One-shot 3D Reconstruction of Moving Objects by Projecting Wave Grid Pattern with Diffractive Optical Element

Ryusuke Sagawa[1], Tatsuya Kawamura[2], Ryo Furukawa[3], Hiroshi Kawasaki[4] and Yoshio Matsumoto[1]

[1] *Intelligent Systems Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8563, Japan*

[2] *ALT INC., 1-21-10 Toyotama Minami, Nerima-ku, Tokyo 176-0014, Japan*

[3] *The Graduate School of Information Science, Hiroshima City University, 3-4-1 Ozuka-Higashi, Asa-Minami-ku, Hiroshima 731-3194, Japan*

[4] *The Department of Information and Biomedical Engineering, Kagoshima University, 1-21-40 Korimoto, Kagoshima 890-0065, Japan*

Abstract:

Recently, the 3D measurement of moving objects becomes an important task in various applications. The potential applications are medical application, multimedia, crash testing, fluid analysis and so on. For these applications, the 3D measurement is required to have properties of high accuracy, high frame-rate, and high density. We have proposed a method based on a projector-camera system that reconstructs a shape from a single image where a static pattern is cast by a projector. The proposed method realizes one-shot 3D reconstruction with a single-colored pattern that consists of vertical and horizontal sinusoidal curves. It finds correspondences in sub-pixel accuracy between projector and camera by using the information implicitly encoded in the grid of wave lines. The reconstructed shape is optimized by interpolating grid patterns to reconstruct dense shape to achieve the above requirements. One of the issues to improve this system is to realize 3D reconstruction under external light that disturbs detection of the pattern. We have developed a laser system to generate the wave grid pattern that can be detected by filtering out external light. The pattern is formed by using diffractive optical element to generate the static pattern power-efficiently without temporal scanning. The power-efficiency can contribute to downsizing of the system. We apply the same approach to reconstruct 3D shapes for medical endoscopy that equips the pattern projector at the tip of the probe. In the experiments, we show the effectiveness of the proposed laser pattern projector by reconstructing 3D shapes.

Keywords: 3D Reconstruction of Moving Objects, One-shot Reconstruction, Wave Grid Pattern, Pattern Generation by Diffractive Optical Element

1. INTRODUCTION

The application of capturing the shapes of moving objects is rapidly increasing. For example, gaming device [1] has become very popular to capture human motion in real-time and realize a device-free interface. A vision system for autonomous vehicles and robots [2, 3] are also ones of the promising application of shape capturing. One of the issues of the scanners for capturing moving objects is that the accuracy and density are lower than the range sensors for static objects. If high accuracy and resolution are realized, they should be more useful for various purposes, e.g., medical application, fluid analysis and so on.

Various methods have been proposed for capturing moving objects, such as stereo methods or time-of-flight (TOF) methods. Especially, structured-light stereo methods are suitable for capturing moving objects and have been widely researched [1, 4, 5].

We have proposed a one-shot scanning method [6] that reconstruct the shape of an object from a single image by casting structured light. Since the method requires only one frame of an image sequence, the shape of the moving object can be reconstructed frame by frame. The pattern used in the method is a wave-shaped grid pattern so that the intersection points of vertical and horizontal sinusoidal curves can be used as features for matching. Instead of explicitly encoding the positional information of a structured light, the proposed pattern implicitly gives information which can make the order on the candidates of corresponding points.

The proposed pattern is a single-colored static pattern, which increases the stability on the colors of the target objects. This feature has additional advantage to improve the system. While using an off-the-shelf video projector is the easiest way to cast the structured-light pattern, wider range of projectors can be used thanks to the static pattern of single color. In this paper, we have developed a laser pattern projector for the structured-light system. The advantages are: 1. the laser projector of single narrow-band light can be used to reduce the interferences from the ambient lights, 2. the laser pattern projector can contribute improving power-efficiency and downsizing of the system.

In the following sections, we first explain the proposed structured-light system with wave grid pattern briefly. Second, the proposed laser projector is described, and we show the results with the projector.

2. RELATED WORK

TOF systems and triangulation based methods (e.g., light-sectioning method or stereo method) are widely known for active measurement systems. To capture dynamic scenes with the active systems, methods based on both approaches have been researched.

In many TOF laser scanning systems, a point laser is projected and the interval time between the emission and the detection time is measured. Since the 3D information is obtained one point at a time, it is unsuitable for capturing a entire scene in a short period of time. To capture dynamic scenes, some TOF devices project temporally-modulated light patterns and acquire a depth image at once by capturing the reflections with many detectors on a 2D image sensor [7]. However, the present systems are easily disturbed by other light sources and the resolution is lower than normal cameras.

There are two types of the methods in structured-light methods. One is temporal-encoding methods and the other is spatial-encoding methods. Since a spatial-encoding method just requires a single input for reconstruction (a.k.a. one-shot scan), it is ideal to capture moving objects with high FPS. Therefore, many researches have been involved in spatial-encoding methods [8]. However, since they require certain areas to encode information on object surfaces, the resolution tends to be low and reconstruction becomes unstable.

One of the approaches to encode information in efficient ways is to use a color code. By using multiple colors, multiple bits of information can be assigned to each pixel of the camera image. A color-based coding is suitable for spatial encoding [4, 8, 9, 10]. It, however, has some limitations and problems. The surface of the target objects must sufficiently reflect each color of the pattern. And, since the RGBs of off-the-shelf video projectors have overlapped spectral distribution, errors in determining colors of pixels are inevitable.

To avoid those problems, several methods are proposed for efficient spatial encoding without using colors, such as dot patterns [1] or grid patterns [6]. The advantage of these methods is to enable the simplification of pattern projector. Although the easiest way to cast a pattern is to use a video projector, it is difficult to apply wide range of applications because a video projector is expensive, large and low power-efficiency. If the pattern is static and single-colored, it can be generated by a simple device. We therefore propose a laser pattern projector in this paper to solve the above issues.

3. A STRUCTURED LIGHT SYSTEM WITH WAVE GRID PATTERN

Our system consists of a single projector and a camera as shown in Fig. 1 . The projector casts a static pattern which is shown in Fig. 2 . The pattern is configured with vertical and horizontal sinusoidal curves to create grid shape. Since the pattern is static with single color, no synchronization is required between the projector and the camera.

We first detect curves from a captured image. We use the curve detection method using belief propagation method

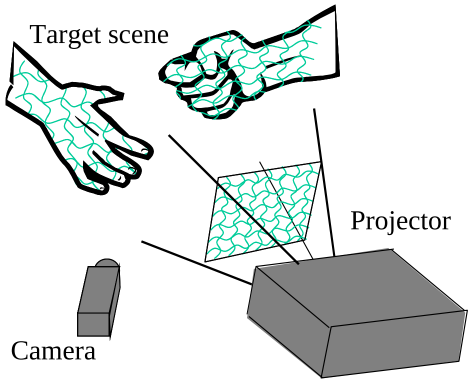

Fig. 1 Scanning system: vertical and horizontal wave lines are projected and their intersections are detected and used for reconstruction.

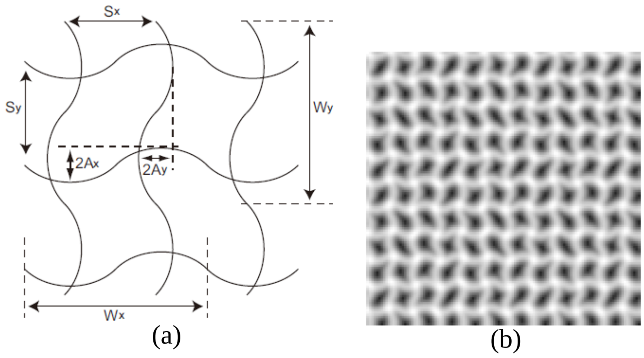

Fig. 2 Parameters of wave grid are shown in (a). S_x and S_y are the intervals between adjacent wave lines, W_x and W_y are the wavelengths of a wave line, A_x and A_y are the amplitudes of waves with respect to vertical and horizontal lines, respectively. (b) is an example of wave grid.

proposed in [11]. With the method, vertical and horizontal lines are robustly detected from a grid pattern from a single color. From the detected curves, intersection points are calculated, and a graph is also constructed by using intersection points as nodes of the graph. Then, for each intersection point, epipolar line on the projected pattern is calculated to find a correspondence. Since multiple candidates of correspondences are usually found, one solution can be determined by our belief propagation based technique. Since the grid pattern captured in a camera image is sparse, the correspondence information is interpolated and optimized by image matching between the camera image and the projector pattern for all the pixels to reconstruct dense 3D shapes.

2.1 Wave Grid Pattern

The wave line is a sinusoidal pattern, which is periodic and self-recurring. The grid of wave lines, however, can give information for finding correspondences. The proposed method uses the intersection points of vertical and horizontal wave lines as feature points. The arrangement of intersection points is determined by the intervals and the wavelength of the wave lines. In the paper, we use the same interval and wavelength for all the vertical and horizontal wave lines.

However, as described in the following, because the interval of the vertical wave lines is not equal to the integral

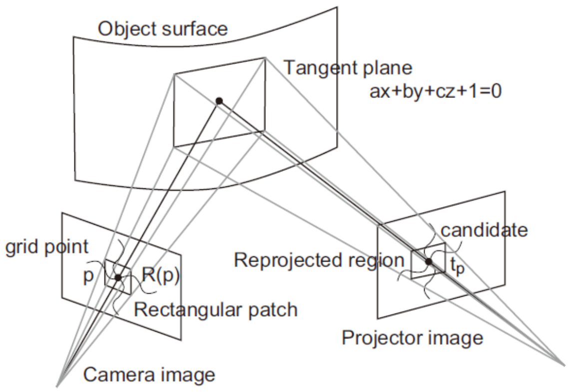

Fig. 3 The rectangular patch around a grid point p is reprojected onto the projector's image plane. t_p is one of candidates of correspondence for p.

multiple of the horizontal wavelength, the intersection points appear at the different phases on the wave pattern; it means that the local pattern around an intersection point has local uniqueness, and it can be used as a discriminative feature. In this paper, we also use `wave patterns' to refer to the wave lines.

The local pattern around an intersection point is not globally-unique in the whole pattern and periodic. Therefore, the same pattern occurs at every N_x and N_y wave lines along the horizontal and vertical axes, where N_x = lcm(S_x, W_x)/S_x, N_y = lcm(S_y, W_y)/S_y where lcm(a, b) is the least common multiple of a and b. Hereafter, subscript letter x means the symbol describes values about horizontal axis, and y about vertical axis. S_x and S_y are the intervals between adjacent wave lines, and W_x and W_y are the wavelengths, as shown in Fig. 2 (a). The patterns, however, can be discriminative in each cycle. Fig. 2 (b) shows an example with $S_x = 10$, $S_y = 11$, $W_x = W_y = 14$, $A_x = A_y = 1$ (pixels), where A_x and A_y are the amplitudes of waves. In this case, each cycle has 7 and 14 wave lines along horizontal and vertical axes, respectively. Consequently, 98 (=7 x 14) intersection points of different patterns exist in the rectangle.

2.2 Finding Correspondence between Projector Pattern and Camera image

The propose method uses two types of information to find the correspondence between the projector pattern and the camera image. The first one is the similarity of the images of the local area around an intersection point. We calculate the similarity by the sum of squared distance (SSD) of two images. The rectangular patch around an intersection point is compared to the reprojected region in the projector pattern as shown in Fig. 3.

The second one is the connectivity of intersection points. If two intersection points in a camera image is connected by the detected curve, their corresponding points in the projector pattern should be on the same wave line. This information is the constraint to determine the correspondence. For this constraint to work properly, the wave line should not be parallel to the epipolar line in the projector image. If they are parallel, the connectivity gives no useful information.

The problem to find the correspondence for each intersection point is defined as the one to choose the one of candidates of correspondence on the epipolar line. Now, the detected wave lines and intersection points forms a grid graph. The intersection points $p \in V$ are connected to another point by $(p, q) \in U$, where V and U are the set of intersection points and their connections. A point p has the candidates of corresponding points $t_p \in T_p$ in the projector image. We define the energy to find correspondences as follows:

$$E(T) = \sum_{p \in V} D_p(t_p) + \sum_{(p,q) \in U} W_{pq}(t_p, t_q)$$

where $T = \{t_p | p \in V\}$. $D_p(t_p)$ is the data term of assigning a candidate t_p to p. W_{pq} (t_p, t_q) is the regularization term of assigning candidates t_p and t_q to neighboring intersection points.

The data term is the SSD of local patch around the intersection point between the projector and camera. The regularization term is defined as follows

$$W_{pq}(t_p, t_q) = \begin{cases} 0 & t_p \text{ and } t_q \text{ are on the same wave line} \\ \lambda & \text{otherwise,} \end{cases}$$

where λ is a user-defined constant. The energy is minimized based on belief propagation [12] in this paper.

The advantage of using energy minimization to enforce structures is that they can be ``soft constraints." This is important because there is always a chance that erroneous grid connections occur in actual case. With our method, removing wrong connections and 3D reconstruction are simultaneously accomplished to achieve dense and better result. This is because not only local features by wave pattern but also epipolar constraint are used for removing wrong connection.

The correspondences of the intersection points are sparse in the camera image. An additional step is accomplished to obtain dense correspondences by using all the pixels. We first calculate depth values for all pixels by interpolating the intersection points. Then, the depth values are optimized by minimizing the difference of intensity for all the pixels between camera and projector images.

3. GENERATING THE PATTERN BY DIFFRACTIVE OPTICAL ELEMENT

An easy way to generate the wave grid pattern is to use an off-the-shelf video projector, which can change the pattern controlling by using a PC. If we use a LCD projector, the pattern can be completely static while a DLP projector changes the pattern temporally. We therefore used a LCD projector because no synchronization was necessary. A typical LCD projector has a lamp bulb as the light source and the LCD masks to generate the pattern. Since more than 70% pixels are blocked by the mask, the large part of the light energy is discarded and the power-efficiency is low. Moreover, the size of a projector can be too large for some practical use, such as mounting it on a robot head.

In this paper, we have developed a laser pattern projector for the proposed structured-light system for improving the power-efficiency and downsizing the projector. Instead of generating the pattern by masking, we used diffractive optical element (DOE) to form the pattern, which works as a beam splitter that divides an input beam into a large number of output beams, which generally has about 80% efficiency of the input light power. The design of a DOE has some constraints to obtain good efficiency. For example, a DOE basically generates a pattern of rotation symmetry, and the

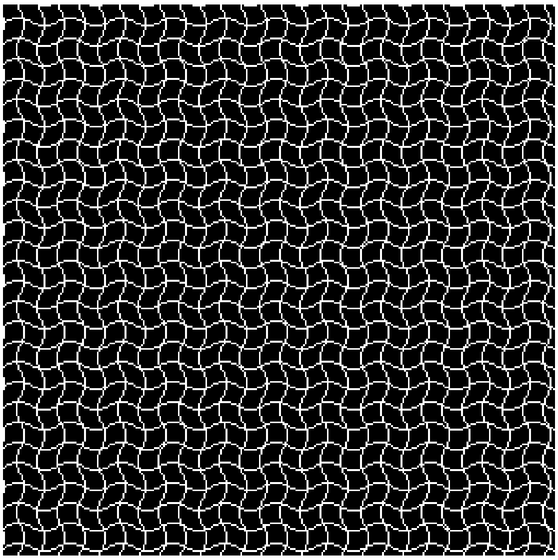

Fig. 4 The pattern is designed with the parameters, S_x = 11, S_y = 11, W_x = W_y = 28, A_x = A_y = 2. This figure shows a part of the pattern of 84 lines along each axis.

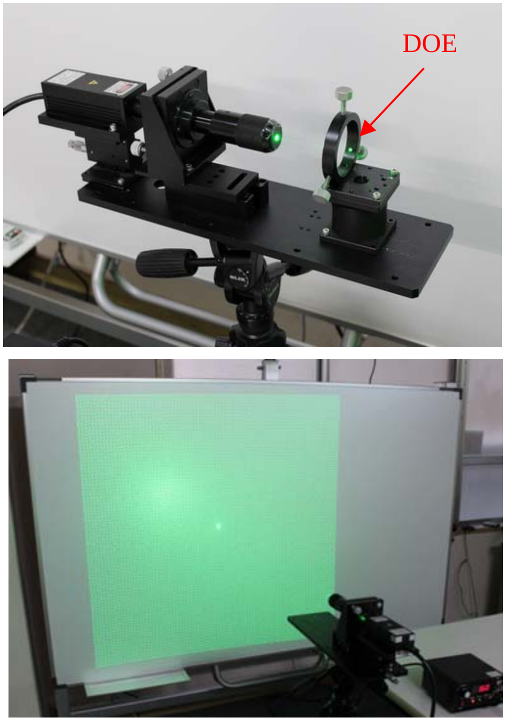

Fig. 5 The experimental pattern projector is developed by combining a 532nm DPSS laser light source and the designed DOE. The bright spot at the center of the screen is the zero order beam of the DOE.

pattern degrades if the number of spots increases. We tested several designs of DOE and determined the specification of the DOE and the light source as follows:

- Laser wavelength: 532nm
- Projection angle: 70 degrees (diagonal)
- Focusing depth range: 1m or larger
- Parameters of the pattern: S_x = 11, S_y = 11, W_x = W_y = 28, A_x = A_y = 2 (pixels)
- Number of wave lines: 84 lines
- Target S/N ratio: 10dB

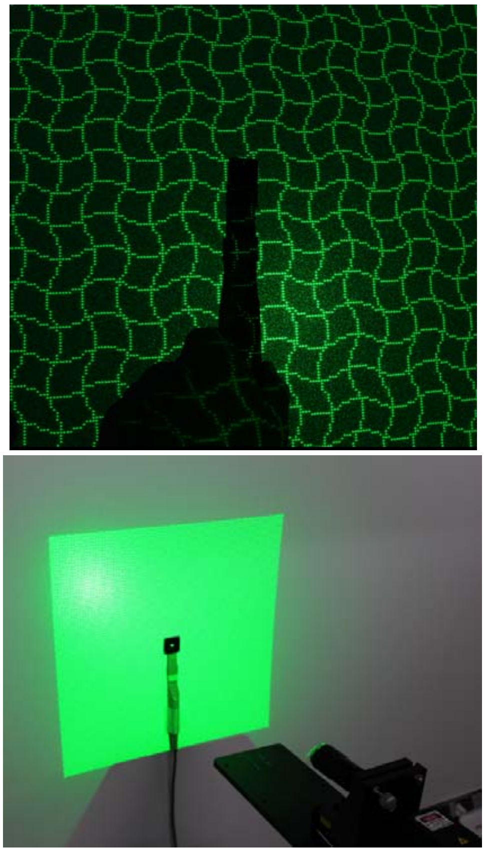

Fig. 6 Top: The S/N ratio is measured by the laser power meter. We compared the laser power on the wave line and that of the background noise. Bottom: The power of zero order beam is measured since the laser safety value of the system depends on the power of the zero order beam.

Since the pattern is represented by the set of dots, we used larger wavelength and amplitude of the wave lines than those of the original design so that the wave pattern is captured clearly in the camera image. Fig. 4 shows a part of the pattern. The total number of dots in the pattern is about 150K points.

Fig. 5 shows the experimental pattern projector developed by combining a 532nm DPSS laser light source of 200mW and the DOE manufactured by Nanocomp Oy Ltd. Although this experimental system has additional optical system, the system can be simplified only with the laser light source and the DOE with additional lenses. The pattern is successfully generated by the DOE, which has flat brightness and low distortion.

4. EXPERIMENTS

We first tested the pattern generated by the experimental pattern projector. Since we detect the lines in the camera image, the quality of the pattern is important. Fig. 6 shows the measurement of the S/N ratio by the laser power meter. We compared the laser power on the wave line and that of the background noise. Since the power on the wave line is 0.4μW and the background noise is 0.04μW, the S/N ratio is about 10dB, which satisfies the target value of the design.

Secondly, we tested the laser safety of the system. Since the laser safety value of the system depends on the power of

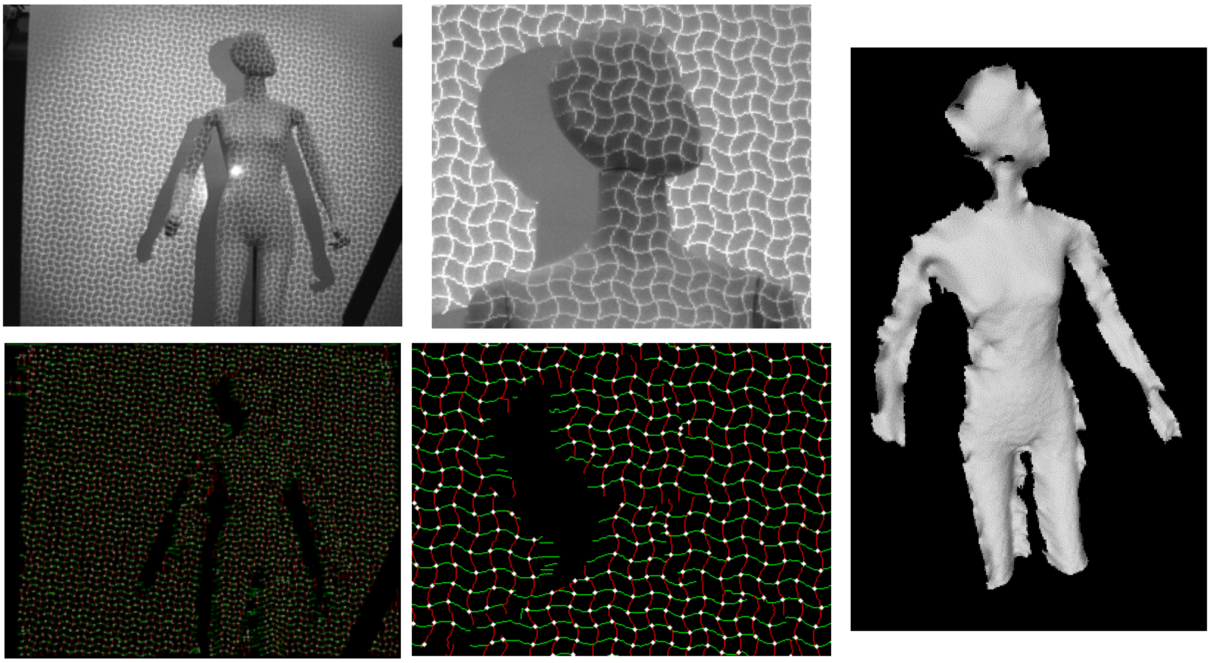

Fig. 7 The wave grid pattern is projected on a mannequin. The left column is the input image and the result of line detection. The middle column is the zoom-up view of the input image and the detected wave lines. The right column is the result of 3D reconstruction.

the zero order beam, we measured it by the laser power meter. The light source has 200mW. The diameter of the input laser beam is 5mm, and the laser is scattered to 5m wide at 5m distance. The laser power with 7mm aperture size becomes less than 1mW if the distance from the projector is larger than 0.4m. We will plan to use the cover to avoid exposure within the range to satisfy the condition of class 2.

Next, we tested 3D reconstruction by the proposed method with the experimental laser pattern projector. Fig. 7 shows an input image and the result of the method. The wave grid pattern is projected on a mannequin. The left column is the input image and the result of line detection. The middle column is the zoom-up view of the input image and the detected wave lines. The right column is the result of 3D reconstruction. The image was capture under a fluorescent light by the camera with the band-pass filter for the laser. The other wavelengths by the external light are filtered out. Since the signal of pattern is sufficiently larger than the noise, the wave lines are detected clearly. The method found the correct correspondences of the intersection points and generated the dense range scan by interpolating the correspondences by using all pixels.

In future work, we apply the proposed laser pattern projector to the medical endoscopic system to reconstruct the shape of objects. We have proposed an endoscopic system of 3D reconstruction [13], which generated the pattern with a mask. To increase the brightness of the pattern by improving the power-efficiency, we replace the mask with the DOE proposed in this paper. Fig. 8 shows the endoscopic system to reconstruct the 3D shape of objects and an example of the endoscopic camera with the pattern generated by the DOE. Our next step is to reconstruct the shape of objects from the images with the pattern.

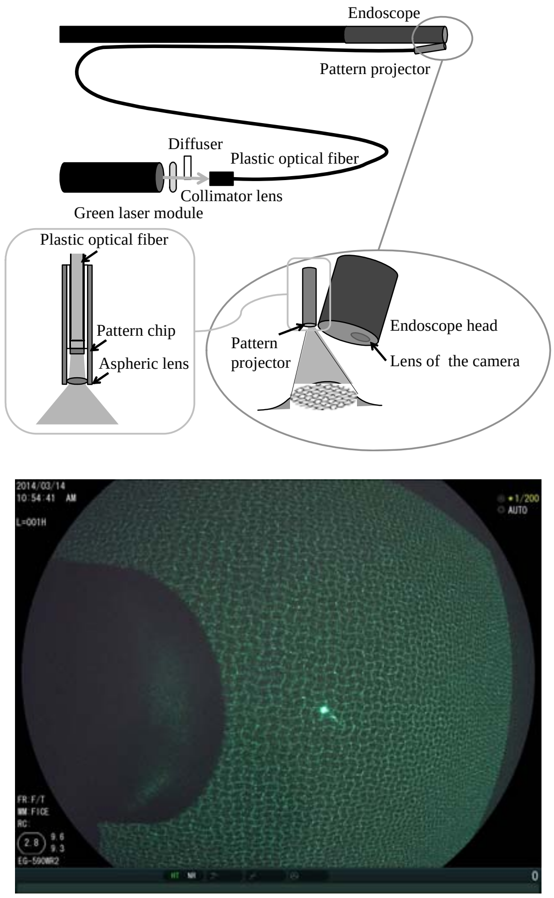

Fig. 8 The laser pattern projector is combined with the medical endoscopic system. The DOE is equipped at the tip of the optical fiber.

5. CONCLUSION

We have proposed a method based on a projector-camera system that reconstructs a shape from a single image where a static pattern is cast by a projector. One of the issues to improve this system is to realize 3D reconstruction under external light that disturbs detection of the pattern. We have developed a laser system to generate the wave grid pattern that can be detected by filtering out external light. The pattern is formed by using diffractive optical element to generate the static pattern power-efficiently without temporal scanning. The power-efficiency can contribute to downsizing of the system. The developed pattern projector is much smaller than an off-the-shelf video projector and we succeeded to reconstruct the shape from the input image captured under external light. We apply the same approach to reconstruct 3D shapes for medical endoscopy that equips the pattern projector at the tip of the probe, which shows this approach is appropriate for downsizing. In future work, we test the proposed method with the endoscopic system.

ACKNOWLEDGEMENTS

This work was supported in part by JST A-STEP No. AS2421093H and NEXT program No.LR030 in Japan.

REFERENCES

[1] Microsoft, "Xbox 360 Kinect," [Online]. Available: http://www.xbox.com/en-US/kinect.

[2] DARPA, *Urban Challenge*, 2007.

[3] DAPRA, *Robotics Challenge*, 2013.

[4] R. Sagawa, H. Kawasaki, R. Furukawa and S. Kiyota, "Dense one-shot 3d reconstruction by detecting continuous regions with parallel line projection," ICCV, 2011.

[5] H. Kawasaki, R. Furukawa, R. Sagawa and Y. Yagi, "Dynamic scene shape reconstruction using a single structured light pattern," CVPR, 2008.

[6] R. Sagawa, K. Sakashita, N. Kasuya, H. Kawasaki, R. Furukawa and Y. Yagi, "Grid-based Active Stereo with Single-colored Wave Pattern for Dense One-shot 3D Scan," 3DIM/3DPVT, 2012.

[7] Microsoft, "Kinect for Windows v2 sensor," [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/.

[8] J. Salvi, J. Pages and J. Batlle, "Pattern codification strategies in structured light systems," Pattern Recognition, 2004.

[9] C. Je, S. W. Lee and R.-H. Park, "High-contrast color-stripe pattern for rapid structured-light range imaging," ECCV, 2004.

[10] L. Zhang, B. Curless and S. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," 3DPVT, 2002.

[11] 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," ICCV, 2009.

[12] P. Felzenszwalb and D. Huttenlocher, "Efficient belief propagation for early vision," IJCV, 2006.

[13] H. Aoki, R. Furukawa, M. Aoyama, S. Hiura, N. Asada, R. Sagawa, H. Kawasaki, S. Tanaka, S. Yoshida and Y. Sanomura, "Proposal on 3-D Endoscope by Using Grid-based Active Stereo," IEEE Engineering in Medicine and Biology Society (EMBC'13), 2013.