Realtime Mixed Reality Representation with a Virtual Light Source based on a Mobile 3D Acquisition

Yoji Watatani† Yoshihiro Yasumuro‡ Hiroshige Dan‡ Masahiko Fuyuki‡

† Graduate School of Science and Engineering,
‡ Faculty of Environmental and Urban Engineering,
Kansai University,
3-3-35, Yamate, Suita, Osaka 564-8680, Japan
{k974859, yasumuro, Dan, fuyuki}@kansai-u.ac.jp

Abstract— Mixed reality (MR) has gathered attention recently as an effective technique for overlaying computer-generated virtual objects on physical scenes. Using MR, this research proposes a realtime imaging system to produce visual illumination effects on physical objects with a virtual light source. The proposed system models the shapes and the color information of a real scene through a realtime process. The illumination influenced by the virtual light source on the scene model are superposed on an actual video image to create MR representation. Experimental results show virtual light-up effects on the physical shapes and colors of the real objects by setting up a non-existing lighting configurations.

I. Introduction

Coordinating and reforming interiors in existing spaces has received broad and continuing interest. The planning of color coordinates and organizing interior items often require an expert's knowledge and ideas. The presentation of the designing plan is also important for communication between the expert and a customer. Designers usually paint perspective figures and use technical drawings to convey concepts of their plans.

Digital presentations using computer aided design (CAD) and computer graphics (CG) software are other powerful tools for visualizing designs in realistic ways. Mixed reality (MR) or augmented reality (AR) techniques are good choices for focusing more closely on the consistency between newly organized design plans and existing real objects [1]. MR is a visual synthesis technique to overlay digital objects of CG on real environmental images with visual consistencies regarding geometries and colors between virtual and real objects. Taking into account the real lighting conditions for rendering CG objects, a number of MR techniques have already been proposed [2, 3]. By sensing real lighting conditions, adjusting the illumination for the CG objects is comparatively easy; however, changing the representation of the real objects by adding a virtual light source, which does not physically exist, is technically difficult when the light equipment is also a major interior item. Therefore, such lighting simulation is basically possible only in a full-CG environment by modeling both existing and candidate objects.

II. Basic Idea

With the recent development of hand-held devices, 3D confiurations can be easily acquired using small-sized imaging or ranging sensors. Using these kinds of devices, this paper shows their feasibility in applying the demands of MR representation for virtual lighting in existing spaces. By utilizing a realtime 3D shape measurement technique, constructing an instant digital 3D copy of the real environment allows lighting simulation on the spot. This research proposes a MR representation system for visualizing virtual illumination in realtime, using a hand-held-sized time-of-flight (TOF) sensor.

Another aspect of this approach is to utilize multi-band optical sensing, which is a type of visible and invisible color processing. Invisible processing using an infra-red (IR) band can capture the 3D information without interfering with the user's view. A color-imaging sensor can capture the appearance of the visual color. In their MR system, Yasumuro et al. also proposed using a combination of an IR-projector and a video camera [5], which has a comparatively large device setup. Kinect by Microsoft [4] is a packaged compact device that employs a similar technique using a coded IR-projection pattern and a color camera to realize fast recognition for a gaming control user-interface. For this coded projection, an invisible code-embedding technique in normal color projection for AR/MR use has been proposed by Park [6].

This paper shows the potential for newly expanded MR usability and its required minimum configuration by a realtime mobile system, which will further development of multi-band optical sensing packages.

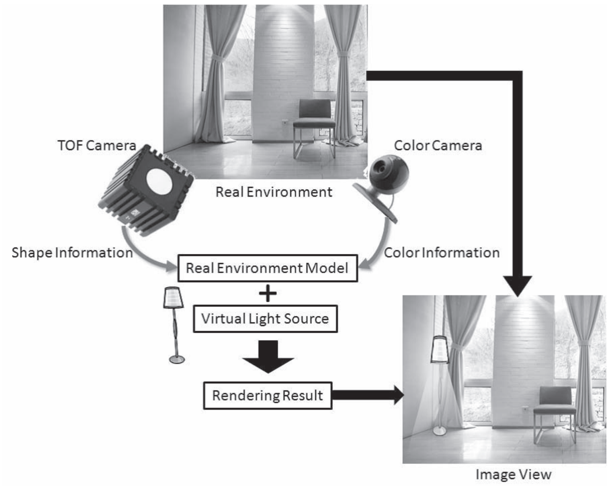

Fig. 1. system image

III. Proposed Method

Our method is a straightforward approach to making the most use of the local illumination model. The local illumination model is capable of representing lighting effects for CG with only directly related elements such as light source configuration, object surface shapes, and a reflection configuration. We instantiate digital colored-shape copies of a real space in realtime, on which the local illumination from virtual lighting sources are applied. The local illumination method computes the color information in an additive color-mixing manner to render the objects. We superimpose the rendered results onto the real scenery images to add virtual light effects in the same manner.

To create the colored-shape models of the real environment, we acquire the shape information by a TOF camera and the color information by a color camera simultaneously from the targeted physical space. The environmental model is generated by integrating the shape and the color information in realtime.

A. *TOF (Time of Flight) camera*

A TOF camera acquires the information of range to the object surface with each pixel. Individual pixel has its own photo-sensor to measure the travel time for the intensity modulated IR rays which are emitted from the camera and reflected back from the objects [7]. Since each pixel independently processes the range in parallel, we can obtain a 3D configuration of a range image as a ranging-pixel set. Eventually, the output from the TOF camera will be a realtime stream of range-image frames. The ranging is performed by searching the phase-shift of a modulated signal pattern in the reflected light. The detection of the phase-shift by wave pattern matching may cause flickering in raw range images. Taking advantage of the fast capture rate of the TOF camera, we apply a moving averaging filter over time and a mean filter for the captured raw data frames. Regardless of these preprocesses, the live video frame rate is still maintained for the range image stream. However, the TOF camera can obtain the luminance and 3D information of the object, but the color information cannot. We need color information to be added to the 3D information to create the environment model of the target. Thus, the combination of a color camera and 3D information, create the environmental model in the manner described in B. Calibration by the obtained 3D information and color information.

Fig. 2. real environment

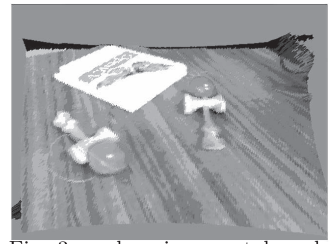

Fig. 3. real environmental model

B. *Calibration*

To integrate the shape and the color information in realtime, the color camera and the TOF camera are initially calibrated and their geometrical relation is fixed before starting the system. A projection matrix can be determined to transform the 2D coordinate of the color camera image to the 3D coordinate data space of the TOF camera to capture an identical physical target space. The projection matrix is found by giving a number of identical target points, whose colors and positions can be identified by both of the two different types of cameras simultaneously. The projection of TOF camera data (X, Y, Z) in 3D space onto the points (x, y) on the 2D image can be written as Equation (1).

$$\begin{bmatrix} hx \\ hy \\ h \\ 1 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{1}$$

To determine the projection matrix C, the least square method can be used to solve Equation (1). From Equation (1), the following two expressions can also be derived to refer to the color information from the 2D image coordinates corresponding to the 3D coordinates.

$$x = \frac{c_{11}X + c_{12}Y + c_{13}Z + c_{14}}{c_{31}X + c_{32}Y + c_{33}Z + c_{34}} \tag{2}$$

$$y = \frac{c_{21}X + c_{22}Y + c_{23}Z + c_{24}}{c_{31}X + c_{32}Y + c_{33}Z + c_{34}} \tag{3}$$

In this paper, the coordinates (x, y) obtained from Equations (2) and (3) sometimes do not come out well on the image coordinates of the coordinate values due to the difference of camera resolutions and noises. In that

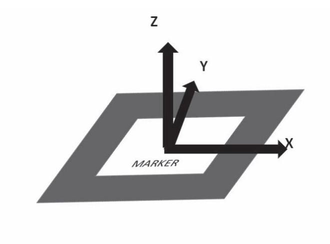

Fig. 4. marker coordinate system

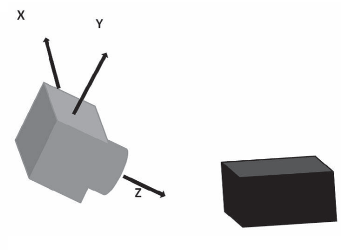

Fig. 5. camera coordinate system

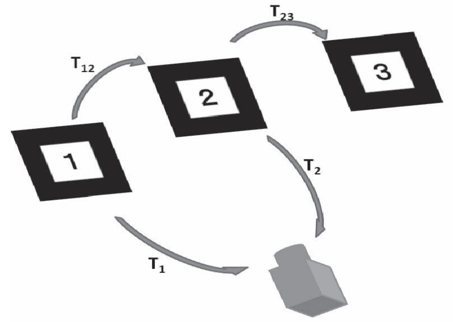

Fig. 6. multi-marker

case, we coded our program to return an error value for exception handling.

C. Superimposing method

To superimpose the environmental model on the actual video image, we use a visible marker to specify a coordinate in physical space. Virtual light source equipment is placed at an arbitrary position relative to the marker. 3D CG objects can be easily placed in physical space by calibrating the provided marker coordinates and the color camera coordinates, similar to the way the ARToolKit library performs for AR/MR representation [8]. The object coordinate system is matched to the marker coordinate system whose origin is at the center of the marker (Fig. 4). The TOF camera has its own coordinate system with its origin at the camera's center (Fig. 5). Using Equation (4), the TOF camera coordinate origin is moved to the marker coordinate and a rotation is performed to place the environmental model properly in respect to the marker. The coordinate system scale is aligned to that of the TOF camera.

$$\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} = \begin{bmatrix} X_e & Y_e & Z_e \end{bmatrix} \begin{bmatrix} X_c - M_x \\ Y_c - M_y \\ Z_c - M_z \end{bmatrix}, \quad (4)$$

where

$[X_m \; Y_m \; Z_m]^T$: *marker coordinate system*

$[X_c \; Y_c \; Z_c]^T$: *camera coordinate system*

X_e, Y_e, Z_e : *basis vectors*

D. Multi-marker system

In this paper, not only a single marker but also additional multiple markers are used for arranging and superimposing the environmental model and virtual light source to allow the camera to move around in the space. In particular, the virtually settled light source may be out of the camera's view, but it still needs to be fixed in a certain position on a ceiling or wall, for instance, to light up all the items in the camera's view. Since the virtual light source is placed based on the marker coordinate in our system, a marker-based compensation of the camera motion is required for properly maintaining the light source position in physical space.

Fig. 7. TOF camera SR4000

Fig. 8. Color camera Webcam C200

We get the coordinate transformation matrix $T_i (i = 1, 2, 3, \ldots)$ from the marker i to the camera as shown in Figure 6. Then the transformation matrix $T_{i,i+1}$ between adjacent markers is expressed as follows.

$$T_{i,i+1} = {T_{i+1}}^{-1} T_i \quad (5)$$

Thus, once the virtual light source is superimposed on the marker 1, even if marker 1 passes from the camera view, known relative positions of the markers can be used and the virtual light source can maintained to be superimposed on the marker 1.

IV. Experiment

We implemented the proposed system and conducted experiments. The devices used for the system are listed as follows.

- TOF camera : MESA SR4000[9]
  Maximum Frame Rate : 50fps (Camera setting dependent)
  Pixel Array Size : 176(h)×144(w)
- Color camera : Logicool Webcam C200
  Maximum Frame Rate : 30fps
  Video Capture : VGA (640×480)

Also, we used the library of ARToolKit for calibration of marker coordinates and the color camera coordinates. We employed the illumination process using specular and

Fig. 9. experiment environment

diffusion reflection parameters that were expected from the color information for the local illumination model.

The following shows the procedure of the experiment. Step 1 is only for the initial camera system setup.

Step 1. Calibrate the two types of cameras
Step 2. Acquire the real environmental data
Step 3. Construct the environmental model
Step 4. Render the environmental model by virtual light settings
Step 5. Superimpose the environmental model on live video frames

A. Time-Moving Average and Mean Filter

We reduced the temporal fluctuations on the TOF camera data by a moving-average for every 20 consecutive frames for each pixel. Then, a 3×3 median value filter was applied to reduce the discontinuity on the surface to form the polygonal mesh of the environmental model.

B. Result

We implemented an experimental system to demonstrate a virtual red spotlight to illuminate the real scene as shown in Figure 10. The results of the superimposed virtual illumination effect are shown in Figure 11. The results represent a state of virtual illumination with a realistic shading that reflects the shapes and colors in the real environment. Also, an interactive trial in realtime, such as moving the virtual light positions and the physical objects in the scene as well as the viewpoint was possible.

V. Conclusion

This paper proposed an interactive MR system framework to visualize virtual illumination effects on physical objects. The experimental results showed a potential for realizing the instant representation of virtual lighting by non-existent light source in realtime. The output quality depends on the resolution and noise-reduction process for the raw data of the sensing modules, particularly 3D information from the TOF camera.

Our future work includes improving the quality of the shape and reflecting properties of the scenery CG model and the light modeling by considering the properties of

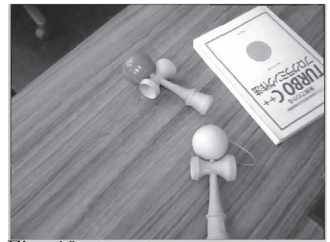
Fig. 10. target scene

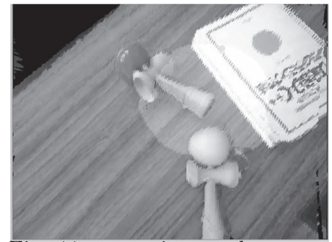
Fig. 11. execution result

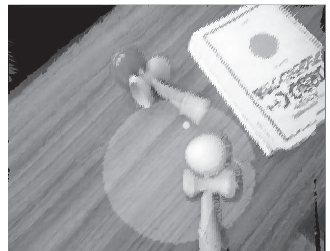
Fig. 12. different angle


Fig. 13. virtual spotlight only

the real illumination equipment. We are also planning to virtually simulate the visibility of the emergency lights in underground constructions as part of disaster-prevention efforts as well.

Acknowledgements

This is a product of research which was financially supported by the Kansai University Research Grants: Grand-in-Aid for Joint Research, 2010-2011.

References

[1] Klein, G., Murray, D., “Parallel tracking and mapping for small ar workspaces,” In Proc. of ISMAR 2007, pp. 225-234, 2007.

[2] Paul Debevec, “Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography”, prod. of SIGGRAPH ’98,pp. 189-198, 1998.

[3] Y. Yasumuro, et al. “Consistent Presentation of Interactive Virtual Objects in Real Space with 3D Markers,” Proceedings of the VIIth@Digital Image Computing: Techniques and Applications (DICTA2003), pp.653-662, 2003.

[4] http://www.xbox.com/en-US/Kinect (last viewed in Oct. 2011)

[5] Yoshihiro Yasumuro, Masataka Imura, Yoshitsugu Manabe, Osamu Oshiro, Kunihiro Chihara: “Projection-based augmented reality with automated shape scanning,” SPIE-IS&T, Vol.5664 pp.555-562, 2005.

[6] H. Park , M.-H. Lee , S.-J. Kim and J.-I. Park ”Surface-independent direct-projected augmented reality”, Proc. ACCV06, p.892 , 2006.

[7] Oggier T., et al. ”Swissranger sr3000 and first experiences based on miniaturized 3DToF cameras,” In Proc. of the First Range Imaging Research Day at ETH Zurich, 2005.

[8] “ARToolKit”http://www.hitl.washington.edu/artoolkit/

[9] ”SR4000 User Manual,” http://www.mesa-imaging.ch/ (last viewed in Oct. 2011)