ONE-SHOT RANGE SCANNER USING COPLANARITY CONSTRAINTS

Ryo Furukawa,[†] *Huynh Quang Huy Viet,*[††] *Hiroshi Kawasaki,*[††] *Ryusuke Sagawa*[†††] *and Yasushi Yagi*[†††]

[†]Faculty of Information Sciences, Hiroshima City Univ., Japan
[††]Faculty of Engineering, Saitama Univ., Japan
[†††]Institute of Scientific and Industrial Research, Osaka Univ., Japan

ABSTRACT

Methods for scanning dynamic scenes are important in many applications and many systems using structured light have been proposed. Many of these systems use either multiple patterns projected rapidly or a single pattern. Although the former allows dense reconstruction with a sufficient number of patterns, it has difficulty in capturing objects in rapid motion. The latter technique uses only a single pattern and have no such difficulties, however, they often have stability problems and their result tend to have low resolution. In this paper, we develop a system to achieve dense and accurate 3D measurement from only a single image. The proposed system also has the advantage of being robust in terms of image processing.

Index Terms— machine vision, shape measurement,

1. INTRODUCTION

Currently active 3D scanners are widely used for actual 3D model acquisition process [1]. Especially, structured light based systems [4] have been intensively researched and commercialized, because systems of this type are relatively simple and realize high accuracy. To scan a 3D shape with motion, such as human body in motion or faces with dynamic expressions, 3D scanners using high-speed structured light systems have been studied in recent years [3, 6]. However, since these systems assume that there is little motion in a scene while a sufficient number of patterns for decoding are projected, it is often difficult to capture objects with rapid motion. In addition, difficulty in designing high-speed synchronization systems is also an issue.

On the other hand, 'one-shot' structured light systems which use only single images have been also studied. Widely used methods in this category are embedding positional information of the projectors' pixels into spatial patterns of the projected images [5]. Although the techniques can resolve the issues of rapid motions and synchronization, they typically use patterns of complex intensities or colors to encode positional information into local areas. Because of the complex patterns, they often require assumptions of smooth surface or reflectance, and the image processing tends to be difficult. If the assumptions do not hold, the decoding process of the patterns may be easily affected and leads to unstable reconstruction.

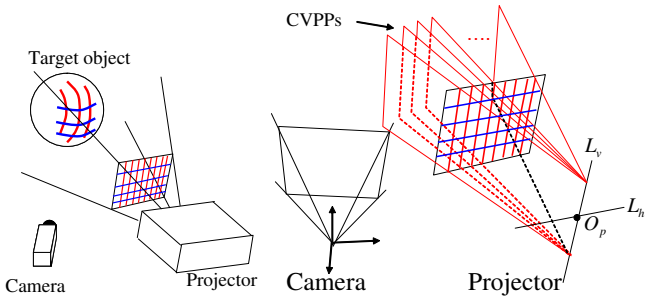

Fig. 1. Scanning system:(left) the system configuration, and (right) CVPPs.

This paper presents a single scanning technique resolving the aforementioned problems. The proposed technique uses a simple grid pattern formed by straight lines distinguishable only as vertical or horizontal lines. Since lines should be identified into only two types, two colors are enough so that image processing is simple and stable for surface texture/color. In addition, there is no need to encode particular information for the local grid pattern itself, so the pattern can be dense as long as it is extractable. Normally, a shape cannot be reconstructed from such a pattern. Thus, a new technique that reconstructs the grid pattern using coplanarity constraints [2] is presented.

2. SHAPE RECONSTRUCTION FROM GRID PATTERN

2.1. Outline

The 3D measurement system proposed consists of a camera and a projector as shown in Fig.1(left). Two types of straight line patterns, which are vertical and horizontal stripes, are projected from the projector and captured by the camera. The vertical and horizontal patterns are assumed to be distinguishable by color.

The straight pattern projected by the projector defines planes in 3D space. Planes defined by a vertical pattern and a horizontal pattern are respectively referred to as a vertical pattern plane (VPP) and a horizontal pattern plane (HPP).

The projector is assumed to have been calibrated. That is, all parameters for the VPPs and HPPs in 3D space are known. A VPP and a HPP with known parameters are referred to as

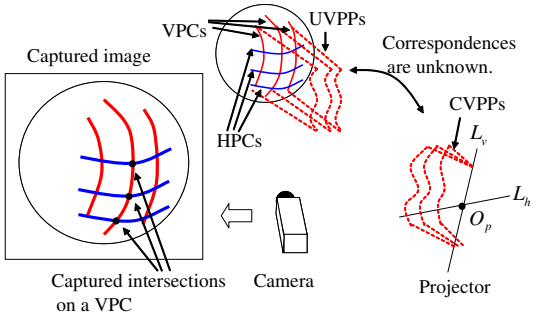

Fig. 2. CVPPs and UVPPs.

a calibrated VPP (CVPP) and a calibrated HPP (CHPP). All the CVPPs contain a single line, as in the figure 1(right). The same assumption holds for all CHPPs. These lines are denoted as L_v and L_h and the optical center of the projector O_p is the intersection of these lines. The point O_p and the direction vectors for L_v and L_h are given by calibration.

The vertical patterns projected onto the surface of the target scene are extracted from images captured by the camera. They are cross sections of the scene by VPPs. Since the correspondence from each of the VPPs generating detected vertical patterns to a particular CVPP is unknown, a VPP that generates a detected vertical pattern is referred to as an Unknown VPP (UVPP) as shown in Fig. 2. An Unknown HPP (UHPP) is similarly defined.

The goal of the problem is to determine correspondences between the UVPPs (UHPPs) and CVPPs (CHPPs) (otherwise described as identifying UVPPs and UHPPs). As a result, 3D positions of all the captured intersections become known.

Let the CVPPs obtained by calibration be represented as $V_1, V_2, \cdots, V_M$, CHPPs be represented as $H_1, H_2, \cdots, H_N$. Also, let the UVPPs and UHPPs obtained from the captured image be represented as $v_1, v_2, \cdots, v_m$ and $h_1, h_2, \cdots, h_n$, respectively. These symbols are used to represent correspondences between UVPPs and CVPPs. In this paper, the correspondence between the k-th UVPP v_k and i-th CVPP V_i is represented as $v_k \rightarrow V_i$, which means v_k is identified as V_i,

The proposed method derives linear equations based on conditions of coplanarity of UVPPs and UHPPs. The simultaneous equations have general solutions with one free parameter. The free parameter can be determined by assuming a correspondence between specific UVPP and CVPP. Then, the positions of all the UVPPs and UHPPs can be determined. Then, by comparing the positions with the CVPPs and CHPPs, each UVPP or UHPP is matched with the nearest CVPP or CHPP. Then, the differences between the UVPPs (or UHPP) and the corresponding CVPPs (or CHPP) are calculated. By regarding these differences as "errors", the correspondences with the minimum sum of squared errors are searched.

2.2. Solving coplanarity constraints

From the intersections of the grid pattern obtained from the captured image, linear equations can be derived. Suppose that the intersection between v_k and h_l is captured and its position on the image in the coordinates of the normalized camera is $\mathbf{u}_{k,l} = [s_{k,l}, t_{k,l}]^\top$. The planes v_k and h_l can be represented by

$$\mathbf{v}_k^\top \mathbf{x} = -1, \;\; \mathbf{h}_l^\top \mathbf{x} = -1 \tag{1}$$

respectively, where 3D vectors $\mathbf{v}_k$ and $\mathbf{h}_k$ are vectors of plane parameters and $\mathbf{x}$ is a point on each plane. Let the 3D position of the intersection $\mathbf{u}_{k,l}$ be $\mathbf{x}_{k,l}$, then $\mathbf{x}_{k,l}$ can be represented using the coordinates of the image as

$$\mathbf{x}_{k,l} = \gamma[\mathbf{u}_{k,l}^\top \;\; 1]^\top, \tag{2}$$

where γ is a scalar that represents the depth of the poiont. By substituting $\mathbf{x} = \mathbf{x}_{k,l}$ in equation (1), equations $\gamma \mathbf{v}_k^\top [\mathbf{u}_{k,l}^\top \;\; 1]^\top = -1$ and $\gamma \mathbf{h}_l^\top [\mathbf{u}_{k,l}^\top \;\; 1]^\top = -1$ are obtained. By dividing both equations by γ and subtracting them,

$$[\mathbf{u}_{k,l}^\top \;\; 1] \, (\mathbf{v}_k - \mathbf{h}_l) = 0 \tag{3}$$

is obtained. Let the simultaneous equations of (3) for all the intersections be described as $\mathbf{Aq} = \mathbf{0}$, where $\mathbf{q} = [\mathbf{v}_1^\top, \cdots, \mathbf{v}_m^\top, \mathbf{h}_1^\top, \cdots, \mathbf{h}_n^\top]^\top$.

In this paper, the plane that include both L_v and L_h is referred to as the projector focal plane (PFP). Let the direction vectors of the line L_v and L_h be represented as $\mathbf{l}_v$ and $\mathbf{l}_h$ respectively, and the plane parameters of the PFP be $\mathbf{p}$. Also, let the 3D coordinate vector of the optical center O_p of the projector be $\mathbf{o}_p$. UVPPs contain the line L_v, UHPPs contain the line L_h, and all the planes contain the point O_p. Thus,

$$\mathbf{l}_v^\top \mathbf{v}_k = 0, \;\; \mathbf{l}_h^\top \mathbf{h}_l = 0, \;\; \mathbf{o}_p^\top \mathbf{v}_k = -1, \;\; \mathbf{o}_p^\top \mathbf{h}_l = -1 \tag{4}$$

are obtained. Let the simultaneous equations (4) for $1 \leq k \leq m, \; 1 \leq l \leq n$ be described as $\mathbf{Bq} = \mathbf{b}$.

Since only the substitution $(\mathbf{v}_k - \mathbf{h}_l)$ appears in the equation (3), the solution of $\mathbf{Aq} = \mathbf{0}$ have freedoms of scaling and addition of a constant 3D vector. Using this fact, the general solution of $\mathbf{Aq} = \mathbf{0}$ can be written as

$$\mathbf{v}_k = s\mathbf{v}'_k + \mathbf{c}, \;\; \mathbf{h}_l = s\mathbf{h}'_l + \mathbf{c}, \tag{5}$$

where $\mathbf{q}'' = [\mathbf{v}_1'^\top, \cdots, \mathbf{v}_m'^\top, \mathbf{h}_1'^\top, \cdots, \mathbf{h}_n'^\top]^\top$ is a special solution of $(\mathbf{Aq}'' = \mathbf{0} \;\wedge\; \mathbf{Bq}'' = \mathbf{b})$, s is an arbitrary scalar, and $\mathbf{c}$ is an arbitrary 3D vector. From $\mathbf{Aq} = \mathbf{0} \;\wedge\; \mathbf{Aq}'' = \mathbf{0} \;\wedge\; \mathbf{Bq}'' = \mathbf{b}$, $\mathbf{l}_v^\top \mathbf{c}' = 0$, $\mathbf{l}_h^\top \mathbf{c}' = 0$, $\mathbf{o}_p^\top \mathbf{c}' = -1$, $\mathbf{o}_p^\top \mathbf{c}' = -1$, where $\mathbf{c}' \equiv \{1/(1-s)\}\mathbf{c}$ is obtained. These forms indicates that the plane $\mathbf{c}'$ coincides with PFP with parameter $\mathbf{p}$. Thus, the general solution of $(\mathbf{Aq} = \mathbf{0} \;\wedge\; \mathbf{Bq} = \mathbf{b})$ can be represented as

$$\mathbf{v}_k = s(\mathbf{v}'_k - \mathbf{p}) + \mathbf{p}, \;\; \mathbf{h}_l = s(\mathbf{h}'_l - \mathbf{p}) + \mathbf{p}, \tag{6}$$

where $\mathbf{x}' = [\mathbf{v}_1'^\top, \cdots, \mathbf{v}_m'^\top, \mathbf{h}_1'^\top, \cdots, \mathbf{h}_n'^\top]^\top$ is a special solution of $(\mathbf{Aq}' = \mathbf{0} \;\wedge\; \mathbf{Bq}' = \mathbf{b})$. $\mathbf{p}$ is known from calibrations and $\mathbf{v}'_k, \mathbf{h}'_l$ can be calculated from $\mathbf{A}, \mathbf{B}$ and $\mathbf{b}$.

2.3. Determining ambiguity

The scalar s of the form (6) can be calculated by assuming a specific correspondence between a UVPP and CVPP. By assuming the correspondence from the k'-th UVPP to the i'-th CVPP (i.e. $v_{k'} \rightarrow V_{i'}$)

$$\mathbf{V}_{i'} = \mathbf{v}_{k'} = s(\mathbf{v}'_{k'} - \mathbf{p}) + \mathbf{p} \tag{7}$$

holds, where $\mathbf{V}_{i'}$ is the parameter vector of the CVPP $V_{i'}$. From this form, s can be calculated by

$$s = ||\mathbf{V}_{i'} - \mathbf{p}||/||\mathbf{v}'_{k'} - \mathbf{p}||, \tag{8}$$

then, all UVPPs and UHPPs are determined using the s.

Let this s of the form (8) be denoted as $s(k', i')$. Then, $\mathbf{v}_k$ and $\mathbf{h}_l$ given the correspondence $v_{k'} \rightarrow V_{i'}$, which we refer to as $\mathbf{v}_k(k', i')$ and $\mathbf{h}_l(k', i')$, respectively, can be calculated by

$$\begin{aligned} \mathbf{v}_k(k', i') &= s(k', i')(\mathbf{v}_k - \mathbf{p}) + \mathbf{p}, \\ \mathbf{h}_l(k', i') &= s(k', i')(\mathbf{h}_l - \mathbf{p}) + \mathbf{p}. \end{aligned} \tag{9}$$

The next step is comparing the calculated UVPPs (or UHPPs) with the CVPPs (or CHPPs). For each UVPP, the difference between the UVPP and the nearest CVPP is calculated as an error. Then, we can calculate the error function of the assumption $v_{k'} \rightarrow V_{i'}$ as the sum of squared errors. By searching the minimum of the error function, we can find the optimum correspondence and the ambiguity can be solved.

The comparisons are executed between UVPPs $\mathbf{v}_k(k', i')$,$(k = 1, \cdots, m)$ and CVPPs $\mathbf{V}_i$,$(i = 1, \cdots, M)$, and also between UHPPs $\mathbf{h}_l(k', i')$,$(l = 1, \cdots, n)$ and CHPPs $\mathbf{H}_j$,$(j = 1, \cdots, N)$. In this paper, comparison is done based on the squared angles between the planes. More specifically, the error function is defined as

$$\begin{aligned} E_{k'}(i') \equiv &\sum_{k=1}^{m} \min_{i=1,\ldots,M} \{D(\mathbf{v}_k(k', i'), \mathbf{V}_i)\}^2 \\ &+ \sum_{l=1}^{n} \min_{j=1,\ldots,N} \{D(\mathbf{h}_l(k', i'), \mathbf{H}_j)\}^2, \end{aligned} \tag{10}$$

where D means the angle between two planes which can be defined as

$$D(\mathbf{v}_k, \mathbf{V}_i) \equiv \arccos((\mathbf{v}_k \cdot \mathbf{V}_i)/(||\mathbf{v}_k||\ ||\mathbf{V}_i||)). \tag{11}$$

Then,

$$i'_{min} \equiv \arg\min_{i'} E_{k'}(i') \tag{12}$$

is searched, and the set of planes $\mathbf{v}_k(k', i'_{min}),\ (k = 1, 2, ..., m)$ and $\mathbf{h}_l(k', i'_{min}),\ (l = 1, 2, ..., n)$ is the solution.

2.4. Projection and detection of a grid pattern

The proposed method can be regarded as a "matching" between UVPPs (UHPPs) and CVPPs (CHPPs). For stable matching, adding irregularities to the arrangement of the CVPPs (CHPPs) are desirable. In this paper, combined patterns of dense vertical lines with uniform intervals and horizontal lines with random intervals as shown in Fig.3(left) are used. Here, vertical and horizontal patterns were respectively colored red and blue so that they were distinguishable. This method only requires simple grid patterns. Because of the simplicity of the projected pattern, this method is less affected by noise, surface texture and color, or object shape.

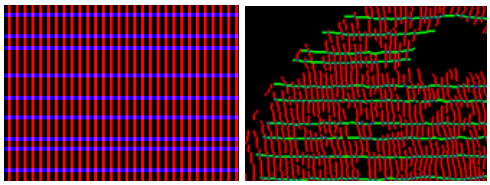

Fig. 3. A projected pattern (left) and detected lines(right). In the right figure, red curves are horizontal patterns, green curves are vertical, and intersection points are blue dots.

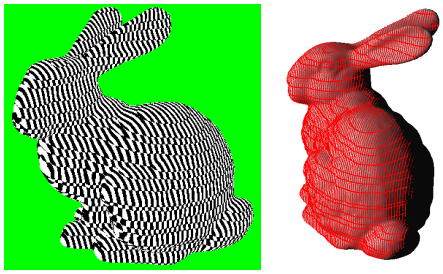

Fig. 4. Result of simulation data:(left) the synthesized image, and (right) the reconstructed points with red color and the shaded ground truth surface.

The detection of the curves of the grid patterns is processed as follows. First, the captured image is scanned horizontally and vertically, detecting peaks of vertical and horizontal patterns, respectively. Then, these peaks are connected using a simple labeling technique. Fig. 3(right) shows example of detected lines and intersection points. Note that we can observe many lines disconnected because of noise, texture color and shape of the object, however, since the lines are connected as network, we can still reconstruct a shape.

3. EXPERIMENTS

Several experiments were conducted to demonstrate the effectiveness of the proposed method.

The first experiment was conducted to confirm the validity of the proposed method using simulated data. The grid patterns of vertical lines with uniform intervals and horizontal lines with random intervals are provided. The synthesized camera images of the grid patterns are shown in Fig. 4(left). The shape obtained from the data with the ground truth is shown in Fig. 4(right). The result shows that the correct shapes was obtained by our method.

An actual 3D scanning system was built. Processing is performed by a PC with a Pentium Xeon 2.5Ghz CPU. Patterns were projected by a projector with a resolution of 1024×768 and scenes were captured by a CCD camera

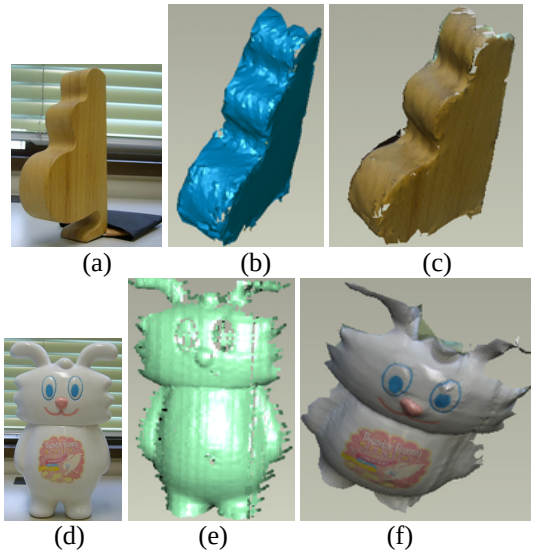

Fig. 5. Reconstruction results of static objects: (a)(d) the target objects, (b)(e) reconstructed shape and (c)(f) textured shape after hole-filling.

Fig. 6. Reconstruction and evaluation results: (left) the captured frame, and (right) the reconstructed model displayed with the ground truth data (red points: reconstructed model and shaded model: ground truth).

(720×480 pixels). Figure 5 shows the captured scenes and results of reconstruction. In the experiment, a ceramic bottle, a paper mask and a ceramic jug with an intricate shape were captured. Unlike previous methods, the proposed technique achieves reconstruction even if jump edges exist and lines are frequently segmented. The processing time for reconstructing the input data was about 1.6 sec. The most of the time was spent for calculation of SVD and the minimum value search for eliminating 1DOF took less than 0.001s.

Next, a scene of a box (size: 0.4 m × 0.3 m × 0.3 m) and a cylinder (height: 0.2 m, diameter: 0.2 m) was measured to evaluate the accuracy. The scene was also measured by a commercial active scanner as the ground truth. Figures 6(a) shows the captured scene and (b) shows both the reconstruction and the ground truth. Although there were small differences between the reconstruction and the ground truth, the shape was correctly restored. The RMS error of the reconstruction from the ground truth was $0.52mm$.

Finally, dynamic scenes were reconstructed with a human face as the target object. The target scene was captured to obtain a series of images while the face was being moved and facial expressions changed freely. Figure 7 shows an example of captured scene and the result. The result indicates that the proposed method successfully restored the human face.

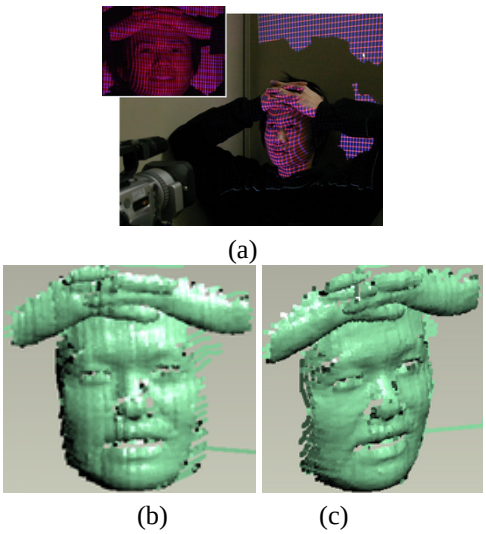

Fig. 7. Reconstruction of facial expressions: (a) a capturing scene, (b),(c) the reconstructed shape.

4. CONCLUSION

This paper proposes a technique to densely measure shapes of dynamic scenes using a single projection of structured light. Since the proposed technique does not involve encoding positional information into multiple pixels or color spaces, as often used in conventional one-shot 3D measurement methods; but into a grid pattern instead, the technique is less affected by an object's texture or shape, providing robust shape reconstruction. Tests were conducted to verify the method.

Acknowledgement

This work was supported in part by SCOPE(Ministry of Internal Affairs and Communications, Japan).

5. 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] R. Furukawa and H. Kawasaki. Self-calibration of multiple laser planes for 3D scene reconstruction. In *3DPVT*, pages 200–207, 2006.

[3] O. Hall-Holt and S. Rusinkiewicz. Stripe boundary codes for real-time structured-light range scanning of moving objects. In *ICCV*, volume 2, pages 359–366, 2001.

[4] S. Inokuchi, K. Sato, and F. Matsuda. Range imaging system for 3-D object recognition. In *ICPR*, pages 806–808, 1984.

[5] C. Je, S. W. Lee, and R.-H. Park. High-contrast color-stripe pattern for rapid structured-light range imaging. In *ECCV*, volume 1, pages 95–107, 2004.

[6] S. Zhang and P. S. Huang. High-resolution, real-time three-dimensional shape measurement. *Optical Engineering*, 45(12):123601, 2006.