ee-405, Ikeuchi Lab., Intstitute of Industrial Science,
4–6–1 Komaba, Meguro-ku, Tokyo, 153–8505, JAPAN
Tel: +81–3–5452–6242 Fax: +81–3–5452–6244 onoshin@cvl.iis.u-tokyo.ac.jp

*1 Institute of Industrial Science, The University of Tokyo
*2 Faculty of Engineering, Saitama University
*3 Center for Collaborative Research, The University of Tokyo

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.

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. 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 acknowledgement and decision parameters in human driving operation by using this simulation system. However, image-based approach is not appropriate for 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.

SYSTEM OVERVIEW

WHOLE SYSTEM

Mixed-Reality Traffic Experiment System is composed by extending conventional frameworks of traffic/driving simulators and by integrating several modules as listed below and shown in Fig.1.

- 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 Space can 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.

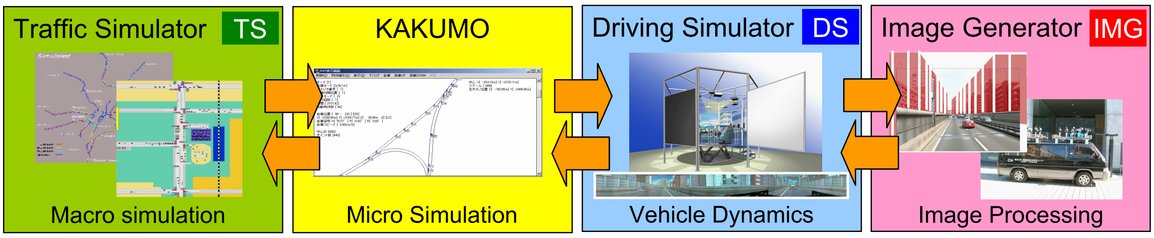

Figure 1: Mixed-Reality Traffic Experiment Space

IMG: IMAGE GENERATOR — A REALISTIC VIEW FOR DRIVER

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.2.

- 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. 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, 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.

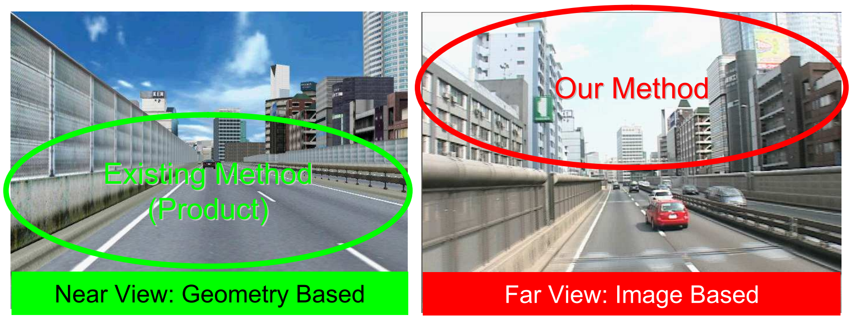

Figure 2: Hybrid model expression: Near-view is represented by geometry-based model, and far-view is represented by image-based model.

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.

REAL-WORLD CAPURING 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.3 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.

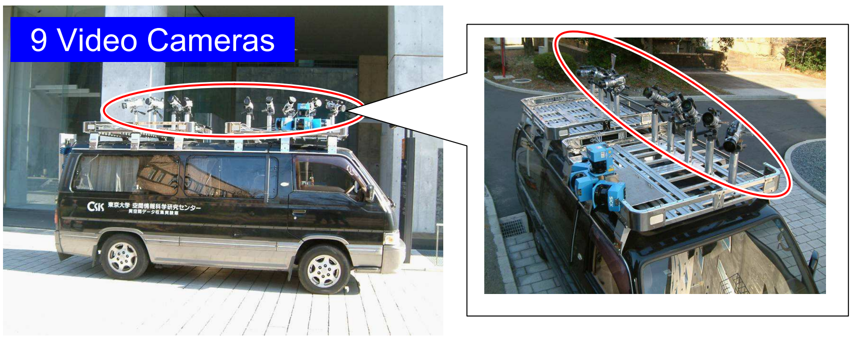

Figure 3: Data acquisition vehicle

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.4 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.4 right): As for camera n, at time t_n.

Fig.5 shows an example of omni-directional image created by mosaicing multiple video camera images.

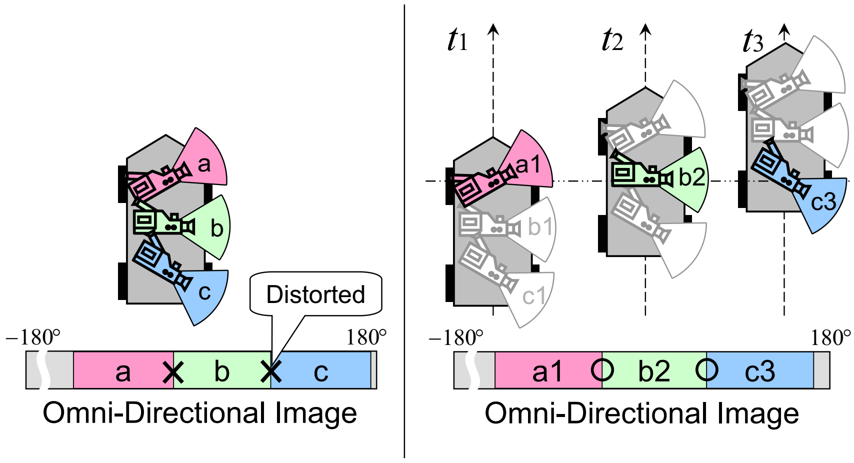

Figure 4: Spatio-temporal agreement of optical centers

Figure 5: Omni-directional image

RECONSTRUCTION OF VIEW FROM USER VIEWPOINT

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.6 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.

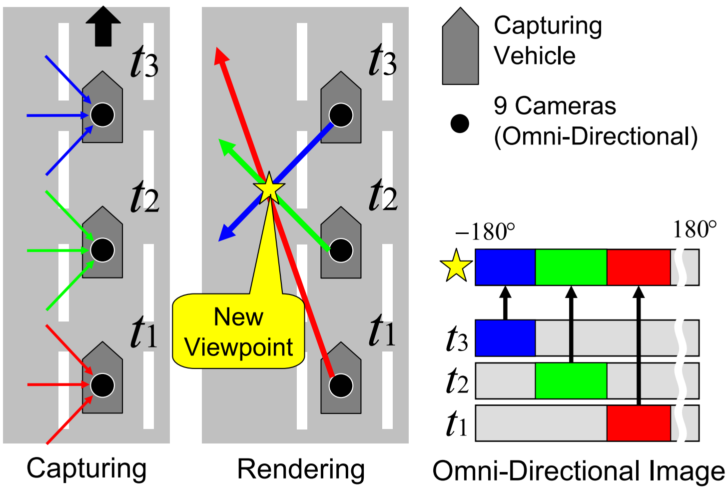

Figure 6: View synthesis from new viewpoint by using omni-directional image

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 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).

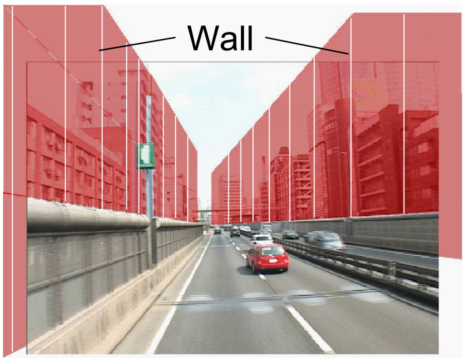

Figure 7: Virtual walls along roadside

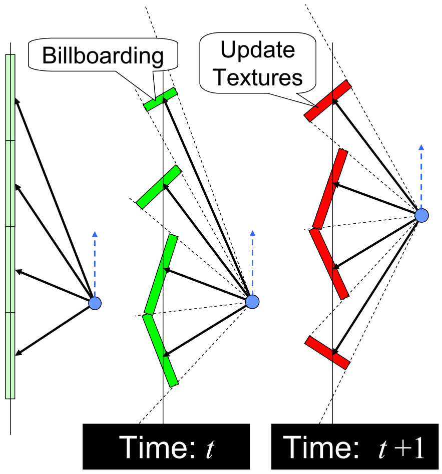

Figure 8: Textures on wall slits are updated according to the position of viewpoint. Slits are rotated to visual line direction of the user.

SPEEDING UP TECHNOLOGY

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 considerably 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.9) per each slit and called according to need.

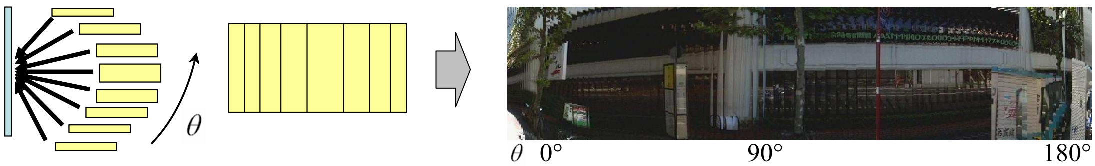

Figure 9: Texture dataset

QUALITY IMPROVEMENT

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.

SYSTEM OUTPUT

The hardware configuration of IMG module is shown in Fig.10. 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. Far-view machines and near-view machine outputs both color values (R, G, B) and depth value (Z). These outputs are integrated and color values per each pixel are determined according to the depth value of each output by a hardware called Compositor, a product of Mitsubishi Precision Co. Ltd., and projected to a screen in front of the user through multiple projectors.

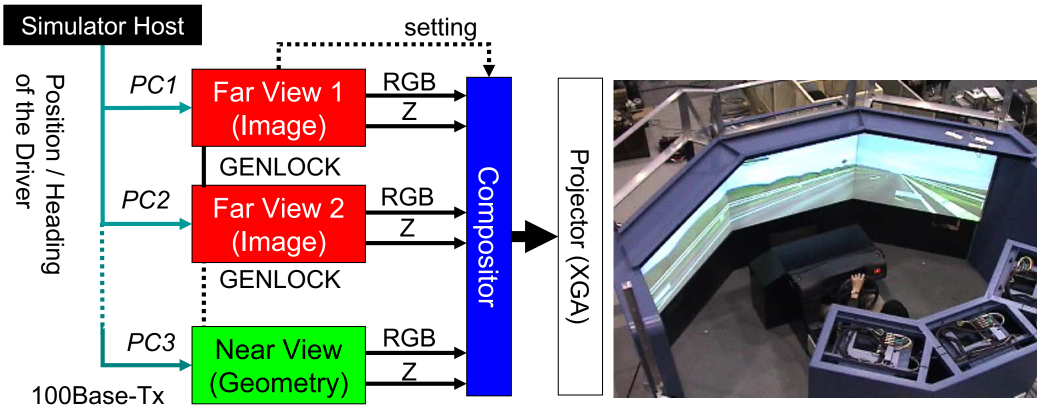

Figure 10: Hardware configuration of IMG module

CURRENT STATUS OF DEVELOPMENT

Currently we have confirmed an effectiveness of rendering algorithm of far-view part. Examples of synthesizing arbitrary viewpoint images are shown in Fig.11. Subsequently we are testing rendering process using omni-directional video images captured at Tokyo Metropolitan Expressway, coordinative distributed processing of multiple machines, overlaying with near-view part, optimization of the structure of texture mapping face.

Figure 11: Examples of arbitrary viewpoint images

SUMMARY

This paper introduced a hybrid-based generation system of user's view for Mixed-Reality Traffic Experiment Space, which we have been developing as a part of Sustainable ITS Project.

This system simultaneously offers photo-realistic and interactive view to users by allocating view image-based and geometry-based view for far-view part and near-view part respectively according to the appropriateness of each approach.

Image-based part can be rendered from arbitrary viewpoint using real video images captured as a form of omni-directional image by traveling along targeting road with video cameras in advance only once.

Geometry-based part, though requires huge human cost for reconstruction, is needed only for near-view part, leading to less development cost than conventional full geometry-based approach.

In the 11th World Congress on Intelligent Transport Systems and Services Nagoya Aichi Japan, we show off the Mixed-Reality Traffic Experiment Space at the demonstration site and have our systems experienced.

References

[1] M. Kagesawa et.al, "Sustainable ITS Project", 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`