A System for Capturing Textured 3D Shapes based on One-shot Grid Pattern with Multi-band Camera and Infrared Projector

Kazuhiro Sakashita, Yasushi Yagi
Osaka University, Osaka, Japan
{sakashita,yagi}@am.sanken.osaka-u.ac.jp

Ryusuke Sagawa
National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan
ryusuke.sagawa@aist.go.jp

Ryo Furukawaa
Hiroshima City University, Hiroshima, Japan
ryo-f@hiroshima-cu.ac.jp

Hiroshi Kawasaki
Kagoshima University, Kagoshima, Japan
kawasaki@ibe.kagoshima-u.ac.jp

Abstract

Recently, several method have been proposed to capture 3D shapes of a moving object at a high frame rate. One of promising approach to reconstruct a 3D shape is a projector-camera system that projects structured light pattern. One of the problem of this approach is that it has difficulty to obtain texture simultaneously because the texture is interfered by the illumination from the projector. The system proposed in this paper overcomes this issue by separating the light wavelength for texture and shape. The pattern is projected by using infrared light and the texture is captured by using visible light. If the cameras for infrared and visible lights are placed at different position, it causes the misalignment between texture and shape, which degrades the quality of the textured 3D model. Therefore, we developed a multi-band camera that acquires both visible and infrared lights from a single viewpoint. Moreover, to reconstruct a 3D shape using multiple wavelengths of light, namely multiple colors, an infrared pattern projector is developed to generate a multi-band grid pattern. Additionally, a simple method to calibrate the system is proposed by using a fixed grid pattern. Finally, we show the textured 3D shapes captured by the experimental system.

1. Introduction

Recently, several method have been proposed to capture 3D shapes of a moving object at a high frame rate. If the shape of a moving object is obtained, the resulting data can be applied in various fields, such as virtual reality (VR), computer vision, gesture recognition, and robotics research. In many applications, simultaneous acquisition of texture and shape is desirable. For example, if the texture and shape of a person in motion are obtained, the model of the person can be easily rendered in a virtual world using 3D CG

In this paper, we propose a system to capture texture and shape simultaneously based on projector-camera system. The main idea is that the proposed system acquires texture and shape by using different light wavelengths. The texture is captured by using visible light and the shape is reconstructed by using infrared (IR) light. By separating the wavelength, the system can capture the texture without affected by the projected structured light, which is a common issue for a projector-camera system.

To generate a textured 3D model from a reconstructed shape and a captured image, alignment of them are required. If the cameras for them are placed at different positions , the misalignment between shape and texture inevitably occur when shape has an error even if they are precisely calibrated. To avoid the problem, we developed a multi-band camera that can capture both visible and IR lights from a single viewpoint. Another advantage of single viewpoint is that image warping using depth information is not required and rendering error near occlusion does not occur.

The method of 3D reconstruction applied in this paper is based on [**?**]. The method uses a fixed grid pattern that consists of vertical and horizontal lines and the colors of lines are determined by a periodic code. In the paper, the method used two colors from RGB (*e.g.* blue and green). To implement the method by using IR light, we developed a projector that projects a grid pattern by using two IR wavelengths. Additionally, since a fixed pattern mask is built

1

in the proposed projector, it is not easy to change patterns like a video projector. Therefore, a method to calibrate the system only by using a fixed pattern is presented, which is useful to simplify the calibration step of the system.

In the following sections, the related work is described in Sec. 2. Next, the system to reconstruct textured 3D shapes by grid pattern is briefly explained in Sec. 3. Then, we describe a system to capture texture of visible light and IR structured light simultaneously in Sec. 4. Finally, an experimental system is tested to obtain the textured 3D model in Sec. 5.

2. Related work

Techniques to acquire the shape of an object are categorized roughly into to two types, active and passive methods. Stereo vision [?] and visual hulls [?] are often used in passive methods. Since these methods measure the shape of an object using multiple synchronized cameras, they are suitable for use in capturing texture and shape simultaneously. Moreover, it is possible to capture data at a high frame rate because all the input data is obtained at the camera frame rate. Although the accuracy and stability of the measured shape is low compared to active methods; it is sufficient to produce the 3D model for computer graphics. One drawback of a visual hull is that a large system is required since many cameras are necessary to observe an object from various viewpoints.

For active methods, laser range finders and structured-light methods are commonly used. Especially, a laser range finder is often used for an accurate measurement of a shape. Laser range finders measure the distance between the sensor and an object based on triangulation or time-of-flight. A time-of-flight sensor that changes the direction of a laser beam by controlling a mirror cannot capture shapes at a high frame rate, because each change takes several seconds or minutes. Although fast range finders [?, ?] have recently been developed, their resolution and accuracy are not sufficient due to the trade-off between time and accuracy.

Structured-light methods cast light patterns on an object using a projector and a reflected patterns are observed by a camera. This technique determines the corresponding coordinates between the projector and the structured-light patterns observed by the camera. The distance from the camera to the object is computed by triangulation based on the correspondences. Some fast methods [?, ?, ?, ?] are proposed based on structured-light methods and they can measure shape at a high frame rate. Especially [?, ?] are suitable for real-time acquisition because they can measure a shape from a single image. However, these methods have difficulty in obtaining texture simultaneously because visible light patterns are projected to measure the shape. Two solutions are proposed for the problem: time-sharing methods [?, ?, ?, ?, ?, ?] and band-separating methods [?, ?, ?].

The time-sharing method uses structured light not only for capturing shape but also for illuminating objects to obtain texture. If the structured light patterns are removed from the images, they can be used as texture images. This is accomplished by computing the average of the images of time-multiplexing patterns, or low-pass filtering of fringe patterns. The advantage of this approach is that texture and shape can be obtained using the same camera. However, since structured light is used for both shape and texture, other illumination in the environment cannot be used. Consequently, the illumination in which these methods are applicable is restricted. Additionally, since the duration of shape and texture capture are different, misalignments between texture and shape may appear if motion is fast.

The band-separating methods uses different light wavelengths for shape and texture. The shape is obtained by using near-IR light, while the texture image is captured with visible light. Since the structured light patterns for NIR light are not detected by the camera for visible light, the acquisition of shape and texture can be accomplished simultaneously. Frueh and Zakhor [?] use vertical stripes and a horizontal line of IR. The vertical stripes are used for shape reconstruction and the horizontal line for identifying the vertical stripes. Hiura et al. [?] proposed a laser range finder that obtains a range image at 30Hz by fast scanning of IR light stripes. A texture image is captured simultaneously with the system. Microsoft Kinect [?] can capture a shape by using IR projector and camera and at the same time a texture by another CCD camera. Although there is a similarity between Kinect and ours such as shape reconstruction from a single image, there are two large differences exist; *i.e.* accuracy of our system is much higher than Kinect and a shape and texture has a single viewpoint with our system.

3. Acquisition of texture 3D shapes by one-shot structured light method

3.1. Overview of the proposed system

The proposed system to capture a 3D shape with texture consists of an IR projector and a multi-band camera as shown in Fig. 1. The IR projector casts a fixed grid pattern that consists of vertical and horizontal lines. The lines have two different patterns using two near-IR lights of 765nm and 850nm, respectively. The multi-band camera acquires images of both visible and IR wavelength from a single viewpoint. For capturing a texture, an external light source is used for illumination. If the external light source contains an IR component, it is cut-off by filter to avoid interference when capturing the shape. The 3D reconstruction from the grid pattern is briefly described in the next section.
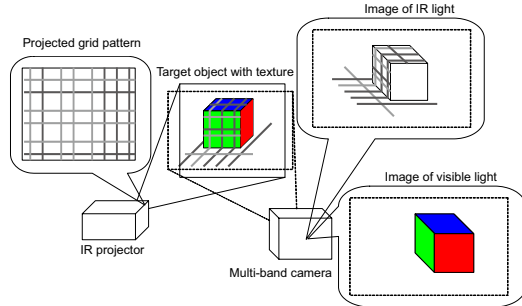
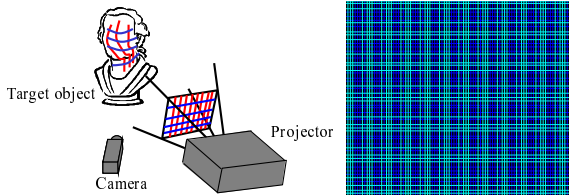Figure 1. The proposed system to capture textured 3D shape consists of an IR projector and a multi-band camera.

Figure 2. (Left) Scanning system: multiple lines are projected and their intersections are detected and used for reconstruction. (Right) projected pattern.

3.2. One-shot 3D reconstruction based on oneshot grid pattern

The proposed 3D measurement system consists of an infrared camera and an infrared projector as shown in Fig. 2 (left). The camera and the projector are assumed to be calibrated (*i.e.*, the intrinsic parameters of the devices and their relative positions and orientations are known). The projected pattern is fixed and does not change, so no synchronization is required. A grid pattern of vertical and horizontal lines is projected from the projector and captured by the camera.

The flow of the reconstruction algorithm is shown in Fig. 3. The algorithm is based on the work by Sagawa *et al*. [**?**]. In the first step, the curves illuminated by the grid lines are detected from the infrared image. In the second step, the intersection points of the detected curves (grid points) are extracted. Each line emitted from the projector sweeps a plane in the 3D space. This plane is called a pattern plane in this paper. From the extracted grid points, simple linear constraints about the crossing pattern planes can be acquired. In the third step, by solving the above linear equations generated from the grid points, solutions of the pattern planes are obtained. In the solution, there remains 1-DOF ambiguity which cannot be solved from information of the grid points [**?**]. In the fourth step, the ambiguity is solved from estimation of matching between the 1-DOF solution and the actual positions of the pattern planes obtained from the projected pattern and the calibration results. Intuitively, this matching process is equivalent to finding correspondences of the grid points using epipolar constraints that are consistent for the whole set of connected grid points, In the final
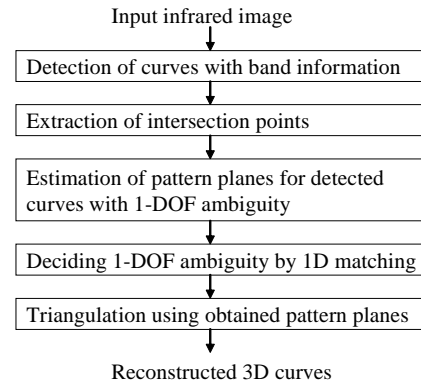
Figure 3. Flow of the reconstruction algorithm.

step, the 3D information of the detected curves are calculated by triangulation using the obtained pattern planes.

In the matching process, additional information attributed to each line pattern is effective for avoiding ambiguous matching, where multiple candidates of the solution remains. In this paper, we use 'color' information (*i.e.*, wavelength band information) of the infrared light. To achieve this, the grid lines are colored by a periodic (thus, not globally unique) coding of de Bruijn sequence. In the reconstruction process, the periodic IDs are decoded and assigned to each detected curves, and the IDs are compared to the projected patterns at the matching stage.

Once the correspondence between a line of the projected pattern and a curve in the camera image is determined, the 3D positions of points in a curve is computed. To texture-map an image of visible light, the shape between lines is linearly interpolated. Since the texture image is acquired from the same viewpoint as described in the next section, the occlusion problem does not occur in the texture mapping process in the proposed system.

4. Proposed system for capturing texture and shape

4.1. Multi-band camera to acquire visible and IR images from a single viewpoints

To generate a textured 3D shape, it is desirable to acquire texture and shape from the same viewpoint because the misalignment between shape and texture can occur due to the error of shape if they are capture from different positions. For solution, the proposed camera has imaging devices for visible and IR lights that are placed coaxially.

Fig. 4 shows an experimental multi-band camera system. The object lens consists of a general C-mount lens and a relay lens. Different fields of view can be obtained by changing the C-mount lens. The relay lens has a long back focus so that incoming ray focuses on the surface of imaging sensors after going through the optical components shown in Fig. 4(b).
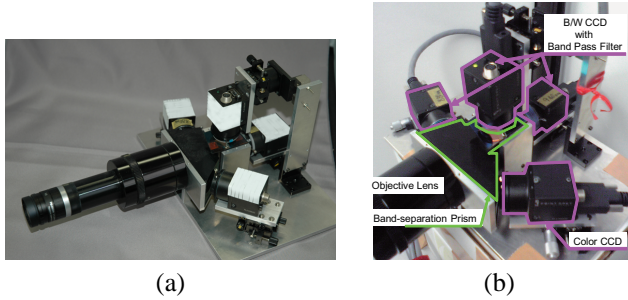
(a)   (b)

Figure 4. (a) an experimental system of the proposed multi-band camera. (b) the zoom up of the optical components. The object lens consists of a general C-mount lens and a relay lens. Different field of view can be obtained by changing the C-mount lens.
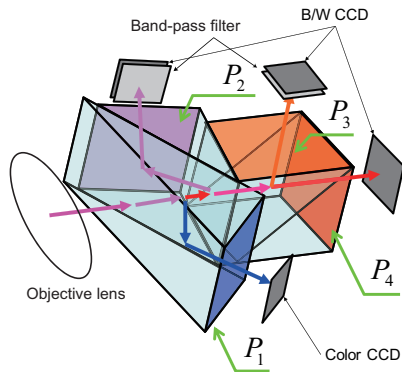

Figure 5. The internal structure of the multi-band camera in the proposed system: The camera consists of an objective lens, a band-separation prism, and image sensors.
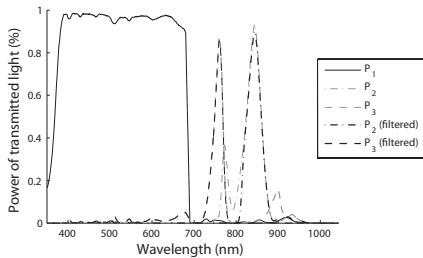

Figure 6. More than 80% of the light power is transmitted to desirable exit faces when the IR light shown in Fig. 9 and white visible light enter from the objective lens.

Fig. 5 shows the internal structure of the multi-band camera. The optical components are composed of an objective lens, a band-separation prism, and imaging sensors. The incoming light passing through the objective lens is first split into visible, 765nm and 850nm IR light by the band-separation prism; a Philips prism that consists of four glass blocks, of which interfaces are coated with membranes to separate different bands of light. Although the prism can separate out 935nm light from the other wavelength, 935nm light is not used in this paper because two wavelengths are sufficient for the current encoding of patterns. The wavelength will be used if the pattern is encoded by using three wavelengths. The exit faces P_1, P_2, P_3 and P_4 in Fig. 5
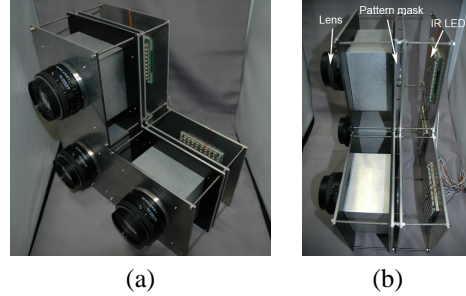

(a)   (b)

Figure 7. The new experimental IR projector system with three arrays of infrared LEDs, pattern masks printed on transparent films, and objective lenses.

corresponds to visible, 850nm, 765nm, and 935nm, respectively.

Fig. 6 shows the power of light transmitted to each exit face when the IR light cast from the IR projector and white visible light from external light sources enter from the objective lens. P_1, P_2 and P_3 indicates the prism faces illustrated in Fig. 5. Since the results of P_2 and P_3 have crosstalk each other, band-pass filters are used to remove this. The efficiency of the light power is more than 80%. The optical path lengths from the objective lens to the four image sensors are set to be equivalent so that the rays from the same point focuses on the imaging sensors. Two monochrome cameras are used to obtain IR images as imaging sensors, and a color camera is used for visible light in this system.

Akasaka *et al.* [?] proposed a similar system that captures IR and visible lights from the same viewpoints. The disadvantage of their system was inefficiency in terms of light energy. Since a simple beam splitter that split incoming light regardless of its wavelength was used in the system, the power of the structured light was reduced to less than half of the power discharged from the projector when splitting and combining light. Consequently, it needs strong light source to acquire image sequence at high frame rate. The proposed system has better performance in terms of the efficiency of light power by modifying the optical components of the camera, and accomplish to capture moving objects at high frame rate.

4.2. Infrared projector to cast grid pattern

To design an IR projector for 3D reconstruction, there are two requirements exist as follows: 1) the projector casts a fixed grid pattern that consists of vertical and horizontal lines. 2) two IR wavelengths are required to encode the pattern. One idea to realize them is to modify a video 3-LCD projector commercially available to cast IR lights instead of visible lights. However, since the optical components of the projectors are specialized for visible light and usually cut-off filter for IR is installed, it is not possible to modify a video projector for our purpose.
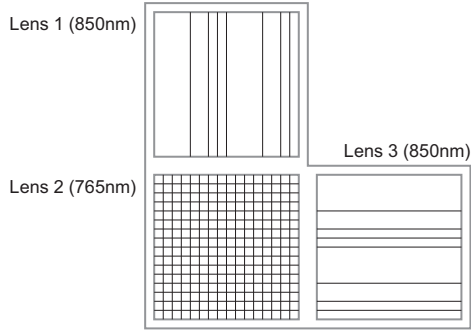
Figure 8. An example of the projected grid pattern is formed by three line patterns that corresponds to the three lenses.

Because the pattern used in the proposed system is static, it is not necessary to change it during scanning process. Therefore, a fixed pattern mask printed on transparent film is sufficient for the system instead of LCD filters used in video projectors.

Another issue is to cast two IR light of different wavelengths. a usual LCD projectors combines lights of different colors by a combination prism before casting through the objective lens. In the proposed system, however, the pattern is simply composed of vertical and horizontal lines. Therefore, it is possible to form the pattern without combining rays by a prism. The proposed system casts lines by using three lenses placed like L-shape as shown in Fig. 7.

An IR LED array and pattern mask are placed behind each lens. Fig. 8 shows an example of the pattern masks for three lenses. 765nm light is cast through the lens 2, and 850nm light is through the lens 1 and 3. If lines cast from the lenses 1 and 2 are parallel to the direction from the lens 1 to lens 2, the pattern planes cast from these lenses are the same. Thus, the lines are encoded by two wavelengths of IR light without combining them before emitting from the objective lenses. This approach contributes the efficiency of light power. The patterns of 850nm are determined based on de Bruijn sequence and aligned to that of 750nm.

Each IR LED array used as a light source has 116 LEDs and the power of each array is 0.96-1.20W, which is much smaller than the light sources used in commercial video projectors. Fig. 9 shows the power of discharged light from each lens, of which the maximum is normalized to one. Since the bandwidth of LEDs is quite narrow, the crosstalk between different wavelengths does not occur; crosstalk is one of the main issue of image processing when a visible light source such as a commercial video projector is used.

4.3. System calibration

4.3.1 Calibration between image sensors

For texture mapping, the texture image must be geometrically calibrated with the system to acquire the shape. If the texture is captured from a viewpoint different from the cam-
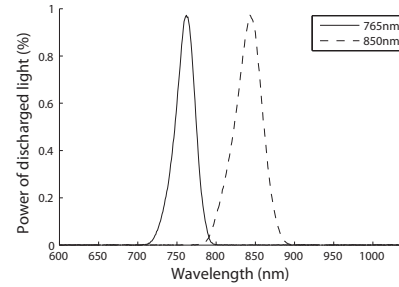

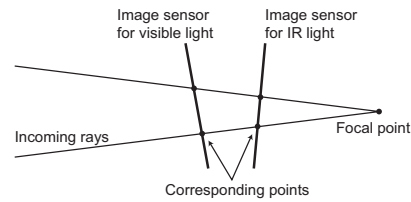Figure 9. The power of discharged light from each lens: the maximum of each band is normalized to one.

Figure 10. The virtual optical path are common for all image sensors. Due to the misalignments, an incoming ray is projected on the different pixel of the image sensors.

era to obtain the shape, the corresponding point between the texture image and shape cannot be determined a priori. On the other hand, since the proposed system can obtain texture and shape from the same viewpoint, the corresponding points are determined by calibrating the geometry between the image sensors.

In the proposed multi-band camera, although the different wavelengths are separated by a prism, the optical structure for all image sensors is virtually the same except the positions of sensors relative to the other optical components. Since the positions of the image sensors have some small misalignments relative to the other optical components, an incoming ray are not projected on the same pixel of image sensors. The situation is illustrated in Fig. 10. Since the virtual optical path are common for all image sensors, the calibration to compute the corresponding points can be done by estimating the homography between the image sensors.

Estimating a homography is similar to the method used in the camera calibration. First, the image of a planar object, such as a flat plate with checker board pattern, is captured by all sensors. By detecting four or more feature points on the object, the homography can be estimated [?].

4.3.2 Calibration of projector-camera system

To calibrate a projector-camera system, the correspondence of the coordinates between the projector and the camera is required. It can be accomplished by capturing coded images projected from the projector. Time-multiplexed pattern, such as a pattern based on Gray code, is often used to obtain the absolute coordinates of the projector by observ-
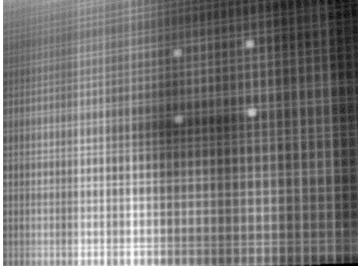
Figure 11. The pattern of one-shot 3D reconstruction is used for calibration. The absolute coordinate of the projector is manually given by using the marker in the pattern.

ing the multiple coded images of the structured light. However, since the proposed system uses a fixed pattern printed on a transparent film, it is not easy to use time-multiplexed pattern by changing the pattern mask. Thus, we utilize the pattern of one-shot 3D reconstruction for calibrating the system. Since the pattern is a periodic pattern, the absolute coordinate of the projector is not automatically determined. For solution, we manually decide the first point to determine the absolute coordinate. Then, remaining points are automatically determined by using the connecting information of curves; those curves are automatically detected by efficient curve detection algorithm proposed in [?]. To ease the manual process to decide the first point, some markers are added on the pattern as shown in Fig. 11. We make the markers to be small so that they do not interfere the curve detection algorithm.

Once the corresponding coordinates between the camera and projector are obtained, the intrinsic of the camera and projector, and extrinsic parameters between them can be estimated as calibration of a stereo system. In this paper, an implementation based on bundle adjustment provided by Snavely [?] is used for calibration. If both calibration of the system and 3D reconstruction is accomplished only by a fixed pattern, we can simplify the projector drastically compared to commercial video projectors, which need optical devices to generate a pattern by LCD or DMD.

5. Experiments

We first estimate the accuracy of the experimental system by measuring a known object. The shape of a cube-shaped object shown in Fig. 12 is acquired. Fig. 12(a), (b) and (c) are the images of visible, 765nm IR, and 850nm IR lights, respectively. Fig. 12(d) is the grid pattern generated by integrating the two IR images. Blue and green lines correspond to 765nm and 850nm lights. The images are aligned each other by using homographies estimated by the method described in Sec. 4.3.1. The cube is 0.3m on a side and the distance from the camera is about 0.8m. There is no crosstalk between visible, 765nm and 850nm bands thanks to the prism of band separation.
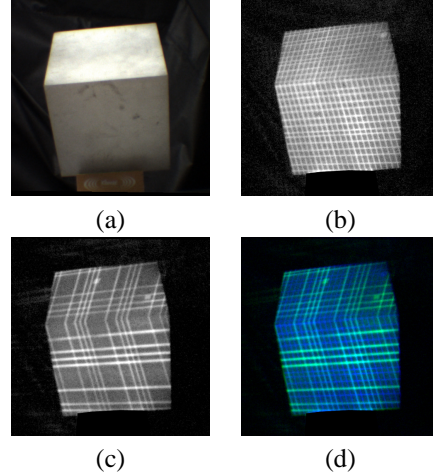

Figure 12. Acquired images of a cube-shaped object: (a) color image of visible light (b) the projected pattern of 765nm wavelength (c) the projected pattern of 850nm wavelength (d) the grid pattern generated by combining 765nm and 850nm.
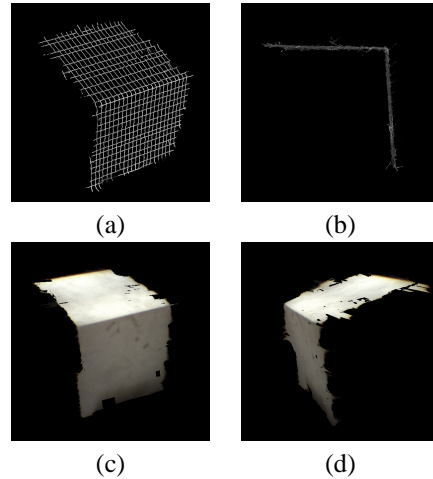

Figure 13. The reconstructed result of the cube-shaped object. (a) and (b) are the points reconstructed for the pixels on the curves in the input image. (c) and (d) are the textured shape.

Fig. 13 shows the reconstructed result of the cube. (a) and (b) are the points reconstructed for the pixels on the curves in the input image. A mesh model is generated after interpolating vertices, and (c) is the textured shape to the mesh model. To evaluate the accuracy of the shape, the two faces of the cube are fit to planes. The root mean square error (RMSE) of the points from the fit planes is 0.48mm. The angle of the two planes is 84.8 degrees (ground truth 90.0). Since our calibration is based only on bundle adjustment, the intrinsic and extrinsic parameters are estimated simultaneously; and thus, it may be a main reason for an angle difference. In our future work, the intrinsic parameters should be calibrated separately to improve the accuracy of the angular error.

Next, we show example comparison with Kinect and our method. Fig. 14 is the result of 3D reconstruction with tex-

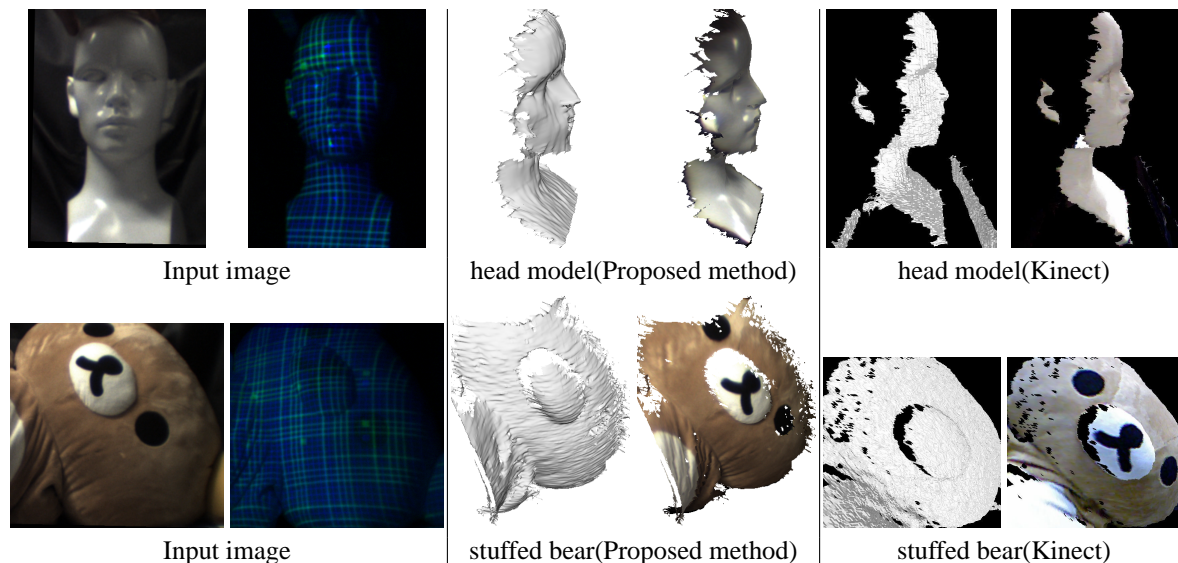| Input image	head model(Proposed method)	head model(Kinect)
Input image	stuffed bear(Proposed method)	stuffed bear(Kinect)

Figure 14. Comparison with Kinect and proposed method on 3D reconstruction with texture mapping of a head model and a stuffed bear. The figures are the image of visible light, the IR image, the shape without texture of proposed method, the textured 3D shape of proposed method, the shape without texture of Kinect, the textured 3D shape of Kinect, respectively.

ture mapping of a head model and a stuffed bear. The figures are the image of visible light, the IR image, the shape without texture of proposed method, the textured 3D shape of proposed method, the shape without texture of Kinect, the textured 3D shape of Kinect, respectively. Since the proposed method is based on light sectioning by using the grid pattern, the surfaces reconstructed by proposed method are smoother than Kinect using the dot pattern. The undulation on the surfaces is artifact caused by interpolation of patterns, which is an issue to be solved in future work.

Finally, we acquired the textured shapes of moving objects. Fig. 15 shows the input images and results of capturing a waving cloth. The images are captured at 30 frames/second. The top row shows texture images of visible light at each frame. The second row shows the images of IR structured light. The third row shows the results reconstructed without texture, and the fourth row show the shapes with texture mapping. The proposed system succeeded in capturing the shape of the moving object with texture mapping. If an object reflects IR light, the system can obtain the shape of black object in visible wavelength as shown in the result.

6. Conclusion

This paper proposes a projector-camera system that acquires textured 3D shapes based on one-shot 3D reconstruction. Since the proposed method obtain texture and shape at the same time, it can capture the shape with texture of moving objects. To accomplish simultaneous capturing texture and shape, we developed a system consisting of a projector of infrared structured light and a multi-band camera. Since the multi-band camera can observe IR and visible lights from the same viewpoint, we can accurately map the texture without affected by the error of 3D reconstruction. The projector proposed in this paper casts structured light coded by two IR wavelengths. Since the pattern is composed simply by lines perpendicular each other, the pattern is realized by multiple-lens system, which contributes to improve the efficiency of light power. Additionally, efficient calibration technique of projector-camera system by using a fixed line pattern of one-shot 3D reconstruction is proposed. With the proposed technique, only a few manual step is required for calibration, which contributes to simplify the system. In the experiments, we tested the accuracy of the proposed system and showed that it can obtain textured 3D shapes of moving objects. In future, we plan to brush up the optical design and improve the quality of the results.

References

[1] K. Akasaka, R. Sagawa, and Y. Yagi. A sensor for simultaneously capturing texture and shape by projecting structured infrared light. In *Proc. The 6th International Conference on 3-D Digital Imaging and Modeling*, pages 375–381, Montreal, Canada, August 21-23 2007.

[2] Canesta, Inc. CanestaVision EP Development Kit. http://www.canesta.com/devkit.htm.

[3] C. Frueh and A. Zakhor. Capturing 21/2d depth and texture of time-varying scenes using structured infrared light. In *Proc. the 5th International Conference on 3-D Digital Imaging and Modeling*, pages 318–325, 2005.

[4] S. Hiura, A. Yamaguchi, K. Sato, and S. Inokuchi. Real-time object tracking by rotating range sensor. In *Proc. 13th*
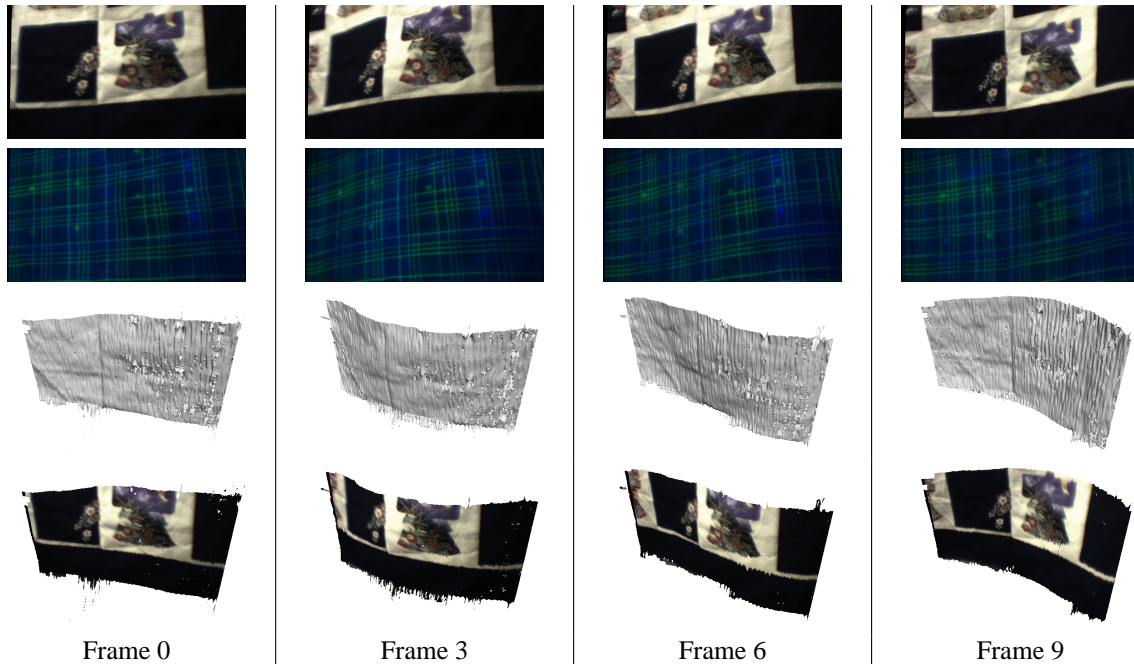
Frame 0   Frame 3   Frame 6   Frame 9

Figure 15. The texture and shape of a waving cloth is acquired. The top row shows texture images of visible light. The second row shows the images of IR structured light. The third and fourth rows show the reconstructed shape without/with texture, respectively.

International Conference on Pattern Recognition, volume 1, pages 825–829, 1996.

[5] C. Je, S. Lee, and R. Park. High-contrast color stripe pattern for rapid structured-light range imaging. In *Proc. Eighth European Conference Computer Vision*, pages 95–107, 2004.

[6] H. Kawasaki, R. Furukawa, , R. Sagawa, and Y. Yagi. Dynamic scene shape reconstruction using a single structured light pattern. In *CVPR*, pages 1–8, June 23-28 2008.

[7] T. Koninckx and L. V. Gool. Real-time range acquisition by adaptive structured light. *IEEE Transaction Pattern Analysis Machine Intelligence*, 28(3):432–445, 2006.

[8] Mesa Imaging AG. SwissRanger SR-4000. http://www.swissranger .ch/index.php.

[9] Microsoft. Xbox 360 Kinect. http://www.xbox.com/en-US/kinect.

[10] P. Narayanan, P. Rander, and T. Kanade. Constructing virtual worlds using dense stereo. In *Proceedings of the Sixth IEEE International Conference on ComputerVision (ICCV'98)*, pages 3–10, January 1998.

[11] J. Pan, P. S. Huang, and F.-P. Chiang. Color-coded binary fringe projection technique for 3-D shape measurement. *Optical Engineering*, 44:3606–+, Feb. 2005.

[12] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future : A unified approach to image-based modeling and spatially immersive displays. In *Proc. ACM SIGGRAPH*, 1998.

[13] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. Real-time 3d model acquisition. In *Proc. SIGGRAPH*, pages 438–446, 2002.

[14] R. Sagawa, Y. Ota, Y. Yagi, R. Furukawa, N. Asada, and H. Kawasaki. Dense 3d reconstruction method using a single pattern for fast moving object. In *ICCV*, 2009.

[15] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring image collections in 3d. In *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*, 2006.

[16] T.Matsuyama, X. Wu, T. Takai, and T. Wada. Real-time dynamic 3d object shape reconstruction and high-fidelity texture mapping for 3d video. *IEEE Trans. on Circuits and Systems for Video Technology*, CSVT-14(3):357–369, 3 2004.

[17] M. Waschbusch, S. Wurmlin, D. Cotting, F. Sadlo, and M. Gross. Scalable 3d video of dynamic scenes. In *The Visual Computer (Proceedings of Pacific Graphics 2005)*. Springer, 2005.

[18] T. Weise, B. Leibe, and L. V. Gool. Fast 3d scanning with automatic motion compensation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[19] L. Zhang, B. Curless, and S. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *Proc. First International Symposium 3D Data Processing Visualization and Transmission*, pages 24–36, 2002.

[20] S. Zhang and P. Huang. High-resolution, real-time 3d shape acquisition. In *Proc. Conference on Computer Vision and Pattern Recognition Workshop*, page 28, 2004.

[21] Z. Zhang. A flexible new technique for camera calibration. In *IEEE Transactions on Pattern Analysis and Machine Intellgence*, volume 22, pages 1330–1334, 2000.

[22] Z. Zhang, D. Zhang, X. Peng, and X. Hu. Color texture extraction from fringe image based on full-field projection. *Optical Engineering*, 42:1935–1939, July 2003.