

3,000-fps 3-D Shape Measurement Using a High-Speed Camera-Projector System

Joji Takei, Shingo Kagami and Koichi Hashimoto

Graduate School of Information Sciences,
Tohoku University

6-6-01 Aramaki Aza Aoba, Aoba-ku, Sendai 980-8579, Japan

swk(at)ic.is.tohoku.ac.jp

Abstract—This paper proposes a high frame rate 3-D shape measurement system based on structured light projection using temporally-coded binary patterns. This system consists of a DMD (Digital Micromirror Device) projector and a high-speed camera, both of which operate at over 6,000 fps. Introducing an image processing algorithm of detecting and tracking the projected light patterns enables frame-by-frame 3-D image measurement. Experimental results show that the proposed algorithms worked effectively, and the system achieved 3-D imaging at 3,350 fps.

I. INTRODUCTION

Three-dimensional measurement based on visual information is in wide use in many fields such as robot vision, industrial and scientific measurement, medical applications, intelligent transporting systems, and human interfaces. For industrial applications in particular, active stereo methods, in which structured light patterns are projected onto the target objects and the projected patterns are captured by cameras, are preferred because accurate measurement can be achieved more easily than by passive stereo methods.

One of the most desired breakthroughs for the active stereo methods is to achieve high-speed 3-D measurement. The goal of this paper is to achieve several thousands of frames (distance maps) per second 3-D measurement so that 3-D information of fast and dynamic phenomena can be obtained.

Many existing methods [1], [2], including the slit line projection method and the Gray code pattern projection method, require projection and capturing of multiple light patterns to obtain a single 3-D distance map, which takes multiple frame times of the involved camera and projector. The multiple frames are used to determine the stereo correspondence between the projector coordinates and the camera coordinates and reconstruct the depth information. The target objects must be still during these multiple frames.

Several recent studies challenged this difficulty by mainly two approaches: achieving high frame rate projection and image capturing, and achieving single frame measurement. The former approach is pursued, in particular, for slit line projection method implementations [3], [4]. Extremely high frame rate (e.g. 50k – 260k fps (frames/s)) custom-made image sensors specialized for the slit line projection method are utilized to generate video frame rate (e.g. 30 fps) 3-D maps. To raise the frame rate further, the exposure time of the sensor must be extremely short. This is technically feasible

[4], but requires a very strong laser light source, which raises safety issues.

The latter approach, single frame measurement, includes simple multi-slit or multi-spot light pattern projection [5], [6]. Since these patterns are in principle not sufficient to determine the projector coordinates, some additional assumptions or constraints are required. Besides, single frame measurement approach includes the techniques to project patterns that convey the projector coordinates information as spatial patterns, grayscale patterns, or color patterns. Spatial patterns, however, are not always decoded easily.

In this paper, we employ both of the two approaches. We introduce a high-speed camera and a high-speed projector that operate at several thousands of frames per second. The projected patterns are binary black/white images, because grayscale or color images are difficult to project at high frame rate, and are also difficult to reproduce the grayscale or color on the imager side at high frame rate because of low S/N ratio due to short exposure time. Then we propose a new technique of interframe tracking of the projected patterns in the camera image space so that we can obtain a 3-D distance map every couple of frames. Thus we obtain, for example, 500-fps 3-D distance maps when we project and capture the patterns at 1,000 fps.

The high-speed camera we use is a commercially available one. On the other hand, because no such high-speed projec-

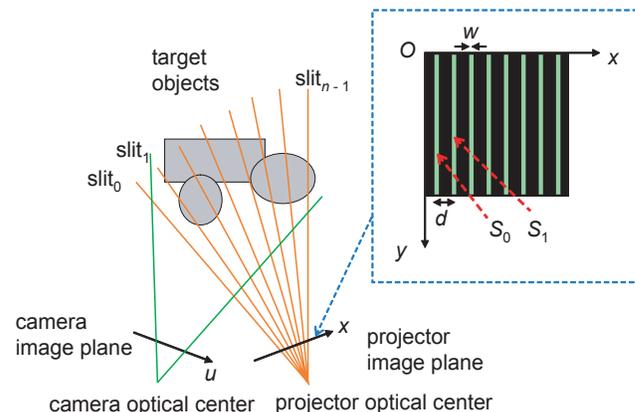


Fig. 1. The camera-projector configuration and the projector image pattern description.

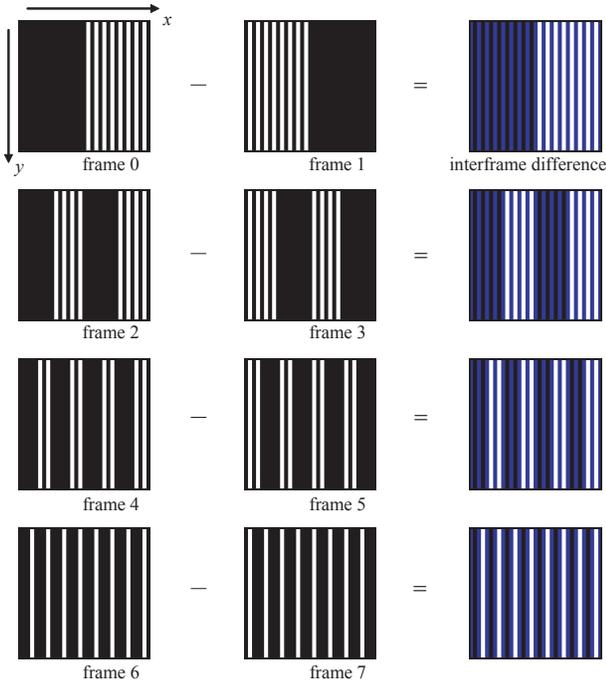


Fig. 2. Temporally-coded binary patterns for projection. The blue areas in the right column represent the regions other than the slits (always dark).

tors are commercially available, we customize a commercial projector so that high frame rate projection is achieved. We use a DMD (Digital Micromirror Device) [7] as a pattern generator (spatial light modulator) for the custom projector.

The DMD is a MEMS device in which several hundreds of thousands of micromirrors are implemented, and each of the micromirrors can be switched to one of two directions rapidly so that it can be used as a spatial light modulator (SLM) in which each pixel can be independently set on and off at 10,000 – 40,000 fps.

The rest of the paper is organized as follows. In Section II, the proposed pattern tracking method is described. In Section III and IV, implementation of the camera-projector system and experimental results are presented. Section V gives the conclusion.

II. TEMPORALLY-CODED MULTI-SLIT PROJECTION METHOD

A. Projected Patterns

The proposed method is based on the multi-slit light projection method (e.g. [5]). Multiple parallel slit lines are projected simultaneously, and the imager captures the line images. If each captured slit line can be correctly identified with one of the projected lines, 3-D positions of the points in the line are uniquely determined by triangulation assuming that the camera and the projector are fully calibrated. However, determining the correspondence of multiple parallel slits is ambiguous. Thus application of the conventional multi-slit projection method is often limited to, for example, continuous surfaces, limited depth ranges, or obtaining partial information on surface properties such as orientations.

To overcome this limit, we apply a temporal coding technique to the multi-slit projection method. The system configuration is shown in Fig. 1. The projected pattern sequence is as follows. Let (x, y) denote the projector image coordinates, and d and w denote the slit spacing and the slit width in the projector image space, respectively. The i -th slit region S_i is defined as $\{(x, y) \mid d(i + 1/2) - w/2 \leq x \leq d(i + 1/2) + w/2\}$. We refer to the suffix i of the slit region S_i as the slit ID.

Our aim is to encode the slit ID in the blinking sequence of the slit region. We use a couple of successive frames to express a bit of the slit IDs in a Manchester-like coding. Assuming that we have 2^n slits in the projector image, let $b_{n-1}^i b_{n-2}^i \cdots b_1^i b_0^i$ be a binary expression of the slit ID i .

The sequence of $2n$ patterns are repeatedly projected frame by frame. In the frames $2k$ and $2k + 1$ modulo $2n$, a pixel in the slit region S_i is bright and then dark (The DMD mirror on and off) if $b_{n-k-1}^i = 1$, respectively in this order, and dark and then bright (off and then on) if $b_{n-k-1}^i = 0$. The pixels in the region other than the slits are always dark (The DMD mirrors are off). The projected image sequence for the case $n = 4$ is shown in Fig. 2.

In this Fig. 2, the natural binary code is used. Instead, in our implementation described in Sections III and IV, the cyclic binary code (or more widely known as the Gray code) is used. Note that the proposed algorithm does not depend on particular formats of binary coding.

In the image processing, the slit image regions are simply detected by calculating the interframe pixelwise difference of successive odd and even frames, because we assume that the projection and capturing frame rates are high enough so that the movements of the target objects are very small in successive frames. This Manchester-like coding also contributes to cancel the influences of background illumination or reflectance of the object surfaces, which is widely utilized in the structured light methods.

Accumulating the frame differences of the pixel brightness, positive or negative, during the $2n$ frames yields the slit ID of the corresponding slit and thus the corresponding projector coordinates. It is, however, difficult to assume that the movements of the target objects are small enough for the duration of the $2n$ frames. Hence the tracking technique should be introduced in the image processing.

In the rest of the paper, a sequence of the $2n$ frames is referred to as a lap, and a successive odd and even frames as a step.

B. Decoding and Tracking of the Images

The image processing flow at the camera side is as follows. As described in the previous subsection, frame difference of successive odd and even frames is computed and slit regions are extracted. The changes of brightness, positive or negative, which correspond to the bits in slit ID codes, are recorded and accumulated to reconstruct the slit IDs. If the target objects move, however, the accumulated slit IDs are mixed sequences of the bits from several IDs, and correct

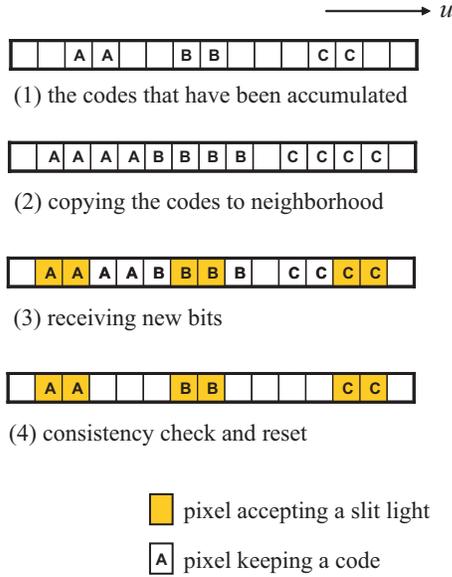


Fig. 3. Flow of the code tracking process.

correspondence of the camera and projector coordinates is not achieved.

Thus a tracking technique based on what we reported previously [8] is introduced. The tracking procedure takes place for each of horizontal lines of the camera image as shown in Fig. 3. The alphabetical letters in the pixels express the slit ID codes that the pixels have accumulated so far, and the pixels painted yellow show the pixels that accept the slit light at the current step.

In Fig. 3 (1), the slit ID codes A, B and C have been kept in the corresponding camera pixels. As we assume high frame rate projection and capturing, the image of each slit will move only in the neighborhood of the current position. Thus we copy the accumulated codes into a certain number of neighbor pixels until it collides with the code from its neighbor slit (in that case, the code copying stops at the midpoint of the two slits in the camera image space), as shown in Fig. 3 (2). If the newly received bit is consistent with the code that has been kept, as shown in Fig. 3 (3) and (4), the code accumulation goes on. If no new bit is received (i.e. no significant brightness change is detected at the pixel) or the new bit is not consistent, the accumulation at the pixel is reset.

Executing this procedure repeatedly frame by frame, the slit IDs can be reconstructed. It should be noted that when a new slit light image appears in the camera image sequence, it will take one lap, that is, $2n$ frames to yield its slit ID. After this first lap, the position of the slit (with known slit ID) in the camera image is updated every two frames, and the updated position is indeed the one at the present time, not the averaged one over the $2n$ frames. Therefore 3-D distance maps are obtained every two frames.

During the first lap for a slit, in which the slit ID code has not been determined yet, the consistency check is not applied.

After the first lap, the consistency check, that is, comparing the new received bit with the one that was received at the corresponding step in the previous lap begins to take place. If they are different, the accumulated code at the pixel is reset. This contributes to suppression of artifacts due to tracking errors. A code for a slit is called a confirmed one if all the bits in the code of the slit ID have been checked to be consistent. In other words, a code for a slit light that has been tracked for two laps is a confirmed one.

C. Tracking Across Jump Edges

A jump edge is an edge of the target object set where the distance from the camera changes discontinuously. Around jump edges, the previously described tracking algorithm does not work correctly, because a slit light image will move discontinuously in the camera image space.

When the tracking of a slit light image fails, the accumulated code corresponding to the slit will be reset within a lap at the latest because the consistency check will detect the failure, and thus the artifacts due to the tracking error are suppressed. On the other hand, this generates void regions, in which no 3D-reconstructed points are available, around jump edges. This may be an undesired behavior in some cases, for example, when structures around edges are of main interest.

Hence we introduce a slight modification for the tracking procedure at around jump edges. In normal cases, a new detected bit must be exactly consistent with the code that has been accumulated. When a jump edge is detected, we make some allowance in the consistency check around the edge.

When investigating a horizontal line of the code image, if we find that two codes adjacent in the line have significantly large difference or that two adjacent codes are accumulated at significantly far away from each other in the line, we presume that a jump edge exists between the two. In this case, we allow one of the codes to change into the other, and moreover, into the codes that are within certain neighborhoods of the two codes (i.e. those of the slits neighboring in the projector image space), because a code that has been occluded behind the jump edge might appear. Similarly, it is allowed to change into the adjacent codes in the upper and lower lines and their neighborhood codes. This technique is applied only to jump edges because it increases the chance of tracking errors.

III. SYSTEM IMPLEMENTATION

The developed experimental system consists of a DMD and its control electronics, a projector optics and a light source, and a high-speed camera. The system is shown in Figs. ?? and 4.

A DMD (DMD 0.7XGA 12 DDR) developed by Texas Instruments and its controller board DMD Discovery 1100 are used as an SLM for the high-speed pattern projection. The DMD has $1,024 \times 768$ mirrors and they can be switched on and off at 10,000 fps at maximum. The projected binary patterns are stored in an external memory board ViALUX ALP-1 connected to the control board Discovery 1100.

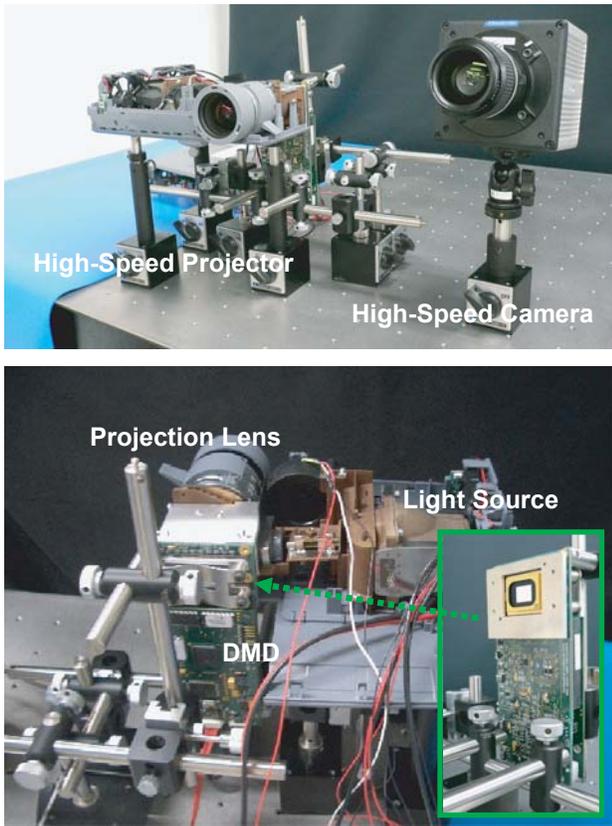


Fig. 4. The developed high-speed camera-projector system.

We introduced a projector optics by removing the original SLM and its associate electronics from a projector InFocus LP600 and attaching the new high-speed DMD module. Because we used the optics designed for a video projector, the current implementation has some limitation as a 3-D measurement system: measured objects must be at the distance of approximately 1 meter or farther in front of the projector, and the depth of focus is limited to a few hundreds of millimeters.

The high-speed camera used is a Photron FASTCAM-1024PCI that captures $1,024 \times 1,024$ pixels monochrome images at 1,000 fps. By reducing the image area size, the frame rate can be accordingly increased. A synchronization signal is sent from the ALP-1 board to the camera via a 3-meters $50\text{-}\Omega$ coaxial cable. The extrinsic and intrinsic parameters of the camera and the projector is calibrated using Zhang's method [9] combined with Gray code pattern projection.

The captured images are stored in a PC and then analyzed off line. Since the proposed algorithm mainly is composed of combination of local image processing, real-time processing will be possible by implementing the algorithm in hardware.

IV. EXPERIMENTS

In this section, the experimental results are shown ¹. The projector was located at about 0.2 m to the right of the

¹A movie is available as an extended object. Movies in other several formats with higher quality are also available at the authors' web site: http://www.ic.is.tohoku.ac.jp/~swk/iros2007_movie/

camera. A white screen ($0.9 \times 0.6 \text{ m}^2$) was set at the distance of about 1.5 m in front of the camera, and a white rectangular solid was randomly moved by hand just before the white screen at the speed of approximately 3 m/s. These objects were selected so as to confirm that planar surfaces or shapes with jump edges are correctly reconstructed.

The parameters used for the projected patterns are: $n = 8$, $w = 1$ [pixel], and $d = 7$ [pixels]. The code copying for tracking is carried out into at most neighboring 4 pixels on a side.

Figures 5 and 6 shows captured images and reconstructed 3-D shapes by 1,000-fps projection and capturing with $1,024 \times 1,024$ camera pixels. Hence the resulting 3-D range maps were obtained at 500 fps. Figure 5 shows the results of the proposed temporally-coded multi-slit projection method, while Fig. 6 shows the results of the conventional Gray code pattern projection method for static scenes [1].

The images in the third columns from the left of the figures show the detected code bit images. The pixels that received valid code bits 1 and 0 are painted white and black, respectively, and the pixels that received no code valid bits (i.e. no significant brightness changes were detected) are painted blue.

In the 3-D reconstruction results, blue points are the candidate points that had not been confirmed yet by the consistency check. Note that the blue points are plotted as bigger points than ordinary red points.

As shown in Fig. 6, the conventional Gray code pattern projection method yielded incorrectly reconstructed results in which the planar surfaces were distorted, and large artifacts appeared at around jump edges. The proposed method, as shown in Fig. 5, on the other hand, succeeded in reconstructing the planar surfaces correctly, and in reconstructing the jump edges as well. It can also be seen that many blue "non-confirmed" points appeared around jump edges, which implies that improving the jump edge processing will further reduce measurement errors.

The working space size, obtained resolution and accuracy depend on the camera-projector configuration. In the present experiment approximately 150 slit lines are projected within 900 mm width at the distance of 1.5 m in front of the camera-projector, Thus the resolution in x direction (horizontal) is approximately 1/6 points/mm, while the accuracy of a point position is expected to be around the width of a projected slit, which is approximately 1 mm. The resolution and accuracy in y direction is determined by the camera resolution and are around 1 mm. The accuracy in z direction (depth) is estimated on the basis of triangulation. In the present experimental setup, the distance to the objects, 1.5 m, is much longer than the baseline length, 0.2 m, which theoretically causes at most 15-mm measurement errors in depth at the center of the field of view. However, from the experimental results, it seems that these worst-case errors seldom appeared and instead most of the artifacts appeared due to tracking errors.

The results with higher frame rates, with the image sizes reduced, are shown in Figs. 7 and 8. Figure 7 shows detected

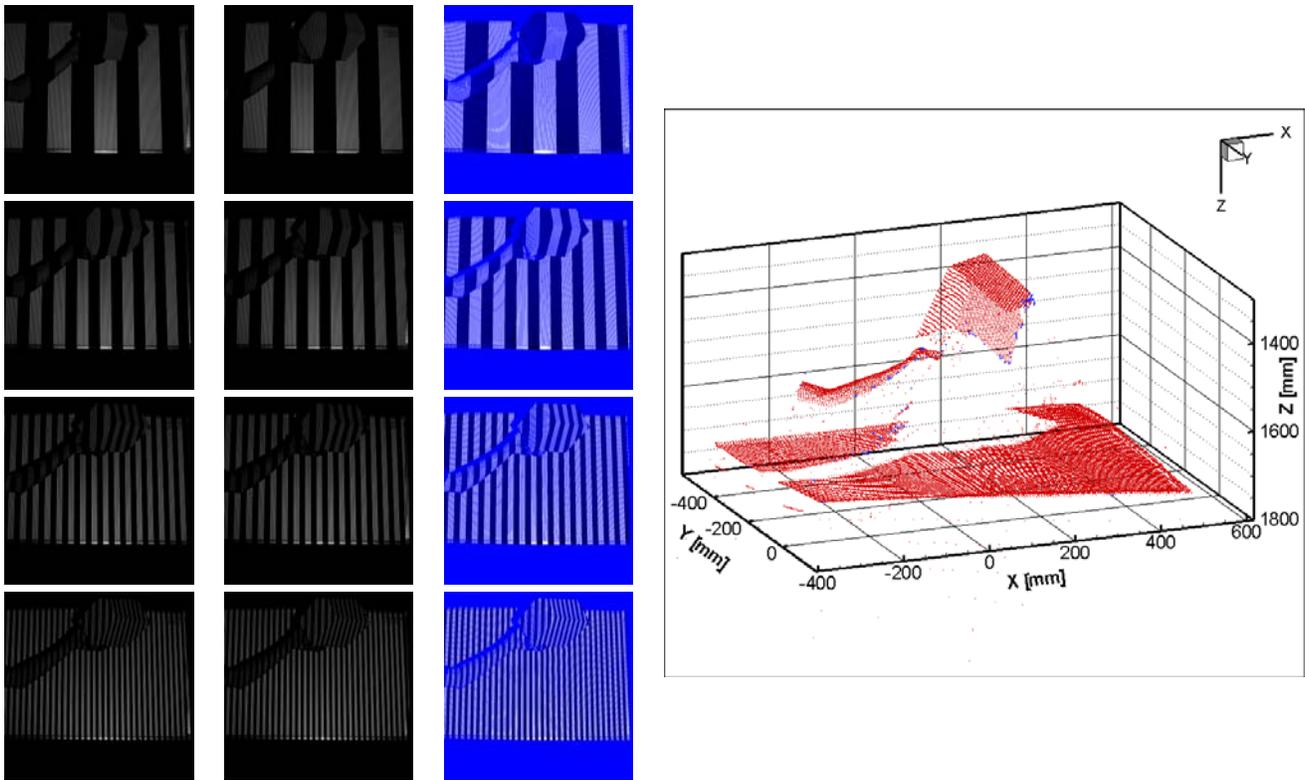


Fig. 5. Result of the proposed method. Obtained camera images (leftmost two columns), detected code bit images (3rd column), and a frame of 3D reconstruction result (rightmost) at 500 fps with $1,024 \times 1,024$ pixels camera images are shown.

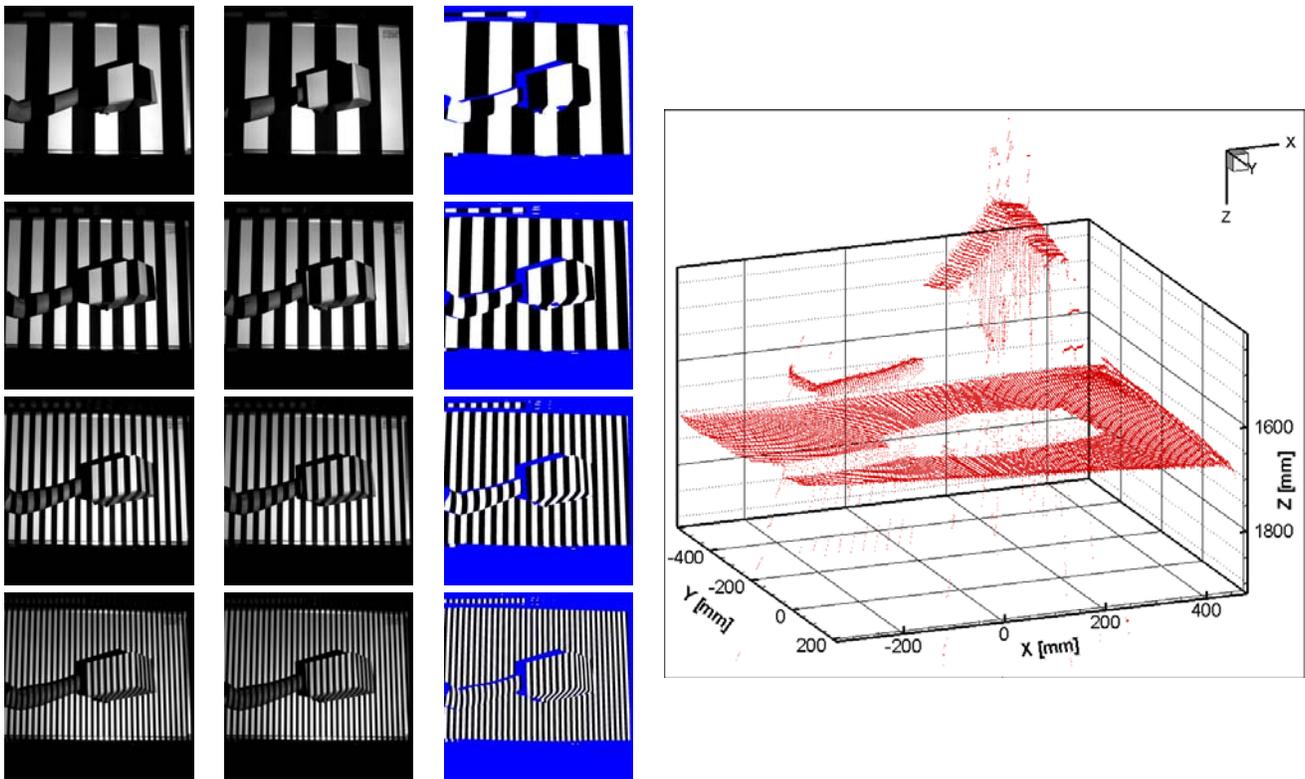


Fig. 6. Result of the conventional Gray code pattern projection method, at 500 fps with $1,024 \times 1,024$ pixels camera images.

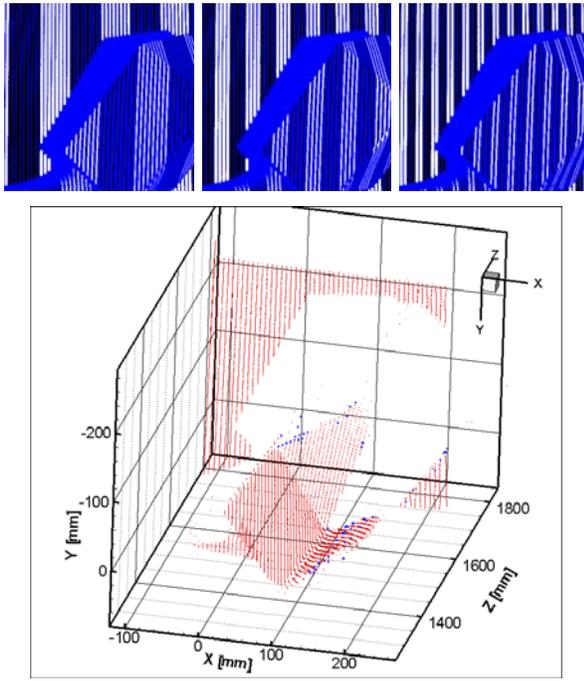


Fig. 7. Result of the proposed method, at 2,500 fps, 384×384 pixels camera images.

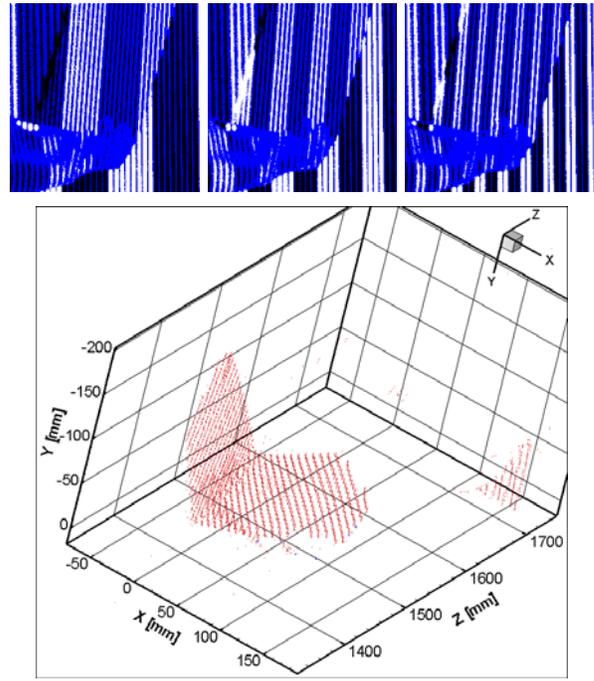


Fig. 8. Result of the proposed method, at 3,350 fps, 256×256 camera images.

code bit images by 5,000-fps projection and capturing with 384×384 camera pixels, and a frame of reconstructed 3-D shape thus at 2,500 fps. Likewise, code bit images obtained by 6,700-fps projection and capturing with 256×256 camera pixels and a frame of 3-D shape at 3,350-fps are shown in Fig. 8. This frame rate, 6,700 fps, is the maximum rate at which the projector and the camera could synchronize with each other using the TTL synchronization signal through the coaxial cable. The 3-D shape of the white rectangular solid was correctly reconstructed, but the human hand, of which the surface has lower reflectance, was hardly detected because of the shorter exposure times of the image sensor.

V. CONCLUSION

A high-speed 3-D shape measurement system has been described. A novel technique called the temporally-coded multi-slit projection method enabled 3-D distance maps to be obtained every two frames, assuming that the system operates at a high frame rate. Implementing a high-speed projector based on an optics system of commercial projector and a DMD development kit, high frame rate binary image projection at several thousand fps has been achieved, being synchronized with an off-the-shelf high-speed camera. As a result, 3-D reconstruction at 500 fps with $1,024 \times 1,024$ camera pixels, or at 3,350 fps with 256×256 camera pixels has been achieved. Future work includes hardware implementation of the image processing algorithm so that the system operates in real time, and to exploit the blank spaces between the slits in the projected patterns so that more dense 3-D distance maps can be obtained.

ACKNOWLEDGMENTS

Part of this study is a joint work with Photron Limited supported by the Grant of the New Energy and Industrial Technology Development Organization (NEDO), Japan.

REFERENCES

- [1] J. Batlle, E. Mouaddib, and J. Salvi. Recent progress in coded structured light as a technique to solve the correspondence problem: A survey. *Pattern Recognition*, 31(7):963–982, 1998.
- [2] Frank Chen, Gordon M. Brown, and Mumin Song. Overview of three-dimensional shape measurement using optical methods. *Optical Engineering*, 39(1):10–22, 2000.
- [3] Shinichi Yoshimura, Toshinobu Sugiyama, Kazuya Yonemoto, and Kazuhiko Ueda. A 48kframe/s CMOS image sensor for real-time 3-D sensing and motion detection. In *Digest of Technical Papers of 2001 IEEE International Solid-State Circuits Conference (ISSCC2001)*, pages 94–95, 2001.
- [4] Yusuke Oike, Makoto Ikeda, and Kunihiro Asada. A CMOS image sensor for high-speed active range finding using column-parallel time-domain ADC and position encoder. *IEEE Transactions on Electron Devices*, 50(1):152–158, 2003.
- [5] M. Asada, H. Ichikawa, and S. Tsuji. Determining surface orientation by projecting a stripe pattern. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):749–754, 1988.
- [6] Kazunori Umeda. A compact range image sensor suitable for robots. In *2004 IEEE International Conference on Robotics and Automation (ICRA2004)*, pages 3167–3172, 2004.
- [7] Dana Dudley, Walter Duncan, and John Slaughter. Emerging digital micromirror device (DMD) applications. *Proceedings of SPIE*, 4985, 2003.
- [8] Shingo Kagami, Masatsugu Shinmeimae, Takashi Komuro, Yoshihiro Watanabe, and Masatoshi Ishikawa. A pixel-parallel algorithm for detecting and tracking fast-moving modulated light signals. *Journal of Robotics and Mechatronics*, 17(4):384–397, 2005.
- [9] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.