2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Simultaneous Projection Mapping Using High-frame-rate Depth Vision

Jun Chen, Takashi Yamamoto, Tadayoshi Aoyama, Takeshi Takaki, and Idaku Ishii

Abstract— **In this paper, we report on the development of a projection mapping system that can project RGB light patterns that are enhanced for three-dimensional (3-D) scenes using a GPU-based high-frame-rate (HFR) vision system synchronized with HFR projectors. Our system can acquire 512×512 depth images in real time at 500 fps. The depth image processing is accelerated by installing a GPU board for parallel processing of a gray-code structured light method using infrared (IR) light patterns projected from an IR projector. Using the computed depth images, suitable RGB light patterns to be projected are generated in real time for enhanced application tasks. They are projected from an RGB projector as augmented information onto a 3-D scene with pixel-wise correspondence even when the 3-D scene is time-varied. Experimental results obtained from enhanced application tasks for time-varying 3-D scenes such as (1) depth-based color mapping and (2) augmented reality (AR) spirit level, confirm the efficacy of our system.**

I. Introduction

Advances in video technology have led to augmented reality (AR) [1] being proposed for realistically overlaying virtual information on real world material, and many real-time video-based AR systems have been developed for human interaction using marker-based tracking [2], [3] and model-based tracking [4], [5]. Most of these systems are designed for real-time visualization of synthesized images via display devices such as head-mount and handheld displays; and images are not directly displayed on the real environment to make it indeed augmented. Due to such inconsistencies between the real environment and images on computer displays, presenting realistic augmented information to operators in real environments is difficult.

Projection mapping [6], [7] is a well-known spatial augmented reality technology for projecting computer-generated light patterns from projectors onto the real environment. It involves complex-shaped or non-monotoned objects and can turn the real environment into a virtual display surface by generating projection light patterns that are matched with the geometry of the real environment. Many illusion works based on projection mapping, such as Virtual Showcase [8], projector-guided painting [9], and architectural projection mapping [10], have been reported. Most of them have only been conducted for static 3-D scenes because they assume that the geometry of the real environment has been measured prior using offline depth sensing such as laser-scan sensing,

Recently, various types of real-time depth sensors operating at video rates, such as Kinect [11], have been developed, and researchers have proposed interactive projection mapping systems such as Beamatron [12] and IllumiRoom [13], which require real-time video processing. However, drawbacks such as delay times on the order of dozens of milliseconds and non-pixel-wise space resolution in 3-D sensing and projection when observing fast human interaction still remain. Okumura et al. [14] developed a galvano-mirror-tracking system for AR projection, and demonstrated the effectiveness of high-frame-rate (HFR) video tracking in projection matching; however, they assume the real environment to be flat due to no pixel-wise acquisition of depth image. To process high-speed 3-D motion, we developed a real-time 500 fps depth vision system with pixel-wise accuracy [15]. We believe that if this HFR depth vision system can be applied to projection mapping, then its enhanced real environment would greatly assist with human tasks.

This paper reports on the development of an HFR camera-projector system that can acquire and process 512×512 depth images in real time at 500 fps and project computer generated light patterns onto time-varying 3-D scenes. The depth image processing is accelerated by installing a GPU board to process the 8-bit gray-code structured light method using infrared (IR) light patterns projected at 1000 fps. RGB light patterns are interactively generated so that the patterns are projected onto the 3-D scene and enhanced with pixel-wise correspondence. Experimental results obtained for several enhanced application tasks verify its efficacy.

II. Real-time HFR Camera-projector System

Our HFR camera-projector system comprises two projectors (DLP Light Commander 5500, Texas Instruments), an HFR vision platform IDP Express [16], a GPU board (Tesla 1060, NVIDIA), and a personal computer (PC). Fig. 1 shows its configuration. The DLP Light Commander 5500 is a development kit for HFR projection based on DMD device technology, which is composed of a high-performance light engine consisting of red (R, 623 nm), green (G, 525 nm), blue (B, 460 nm), and infrared (IR, 850 nm) LEDs, a DLP 0.55 XGA Chipset, and its controller. The two projectors, located on the horizontal shelves, project light patterns onto a common workspace. One projector, hereinafter referred to as the "IR projector," was used for fast structured light 3-D measurement to project 1024×768 IR binary light patterns at 1000 fps in synchrony with the HFR vision platform, while the other, hereinafter referred to as the "RGB projector," was used for projection mapping as a standard XGA projector connected to a PC video card, which is visible to the human eye. The IR projector was placed on the lower shelf, while the RGB projector was placed on the upper shelf. Two F-mount 50 mm-focal-length lenses (Ai AF NIKKOR 50 mm

J. Chen, T. Yamamoto, T. Aoyama, T. Takaki, and I. Ishii are with Hiroshima University, Hiroshima 739-8527, Japan. (Corresponding author (J. Chen) Tel: +81-82-424-7692; e-mail: j-chen@robotics.hiroshima-u.ac.jp).

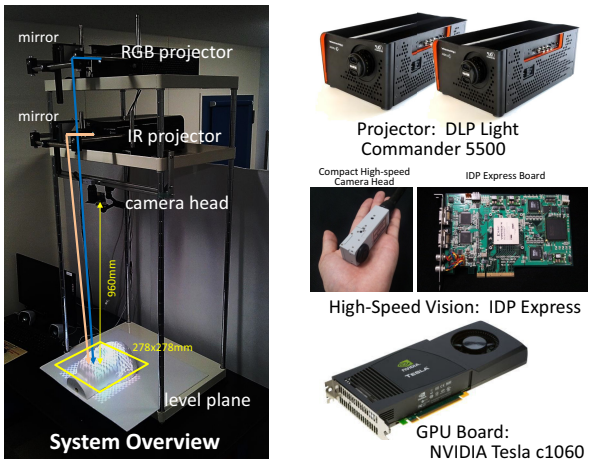


Fig. 1. System overview.

f/2.8D lenses, Nikon) were mounted on the projectors. Right-angle aluminum-coated mirrors were used to change the vertical direction of the projections from the projectors.

The IDP Express [16] was designed to implement real-time video processing and recording of 512×512 images at 2000 fps. It comprises a camera head to capture 8-bit gray 512×512 images at 2000 fps, and a dedicated FPGA board (IDP Express board). The IDP Express board has two camera inputs and trigger I/Os for external synchronization. Two 512×512 images and their processed results can be mapped onto the PC memory at 2000 fps via the PCI-e bus. A C-mount 17 mm-focal-length lens with an IR bandpass filter, whose center wavelength and full width-half max are 830 nm and 260 nm, respectively, was mounted on the camera head. RGB light patterns and other daily lightings were reduced in the images, and only IR light patterns captured at 1000 fps for the structured light 3-D measurement.

The Tesla 1060 is a computer processor board accelerated by an NVIDIA Tesla T10 GPU. It has a processing performance of 933 Gflops/s using 240 processor cores operating at 1.296 GHz and a bandwidth of 102 GB/s with its inner global memory of 4 GB and fast shared memory of 16 kB. A PC with 16-lane PCI-e 2.0 buses and a processor chipset with DMA were adopted to transfer memory-mapped data between standard memory and the Tesla 1060 via the PCI-e bus at high speed. We used a PC with the following specifications: ASUS P6T7 WS SuperComputer motherboard, Intel Core (TM) i7 3.20 GHz CPU, 3 GB RAM, two 16-lane PCI-e 2.0 buses, and graphic video card (QuadroFX 380, NVIDIA). Structured light 3-D measurement was accelerated using parallel-processing software on the Tesla 1060. We used a CUDA IDE provided by NVIDIA to code the algorithms with dedicated API functions for the IDP Express in Windows XP (32 bit), which enabled us to access memory mapped data. A DVI video output of the Quadro FX 380 was connected to the HDMI video input of the RGB projector, and the RGB light patterns generated for projection mapping were projected at dozens of frames per second.

For the xyz-coordinate system in the projection mapping workspace, the xy-plane was set on a level plane at $z = 0$ mm. The origin was set to the point of intersection of the optical axis of the camera lens and the xy plane. The optical axes of the camera, IR projector, and RGB projector lenses were parallel to the z-axis. The optical center of the camera lens was set at a height of $z = 899$ mm. Via the right angle mirrors, the optical center of the IR projector lens was virtually set at $z = 1204$ mm, and that of the RGB projector lens was virtually set at $z = 1504$ mm. In order not to disrupt the RGB projection with the right angle mirror for IR projection, the distance between the optical axes of the IR projector and the RGB projector lenses was set to 60 mm. The IR projector projects a 1024×768 IR light pattern in a 272×205 mm region on the level plane. The projected IR light pattern was captured in a 512×512 image using the IDP Express. The image area, a 278×278 mm square on the level plane, and depth information over a 272×205 mm region were observed. The RGB projector projects a 1024×768 RGB light pattern in a 343×257 mm region on the level plane, and covers the depth-measurable 272×205 mm region.

III. Implemented Algorithms

To project computer-generated patterns that respond to time-varying 3-D scenes, we implemented (a) structured light 3-D measurement and (b) depth-based projection mapping on our HFR camera-projector system. In the structured light 3-D measurement based on Inokuchi's method [17], 1024×768 IR binary light patterns coded with an 8-bit gray code are projected at 1000 fps; 512×512 images of the light patterns are captured at 1000 fps in synchrony with the IR projector. In the depth-based projection mapping, the depth images and features captured at a high frame rate in process (a) were used to generate RGB light patterns to project onto the 3-D scene, corresponding to the depth-based color mapping and AR spirit level applications used in our study.

A. Structured Light 3-D Measurement

1) Projection of positive/negative light patterns

The IR projector projects eight pairs of positive and negative light patterns with an 8-bit gray code in the order $\{g_0, g_1\}, \cdots$, and $\{g_{14}, g_{15}\}$ as follows:

$$\begin{aligned} g_{2i}(X', Y') &= \left\lfloor \frac{2^i \cdot Y'}{768} + \frac{1}{2} \right\rfloor \bmod 2 \\ g_{2i+1}(X', Y') &= \overline{g_{2i}(X', Y')} \end{aligned} \quad (i = 0, \cdots, 7), \quad (1)$$

where $\lfloor x \rfloor$ is an integer that is greater than or equal to x; i denotes the LSB and MSB order in the gray code. $(X,' Y')$ is a pixel coordinate value in the IR projection images.

2) Image acquisition of projected light patterns

Corresponding to the projection image of $g_j(X', Y')$ $(j = 0, \cdots, 15)$, a gray-level 512×512 image is captured at time $t = k\tau$ as follows:

$$I(X, Y, k) = \text{Proj}(g_{k \bmod 16}(X', Y')), \quad (2)$$

where k is the frame number of the captured images, and the frame interval is τ = 1 ms. (X, Y) is the coordinate value of a pixel in the captured images.

3) Binarization

$I(X,Y,2k')$ and $I(X,Y,2k'+1)$ are differentiated on the basis of robustness to ambiguities due to nonuniform brightness at 2 ms intervals. The binary image for space encoding is obtained with a threshold θ_b as follows:

$$G(X,Y,2k'+1)=\begin{cases} 1 & I(X,Y,2k')-I(X,Y,2k'+1)>\theta_b \\ 0 & I(X,Y,2k')-I(X,Y,2k'+1)<-\theta_b \\ \phi & \text{(otherwise)} \end{cases}, \quad (3)$$

where ϕ denotes the ambiguous binarization state.

4) Gray-to-binary conversion

According to the $(k' \bmod 8)$-th bit of the 8-bit gray code, $G(X,Y,2k'+1)$ is converted with a pure 8-bit binary code using $G(X,Y,2(k'-i)+1)$ at the current and previous frames $2(k'-i)+1$ $(i=0,\cdots,7)$ as follows:

$$B(X,Y,2k'+1)=\left(\sum_{i=k' \bmod 8}^{7} G(X,Y,2(k'-i)+1)\right) \bmod 2, \quad (4)$$

where we consider the unmeasurable state when there are two or more ambiguous binarizations between the eight frames; 0 is substituted for ϕ when there is one ambiguous binarization.

5) Space-code image generation

A space-code image $C(X,Y,2k'+1)$ is obtained using $B(X,Y,2(k'-i)+1)$ $(i=0,\cdots,7)$ at the current and previous frames as follows:

$$C(X,Y,2k'+1)=\sum_{i=0}^{7} 2^{(k'-i) \bmod 8} B(X,Y,2(k'-i)+1). \quad (5)$$

The space code image is filtered with a 3×3 median filter.

6) Triangulation

The space code image $C(X,Y,2k'+1)$ is transformed to a depth of $z=D(X,Y,2k'+1)$ by solving a simultaneous equation with a 3×4 camera transform matrix T_C and a 2×4 projector matrix T_P for the IR projector.

$$H_C\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}=T_C\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix},\ H_P\begin{pmatrix} C(X,Y,2k'+1) \\ 1 \end{pmatrix}=T_P\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \quad (6)$$

where H_C and H_P are parameters. The matrices T_C and T_P are obtained by prior calibration. The depth image $D(X,Y,2k'+1)$ is filtered with a 5×5 median filter.

7) Target region extraction

By differentiating the depth image with a reference depth image $D_R(X,Y)$ for background reduction, the target region is extracted as follows:

$$Q(X,Y,2k'+1)=\begin{cases} 1 & D(X,Y,2k'+1)-D_R(X,Y)>\theta_R \\ 0 & \text{(otherwise)} \end{cases}, \quad (7)$$

where θ_R is a threshold to extract a target region, and $D_R(X,Y)$ is the 3-D background scene provided prior.

8) 3-D position and orientation acquisition

Using the following zeroth-, first-, and second-order moment features M_{pqr} of the 3-D point group $S_{2k'+1}$ that satisfies $Q(X,Y,2k'+1)=1$,

$$M_{pqr}=\sum_{(x,y,z)\in S_{2k'+1}} x^p y^q z^r \quad (p+q+r\le 2), \quad (8)$$

the 3-D position and orientation of the target object are calculated as follows

$$(\bar{x},\bar{y},\bar{z})=\left(\frac{M_{100}}{M_{000}},\frac{M_{010}}{M_{000}},\frac{M_{001}}{M_{000}}\right), \quad (9)$$

$$\phi_x=\frac{1}{2}\tan^{-1}\frac{2(M_{011}M_{000}-M_{010}M_{001})}{(M_{002}-M_{020})M_{000}-M_{001}^2+M_{010}^2}, \quad (10)$$

$$\phi_y=\frac{1}{2}\tan^{-1}\frac{2(M_{101}M_{000}-M_{001}M_{100})}{(M_{200}-M_{002})M_{000}-M_{100}^2+M_{001}^2}, \quad (11)$$

$$\phi_z=\frac{1}{2}\tan^{-1}\frac{2(M_{110}M_{000}-M_{100}M_{010})}{(M_{020}-M_{200})M_{000}-M_{010}^2+M_{100}^2}, \quad (12)$$

where $(\bar{x},\bar{y},\bar{z})$ is the averaged 3-D position of the target object, and ϕ_x, ϕ_y, and ϕ_z are rotation angles around the x-, y-, and z-axes, respectively.

B. Depth-based Projection Mapping

According to the enhanced application tasks, the following depth-based projection mapping algorithms are implemented on our camera-projector system.

1) Depth-based color mapping

a) Assignment of color properties

Corresponding to the 512×512 depth image $D(X,Y,k)$, a color property $P(x,y,z)$ is assigned for 3-D points (x,y,z) using the following color map function with respect to depth:

$$P(x,y,z)=\begin{cases} Cm(z) & (x,y,z)\in S_k \\ \emptyset & \text{(otherwise)} \end{cases}. \quad (13)$$

In our study, several color map functions $Cm(z)$ such as a cyclic jet color map and a reduced-color-depth map were implemented for sensitive and distinct depth visualization.

b) Projection of RGB light patterns

The color properties $P(x,y,z)$ were converted into a 1024×768 RGB light pattern $P(X'',Y'',k)$. (X'',Y'') is a pixel coordinate value in the RGB projection images. This conversion is conducted with a 3×4 projection matrix $T_{P'}$ for the RGB projector, which indicates the relationship between the xyz- and the $X''Y''$-coordinate systems as follows:

$$H_{P'}\begin{pmatrix} X'' \\ Y'' \\ 1 \end{pmatrix}=T_{P'}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \quad (14)$$

where $H_{P'}$ is a parameter, and the projection matrix $T_{P'}$ is obtained by prior calibration. Thus, 1024×768 RGB light patterns are projected for pixel-wise projection mapping from the RGB projector onto the measured 3-D scene.

2) AR spirit level

a) Generation of CG patterns

The computer graphic (CG) pattern $G(x',y')$ is designed for the AR spirit level, in which two pointers are movable in a guide circle of radius R. The two pointers are located at $(a\phi_x,0)$ and $(0,a\phi_y)$ on the vertical and horizontal axes of

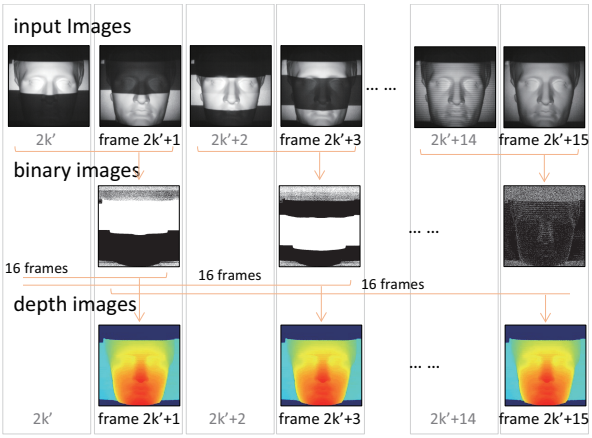


Fig. 2. Pipelining-output of depth images.

the circle in $G(x', y')$ so that their distances from the center of the circle increase sensitively with a large proportionality constant a, even when the rotation angles around the x- and y-axes are slightly small.

The CG pattern $G(x', y')$ is projected as color properties $P(x, y, z)$ of 3-D points (x, y, z) on an approximated tangent plane of the target object. The approximated tangent plane involves the averaged 3-D position $(\bar{x}, \bar{y}, \bar{z})$, and its normal vector corresponds to the rotation matrix $R(\phi_x, \phi_y, \phi_z)$, expressed by the rotation angles ϕ_x, ϕ_y, and ϕ_z around the x, y, and z-axes, respectively. The coordinate value (x, y, z) on the tangent plane is converted from (x', y') as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R(\phi_x, \phi_y, \phi_z) \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix} + \begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix}. \quad (15)$$

b) Projection of RGB light patterns

The same process as that in 1-b), is conducted.

C. Specifications

In the structured light 3-D measurement, steps (2)–(8) are accelerated by executing them in parallel with 512 blocks of 1×512 pixels on the GPU board. The 512×512 depth image is outputted at 500 fps using pipelined parallel-processing for input images between 16 frames, as shown in Fig. 2. Table I shows the execution time for the structured light 3-D measurement, including the transfer time from the PC memory to the GPU board for a 512×512 input image. The total execution time is within 0.39 ms.

For depth-based projection mapping, 1024×768 RGB light patterns to be projected were also generated by executing processes in parallel on the GPU board. In the case of depth-based color mapping, color properties were assigned within 0.01 ms and RGB light patterns were generated within 0.01 ms. In the case of AR spirit level, CG patterns were generated within 0.03 ms and RGB light patterns were generated within 0.01 ms. In both tasks, the transfer time from the GPU board to the PC memory for RGB light patterns was 1.46 ms. Including the depth image acquisition in the structured light 3-D measurement, the total execution times for (1) depth-based color mapping and (2) AR spirit level, were 1.88 ms and 1.89 ms, respectively. Thus, our HFR camera-projector system can generate projection patterns for projection mapping with low latency at the millisecond level, interactively with time-varying 3-D scenes. Here, the projection rate of the RGB projector was limited at 60 fps when 1024×768 RGB light patterns were transferred from the PC via the HDMI video output. Thus in this study, the RGB light patterns were projected with a time delay of 2.0 ms at intervals of 16.7 ms, whereas the 512×512 depth images were obtained in real time at 500 fps.

TABLE I

EXECUTION TIME OF STRUCTURED LIGHT 3-D MEASUREMENT.

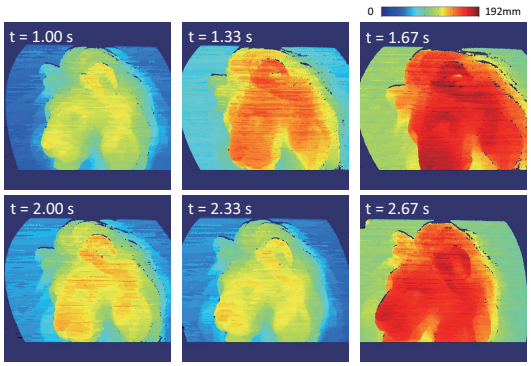
	Time [ms]
(2) Image acquisition	0.03
Transfer to GPU	0.08
(3) Binarization	0.04
(4) Gray-to-binary conversion	0.08
(5) Space-code image generation	0.02
(6) Triangulation	0.01
(7) Target region extraction	0.01
(8) 3-D position and orientation calculation	0.12
Total time	0.39

IV. EXPERIMENTS

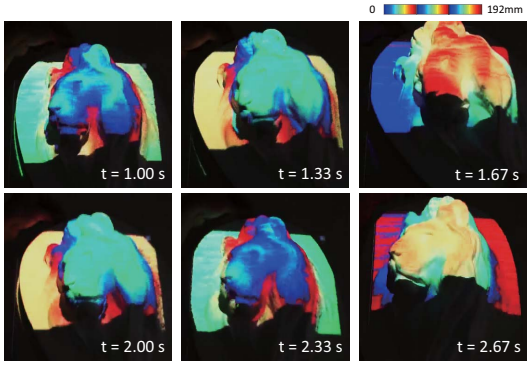
A. Depth-based Color Mapping

In this section, we look at experimental results obtained for projection-mapping of (a) a moving plaster lion relief, and (b) a moving human hand over a desktop when the depth-based color mapping was implemented. In the structured light 3-D measurement, the binarization threshold was θ_b = 1 when the LSB light pattern was binarized, otherwise θ_b=2. The threshold to extract a target region was θ_R = 1 mm. The exposure time of the camera head was 1 ms.

The plaster lion relief to be enhanced was moved above the level plane with periodic up-and-down motions once or less per second and slight rotations around the z axis by human hands. The height, width, and depth of the lion relief were 31 cm, 27 cm, and 10 cm, respectively. The reference depth image was provided as the level plane, and the color map function was set to a cyclic jet color map. Fig. 3 shows (a) the color-mapped 512×512 depth images, and (b) the experimental scenes captured using a standard digital video camera, taken at intervals of 0.33 s for t = 1.00–2.67 s. t = 0 was the start for the observation time. The x, y, and z coordinates for t = 0.0–5.0 s were measured as shown in Fig. 4. It can be seen that the 3-D position of the relief was periodically changed in the z direction whereas the x and y coordinates were not changed as much, corresponding to its up-and-down movement. Corresponding to the measured position and orientation, the 3-D shape information of the relief was correctly measured in the depth images in Fig. 3(a), even when the relief was moved up-and-down with slight rotation. In Fig. 3(b), the white-surface relief was enhanced with a cyclic jet color map, which can directly visualize its detailed height information for the human eye, and the RGB light patterns were correctly projected for pixel-wise projection mapping on the moving lion relief. In the experiment, there

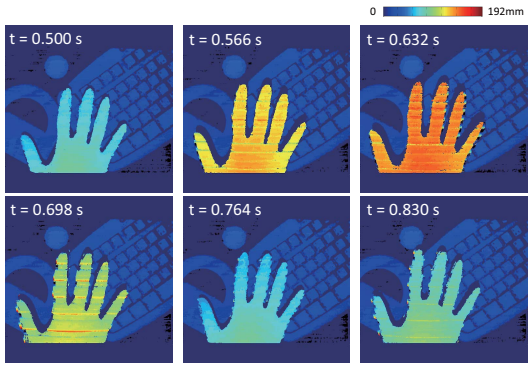


(a) depth images

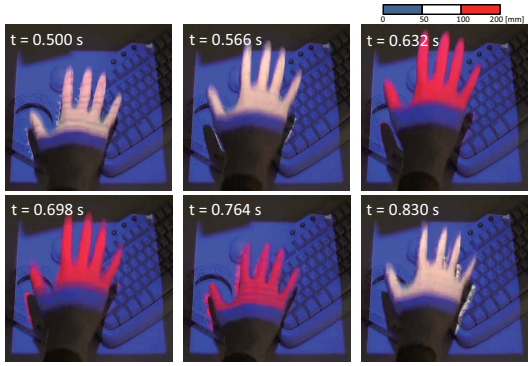


(b) color-mapped scenes

Fig. 3. Experimental results for a moving lion relief.



(a) depth images



(b) color-mapped scenes

Fig. 5. Experimental results for a moving human hand.

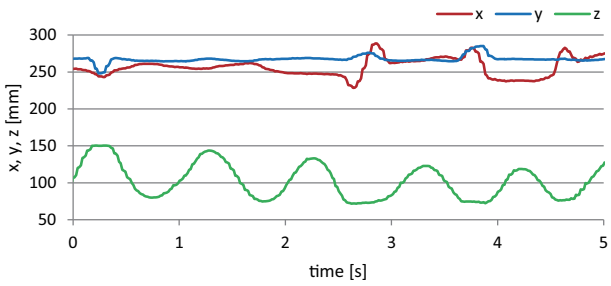


Fig. 4. 3-D position of a moving lion relief.

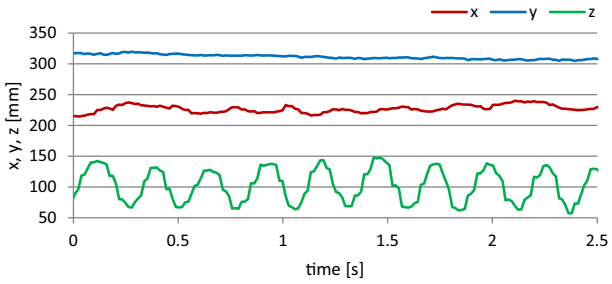


Fig. 6. 3-D position of a moving human hand.

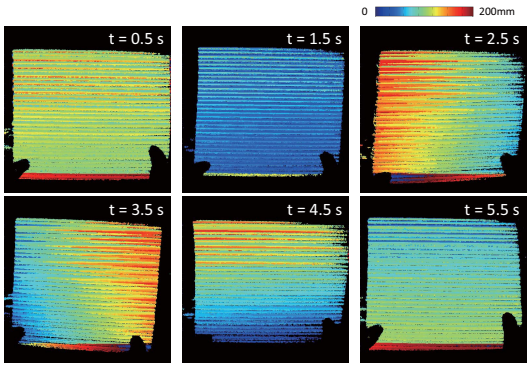
remained slight displacements between the lion relief and the projected color map information when the relief was moved rapidly. These displacements were caused by the latency of the projector and video card used in the projection, which had a delay time of dozens of milliseconds.

Let us look at the experimental results for a human hand periodically moved over the desktop. The right hand was moved periodically at a frequency of approximately 3 Hz in the z direction from 55 mm to 145 mm above the level plane. On the level plane, a computer keyboard, books, and many 3-D objects were placed as background objects. The reference depth image was provided previously as the same background scene in the real-time experiment, and the color map function was set to a three-color-depth color map; the target object was red-mapped when $z \geq z_{top} =$ 100 mm, otherwise the target object was white-mapped; the background scene was always blue-mapped. Fig. 5 shows (a) the depth images, and (b) the experimental scenes, which were taken at intervals of 0.066 s for $t = 0.50$–0.83 s. Fig. 6 shows the x, y, and z coordinates of the human hand for $t = 0.0$–2.5 s. The 3-D position of the human hand was periodically changed three times per second in the z direction, corresponding to the periodic movement of the human hand. The 3-D shape information of the human hand and background objects were measured in Fig. 3(a). In Fig. 3(b), it can be seen that the human hand was highlighted with white or red colors depending on the 3-D position of the human hand, whereas the background scene was always lighted in blue.

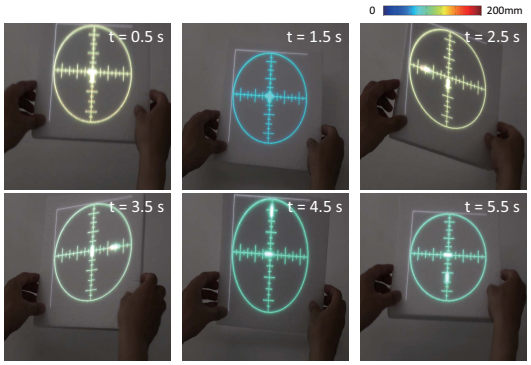
These results indicate that our system can execute real-time pixel-wise projection mapping on a moving 3-D object for depth-enhanced visualization.

B. AR Spirit Level

Let us look at the experimental results for the AR spirit level. The parameters in the structured light 3-D measurement were the same as those in the depth-based color



(a) depth images



(b) projection-mapped AR spirit level

Fig. 7. Experimental results for AR spirit level.

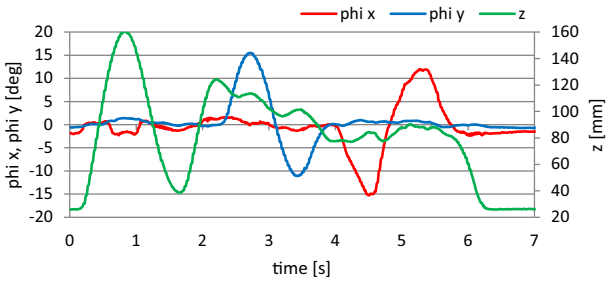


Fig. 8. 3-D orientation and z-coordinate of a white plate.

mapping experiments. A 50 mm$\times$50 mm white plate was manually moved over the level plane; it was moved up and down, and then alternately tilted around the y axis and the x axis. The graphic pattern to be projected involved a guide circle of radius 100 mm and two 20-mm-diameter pointers; the pointers moved 6 mm per degree for the rotation angles around the x and y axes, and the color of the graphic pattern was determined by the z coordinate of the target object. The reference depth image was provided as the level plane.

Fig. 7 shows (a) the depth images, and (b) the experimental scenes, taken at intervals of 1.0 s for t = 0.5–5.5 s. The rotation angles around the x- and y-axes and the z coordinate of the target object for t = 0.0–6.0 s were measured as shown in Fig. 8. The graphic pattern was projected at the center of the white paper and its pointer positions and color were correctly changed, corresponding to the rotation angles and the z coordinate of the white plate; the z coordinate was changed from 20 mm to 165 mm, and then the angles were alternately changed in the range from approximately -15 to 15 degree. Thus, it can be seen that the slight rotation angle at sub-degree level is enhanced for easy visualization on the white plane at AR spirit level, even when its rotation angle is too small for the human eye to inspect its slanted tendency without projection mapping.

V. Conclusion

In this paper, we reported on the developement of a real-time projection mapping system that can acquire 512$\times$512 depth images in real time at 500 fps and project depth-based light patterns to be enhanced for time-varying 3-D scenes. Experimental results from enhanced application tasks for dynamic 3-D scenes verified the efficacy of our system. On the basis of the experimental results obtained, we plan to improve our camera-projector system for more responsive 3-D projection mapping by accelerating CG image generation for a short time-lag projector, and to extend enhanced applications for various AR-based human-computer interactions.

References

[1] R.T. Azuma, "A Survey of Augmented Reality," *Teleoperators and Virtual Environments* 6, 4, pp. 355–385, 1997.
[2] H. Kato and M. Billinghurst, "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System," *Proc. IEEE/ACM Int. Workshop on Augmented Reality*, 85–94, 1999.
[3] X. Zhang, S. Fronz, and N. Navab, "Visual Marker Detection and Decoding in AR Systems: A Comparative Study," *Proc. Int. Symp. Mixed and Augmented Reality*, 97–106, 2002.
[4] J. Chandaria, G.A. Thomas, and D. Stricker, "The MATRIS Project: Real-time Markerless Camera Tracking for Augmented Reality and Broadcast Applications," *J. Real-Time Image Process.*, 2(2–3), 69–79, 2007.
[5] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Real-time Detection and Tracking for Augmented Reality on Mobile Phones," *IEEE Trans. Vis. Comput. Graph.*, 16(3), 355–368, 2010.
[6] O. Bimber and R. Raskar, *Spatial Augmented Reality, Merging Real and Virtual Worlds*, AK Peters, 2005.
[7] M. Mine, D. Rose, B. Yang, J. van Baar, and A. Grundhofer, "Projection-Based Augmented Reality in Disney Theme Parks," *IEEE Comput.*, 45(7), 32–40, 2012.
[8] O. Bimber, B. Frohlich, D. Schmalstieg, and L.M. Encarnacao, "The Virtual Showcase," *IEEE Comput. Graph. Appli.*, 21(6), 48–55, 2001.
[9] M. Flagg and J.M. Rehg,"Projector-Guided Painting," *Proc. Annu. ACM Symp. User Interface Software and Technology*, 235–244, 2006.
[10] S. Chon, H. Lee, and J. Yoon, "3D Architectural Projection, Light Wall," *Leonardo*, 44(2), 172–173, 2011.
[11] Microsoft, Xbox 360 Kinect, http://www.xbox.com/en-US/kinect
[12] A. Wilson, H. Benko, S. Izadi, and O. Hilliges, "Steerable Augmented Reality with the Beamatron," *Proc. Annu. ACM Symp. User Interface Software and Technology*, 413–422, 2012.
[13] B. Jones, H. Benko, E. Ofek, and A. Wilson, "IllumiRoom: Peripheral Projected Illusions for Interactive Experiences," *Proc. ACM SIGGRAPH 2013 Emergent Technologies*, 7, 2013.
[14] K. Okumura, H. Oku, and M. Ishikawa, "Lumipen, Projection-based Mixed Reality for Dynamic Objects," *Proc. IEEE Int. Conf. Multimedia and Expo*, 699–704, 2012.
[15] H. Gao, T. Takaki, and I. Ishii, "GPU-based Real-time Structure Light 3D Scanner at 500 fps," *Proc. SPIE 8437* , 84370J–84370J-9, 2012.
[16] I. Ishii, T. Tatebe, Q. Gu, Y. Moriue, T. Takaki, and K. Tajima, "2000 fps Real-time Vision System with High-frame-rate Video Recording," *Proc. IEEE Int. Conf. Robot. Automat.*, 1536–1541, 2010.
[17] S. Inokuchi, K. Sato, and F. Matsuda, "Range Imaging System for 3-D Object Recognition," *Proc. Int. Conf. Patt. Recog.*, 806–808, 1984.