

Dense 3D reconstruction method using a single pattern for fast moving object

Ryusuke Sagawa, Yuichi Ota, Yasushi Yagi
Institute of Scientific and Industrial Research,
Osaka University, Japan
{sagawa,yagi}@am.sanken.osaka-u.ac.jp

Ryo Furukawa, Naoki Asada
Faculty of Information Sciences,
Hiroshima City University, Japan
{ryo-f, asada}@hiroshima-cu.ac.jp

Hiroshi Kawasaki
Faculty of Engineering, Saitama University, Japan
kawasaki@cgv.ics.saitama-u.ac.jp

Abstract

Dense 3D reconstruction of extremely fast moving objects could contribute to various applications such as body structure analysis and accident avoidance and so on. The actual cases for scanning we assume are, for example, acquiring sequential shape at the moment when an object explodes, or observing fast rotating turbine's blades. In this paper, we propose such a technique based on a one-shot scanning method that reconstructs 3D shape from a single image where dense and simple pattern are projected onto an object. To realize dense 3D reconstruction from a single image, there are several issues to be solved; e.g. instability derived from using multiple colors, and difficulty on detecting dense pattern because of influence of object color and texture compression. This paper describes the solutions of the issues by combining two methods, that is (1) an efficient line detection technique based on de Bruijn sequence and belief propagation, and (2) an extension of shape from intersections of lines method. As a result, a scanning system that can capture an object in fast motion has been actually developed by using a high-speed camera. In the experiments, the proposed method successfully captured the sequence of dense shapes of an exploding balloon, and a breaking ceramic dish at 300–1000 fps.

1. Introduction

Dense and precise shape acquisition of fast moving object with high frame-rate is strongly required in wide fields. For example, if changing shapes of an exploding balloon or dense shape of fast rotating turbine's blades are acquired with high frame-rate, it is expected to make a great contribution to body structure analysis and accident avoidance.

For acquiring dense and accurate 3D shape, many laser-based 3D scanning systems have been developed. However, these are not suitable for fast-moving objects, because the laser lights should be physically moved in one or two dimensions for scanning and essentially need time.

Currently, 3D scanners that use area light sources such as

video projectors to reduce the scanning time are actively researched for scanning moving objects [20, 7]. For those systems, acquisition of correspondences between a camera and a projector is critical. Usually, such information is encoded into the projected pattern and decoded by image analysis; the encoding methods can be categorized into two types, such as spatial and temporal coding. Among them, spatial encoding methods that use only a single input image are more suitable for scanning extremely fast moving object.

A typical example of spatial encoding methods is a color coding method with which reconstruction is achieved by stereo with a window matching technique [3, 20]. However, many color coding methods have a critical problem for density and precision on shape reconstruction. In these methods, to retrieve correspondences between a camera and a projector, positional information to identify each pixel are encoded in each window. Since pattern is usually distorted and compressed depending on the object's shape, it is difficult to use a dense pattern for actual scan, resulting in a sparse reconstruction. In addition, to encode positional information efficiently in a small window, a number of colors increases. This makes the technique to be influenced easily by textures of the surfaces, and thus, shape reconstruction tends to be unstable.

If a shape reconstruction using a small number of colors with dense pattern is realized, aforementioned problems are eliminated and dense and high precision reconstruction for fast moving objects can be achieved. In this paper, we propose a method to realize this by combining two approaches; the first approach is an efficient pattern encoding technique to allow a periodical identification of line patterns using just two colors of de Bruijn sequence and the second is a shape reconstruction technique that does not require explicit correspondences between a camera and a projector.

About the first approach, a dense grid pattern using de Bruijn sequence is proposed. Normally, the number of required colors for de Bruijn sequence increases if unique identification for all pixels is necessary. In the proposed method, however, smaller number of colors is sufficient, be-

cause non-unique identification with the same IDs appearing periodically is allowed for restoring the scene with our reconstruction method (the second approach). Moreover, the grid pattern can be dense because errors of the identification are allowed up to some extent with our reconstruction method. To achieve stable extraction of the grid pattern, vertical and horizontal relationships of adjacent pixels are considered using a belief propagation (BP) technique.

The second approach relates to shape reconstruction techniques that use intersections of line patterns [2, 11, 10]. Although these techniques have great advantages that they require no identification of line patterns, they require large computational costs. In addition, these techniques can reconstruct scenes only with some indeterminacies [12]. Thus, Euclidean upgrade is required for final result and it is an unstable process. Because of such drawbacks, reconstructed shapes are sparse and computational costs are high. In this paper, we solve the problem by a new and efficient calculation technique that can drastically reduce a number of equations for 3D reconstruction.

Note that a complementary combination of the two approaches has a potential for dense and stable one-shot scanning for the first time; the technique can be applied for scanning an extremely fast moving object with high density and accuracy. The main contributions of the paper are as follows: (1) Dense shape reconstruction from a single image using a dense and simple pattern is realized, (2) robust edge detection algorithm based on a belief propagation is presented, (3) an efficient and robust shape reconstruction method that uses both geometric constraints of line patterns with structures of pencils of planes and de Bruijn sequence is proposed, and (4) an actual system which can capture an extremely fast moving object is constructed to show the strength of the method.

2. Related work

Shape reconstruction techniques with a structured light with temporal and spacial coding are summarized in Ref. [1]. Systems using only temporal coding is easy to implement, accurate, dense and robust, so it has commonly been used for real applications [8]. Since the technique needs to capture multiple images, it is not suitable for high-speed capturing. Recently, several methods for high-speed capturing were proposed by using a DLP projector and a high-speed camera [17, 14]. Also, several research reducing the required number of patterns using both temporal and spacial changes were presented [7, 18]. Since these approaches work for objects that have limited speeds, they are not suited for capturing objects with extremely high speed.

Some methods use projectors only to provide textures that change over time and 3D information is restored using a passive stereo technique [3, 20], although they are not strictly structured light systems. Since they still require

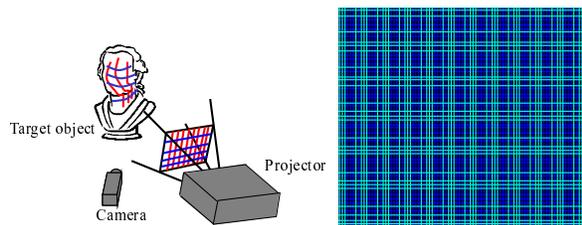


Figure 1. (Left) Scanning system: multiple lines are projected and their intersections are detected and used for reconstruction. (Right) projected pattern.

several patterns for identification, they are not suited for extremely fast objects.

Techniques using only spatial encoding of a pattern are suitable for fast-moving objects, since they use only a single-frame image [9, 16]. On the other hand, the problems are that they typically need complex patterns or colors to encode positional information. To determine the spatial codes uniquely, the size of a code becomes large. Such patterns are easily affected by textures, shape discontinuities, image compression caused by tilted surfaces. Therefore, density of patterns should be inevitable low, and thus, 3D reconstruction tends to be sparse and unstable in reality.

There are several researches that achieve shape reconstruction from single images composed of multiple lines or stripes. Koninckx *et al.* proposed a technique for dense shape reconstruction using a simple pattern, *i.e.*, a set of stripes [13]. Their method depends on relative numbering of the dense patterns, which assumes local smoothness of the surface and may be disturbed by shape discontinuities and line detection failures. Similarly, Frueh and Zakhor [6] used vertical stripes and a horizontal line. The vertical stripes are used for shape reconstruction and the horizontal lines for identifying the vertical stripes. Some techniques reconstruct 3D scenes, by representing projected line patterns as an intersection between 3D planes and the scenes [2, 11, 10], however, their technique can only applied for sparse pattern; the reason will be discussed in Sec. 5.

3. System configuration

The proposed 3D measurement system consists of a camera and a projector as shown in Fig. 1 (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 projector 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.

In this study, the method of reconstructing 3D information efficiently is proposed with this system configuration. First, the projected grid pattern is extracted from the captured image. To detect them stably, a method that utilizes belief propagation and de Bruijn sequence is proposed. In the method, pixels are labeled using BP, then, curves of

boundaries are extracted in sub-pixel accuracy at the boundaries of the labels (Sec. 4).

Each detected curve is illuminated by a grid line emitted from the projector. A emitted line from the projector goes through a plane in the 3D space. The intersection points of the grid pattern (grid points) can be extracted from the detected vertical and horizontal curves. From these grid points, simple constraints about the vertical and horizontal planes are acquired. By solving all the constraints for the entire grid pattern, solutions of the vertical and horizontal planes are obtained, and the scene can be reconstructed (Sec. 5).

4. Detection of dense grid pattern for one-shot active stereo

The detection of color line patterns proposed in this paper consists of two steps. The first step is detecting lines regardless of the color and the second step is decoding the color code based on the de Bruijn pattern shown in 1 (right) and 2 (left).

In the previous studies, the stripe patterns were used as shown in Fig. 2(middle). In that case, the cross talk of RGB planes becomes a problem. Fig. 2(right) compares the green values of the left and middle figures in Fig. 2. The ratio of changes of the solid lines, which represents the profile of an intensity of image for the proposed pattern, are larger than that of the dotted lines, which represents the profile of the stripe pattern used for the previous study. Thus, thin lines used on our method are preferable for detecting dense lines. To enable us to detect dense lines, we define the line to be detected as a peak of intensity in the perpendicular direction of the line. By using this definition, we can ideally detect dense lines that exist every other pixel. Moreover, we can double the density of lines by detecting the negative peak, namely local minimum of the intensity, as the projected lines. Therefore, we can detect lines on every pixel at maximum.

Advantages of the proposed method are as follows. (1) Projected grid lines can be detected while the vertical and horizontal lines are designated, even if they are drawn in the same color. (2) Since the label of each pixel is determined using information of the adjacent pixels, the method is stable against noise.

4.1. Detecting dense line patterns

To take the continuity of a line into consideration, we propose a method based on the belief propagation, which is an energy minimization problem on a graph defined by [5]

$$E(f) = \sum_{p \in V} D_p(f_p) + \sum_{(p,q) \in U} W_{pq}(f_p, f_q), \quad (1)$$

where f is the set of labels to be determined, V is the set of nodes, U is the set of edges, and p and q are the nodes of the graph. $D_p(f_p)$ is the data cost of assigning label f_p to

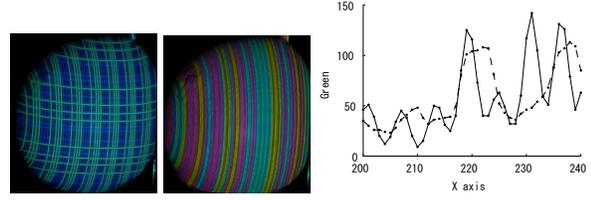


Figure 2. Line detection: (left) projected pattern of the proposed system, (middle) projected pattern of a previous study, and (right) the profiles for the images for the proposed pattern (solid lines) and the previous study (dotted lines).

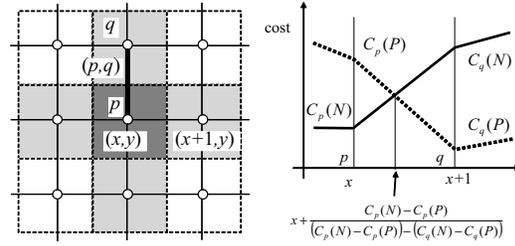


Figure 3. Line detection: (left) Graph for BP, and (right) subpixel calculation.

p . $W_{pq}(f_p, f_q)$ is the discontinuity cost of assigning labels f_p and f_q to neighboring nodes.

The proposed method separately detects vertical and horizontal lines projected by the sensor, and now we consider the detection of vertical lines. In detecting lines, the nodes correspond to pixels of the camera, and the edges are the connection to 4-neighboring pixels (the hatched pixels in Fig. 3(left)). The proposed method discriminates every pixel to three labels based on the derivative of the intensity along x-axis (horizontal) of the camera. The labels are positive (P), negative (N), and nearly zero (0). From the definition of the line defined above, the line is detected as the boundary of the labels P and N .

The data cost $D_p(f_p)$ is computed by the forward difference

$$D_p(f_p) = \begin{cases} I(x+1, y) - I(x, y) & \text{if } f_p = N \\ |I(x+1, y) - I(x, y)| & \text{if } f_p = 0 \\ -(I(x+1, y) - I(x, y)) & \text{if } f_p = P \end{cases}, \quad (2)$$

where $I(x, y)$ is the intensity at a pixel $p = (x, y)$. The discontinuity cost $W_{pq}(f_p, f_q)$ depends on the direction of the edge as follows:

$$W_{pq}(f_p, f_q) = \begin{cases} -\lambda(f_q - f_p)(I(x+1, y) - I(x, y)) & \text{if edge } (p, q) \text{ is along x-axis} \\ |f_q - f_p| & \text{if edge } (p, q) \text{ is along y-axis} \end{cases}, \quad (3)$$

where f_p and f_q are 0, 1, and 2 for the labels N , 0 , and P , respectively. λ is a user-defined parameter. Because of changing the discontinuity cost with respect to the direction of the edges, the proposed method can detect vertical lines while the horizontal lines are ignored even if they are the same color. The solution is given by iterative computation

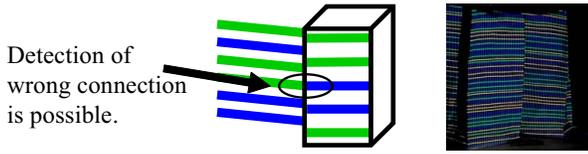


Figure 4. (Left) Using color stripes for detection of wrong connection and (right) actual example image.

of passing messages between the nodes based on the maximum product BP algorithm.

The label of each node is determined by choosing the label of the minimum cost. If the boundary of the labels from P to N is detected, the subpixel position of the line is computed by

$$x + \frac{C_p(N) - C_p(P)}{(C_p(N) - C_p(P)) - (C_q(N) - C_q(P))}, \quad (4)$$

where $C_p(f_p)$ is the final cost, which is computed as the sum of the data cost and the messages (Fig. 3(right)). The x-coordinate of q is $x + 1$ where that of p is x .

The horizontal lines are detected in the similar manner by swapping the directions. Detecting the negative peaks is also the same. Instead of finding the boundary from P to N , the negative peak is detected as the boundary from N to P .

4.2. Decoding color codes

Next, we decode the color code based on the de Bruijn sequence [9, 15, 19]. A q -ary de Bruijn sequence of order n is a sequence of length q^n consisting of an alphabet of size q in which every possible subsequence of length n is present exactly once. If a projected pattern is encoded by two or more symbols distinguished in a camera image, the correspondence between an element in the projected pattern and the observed pattern is uniquely determined by matching subsequences of length n in a de Bruijn pattern.

There are two advantages to use color patterns. First, wrong connection of lines can be reduced by color information. Even if different lines are connected as a continuous line, the discontinuity, which usually occurs at occluding boundary, can be detected by recognizing the color patterns as shown in Fig. 4.

The second advantage is that matching the sets of planes becomes easy by using color patterns; actual algorithm will be described in Sec. 5. Without color patterns, all combinations between projected and observed patterns must be compared to determine the ambiguity. If the de Bruijn pattern is used, it is sufficient that an observed line is compared to every q^n projected lines. This means the density of lines can be increased q^n times in principle. Moreover, since the previous methods that use color patterns need to recognize many lines uniquely, q and n must be large number. However, periodic patterns generated with small q and n are sufficient for the proposed method. In this paper, we used the number of colors $q = 2$ and the length of codes $n = 3$ as

shown in Fig. 1 (right). Namely, each cycle of the pattern consists of eight lines.

For robust decoding of the sequence, a DP matching is used in [19]. Since the proposed method uses vertical and horizontal lines, we can more robustly decode the sequence by utilizing two dimensional regularization based on the belief propagation. Once the vertical and horizontal lines are detected, the intersecting points are computed. Therefore, we use the intersection points as the node of a graph and the edges are determined by the detected lines. Since each cycle consists of eight lines, the number of labels is eight.

The data cost $D_p(f_p)$ for decoding the de Bruijn sequence in the horizontal lines are defined as follows:

$$D_p(f_p) = |H(p) - H(f_p)|, \quad (5)$$

where $H(p)$ is the hue at the point of between the intersection point p and neighbor intersection point, and $H(f_p)$ is the hue of the projected light for the line of label $f_p = 0, \dots, 7$.

The discontinuity cost $W_{pq}(f_p, f_q)$ is given by

$$\begin{aligned} W_{pq}(f_p, f_q) &= \min(|(f_p + d(p, q)) \bmod 8 - f_q|, \\ &|8 - (f_p + d(p, q)) \bmod 8 - f_q|)^2 \end{aligned} \quad (6)$$

where $d(p, q)$ is 1 if q is on the next horizontal line in the $+y$ direction, -1 if q is on the next horizontal line in the $-y$ direction, and 0 otherwise. Namely, $W_{pq}(f_p, f_q)$ depends on the direction of the edge again.

5. Efficient shape from line pattern based on pencils of planes

5.1. Problem definition

In this section, the approach for reconstruction from stripes without identification of each line is described. In the proposed method, a set of patterns of parallel lines is used as a structured light. Each of the line pattern is emitted from the projector and spreads through a plane in a 3D space. We use a *light plane* to represent such a plane.

As described in Sec. 4, lines projected onto the target scene are detected from the image. Each detected line corresponds to a light plane described above. If the 3D position of the light plane corresponding to a detected line were known, the 3D shape of the detected line could be calculated by triangulation; but in reality the position of the light plane is unknown. Thus, we need to estimate the light planes to reconstruct the scene.

Kawasaki and Furukawa examined a problem of estimating light planes generated by a freely moved line laser projector [12]. They pointed out that reconstruction results include at least 4-DOF indeterminacies. Several researches are conducted to achieve Euclidean solution by using additional geometric constraints for similar problems [2, 11, 4, 10].

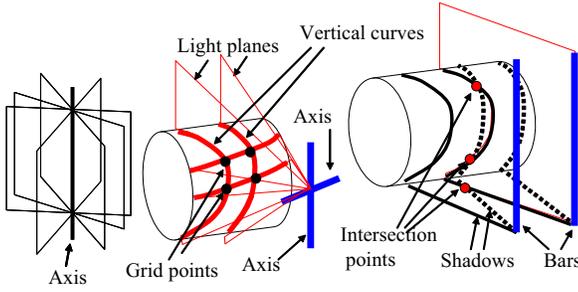


Figure 5. Left:Pencil of planes. Middle:Example of shape from cast-shadows. Right:Example of one-shot active stereo.

In the above researches, reconstruction is based on representing each light plane with 3 parameters since no constraints are assumed for these planes. If a dense grid pattern is used, the 3-parameter representation causes a rapid increase of the variables and equations used for reconstruction; thus, it results in increase of computational costs and instability. A straightforward approach to this problem is using 2-DOF or less parametrization of planes by imposing certain geometrical constraints to the light planes. Bouguet et al. used this approach; they used 2-parameter representation of planes by assuming that all the light planes were generated by a fixed point light source (the assumption means that the planes share a single point) [2].

A set of planes that share a single line is called a pencil of planes(Fig. 5(left)). A plane in a pencil of planes can be represented by 1 parameter¹. Since the light planes associated with the projected grid pattern used in the proposed method have structures of two pencils of planes (one is for the set of all the vertical lines and the other is for the horizontal lines), each of the planes can be parametrized by 1 parameter. By using the 1-parameter representation, there exists a simple proportional relationship between a parameter of *vertical* and *horizontal planes*, which we use to represent light planes generated by vertical and horizontal grid lines, respectively. For each grid point detected from the image, a single proportional relationship between the light planes is obtained. Using this representation and other mathematical techniques, we can decrease the number of variables.

The proportional relationships obtained from the grid points form a system of simultaneous equations. Since it can be shown to have a solution having 1-DOF ambiguity, the remaining ambiguity should be solved by using other constraints. In this paper, the ambiguity can be resolved using the condition that the arrangement of the vertical and horizontal planes and ID of de Bruijn sequence obtained from the grid points must coincide with the given grid pat-

¹A plane in the 3D space corresponds to a point in the dual space. A pencil of planes in the 3D space corresponds to a line in the dual space. 1 parameter representation of the planes corresponds to representing points on a line using 1 parameter in the dual space.

tern.

Note that our method is not specialized only for active stereo using video projectors, but it can be used for more general purposes, such as shape from cast-shadow generated by parallel edges and an uncalibrated moving light source (Fig. 5(right)).

5.2. Representation of planes

In this section, a 1-parameter representation of a light plane is described.

We assume both vertical and horizontal planes do not include the optical center of the camera. Thus, a vertical plane can be represented by

$$v_1x_1 + v_2x_2 + v_3x_3 + 1 = \mathbf{v} \cdot \mathbf{x} + 1 = 0, \quad (7)$$

where $\mathbf{v} \equiv (v_1, v_2, v_3)$ is a parameter vector of the plane and $\mathbf{x} \equiv (x_1, x_2, x_3)$ is a 3D point represented by the camera coordinates. In this paper, we use \mathbf{v} to represent a plane itself, as well as the parameter vector of the plane. All the vertical planes include the axis of the pencil of planes. Let a point on the axis be $\mathbf{p}_v \equiv (p_{v1}, p_{v2}, p_{v3})$. Let the direction vector of the line (the vertical direction of the projector) be $\mathbf{q} \equiv (q_1, q_2, q_3)$. Then, \mathbf{v} should fulfill $\mathbf{v} \cdot \mathbf{p} + 1 = 0$ and $\mathbf{v} \cdot \mathbf{q} = 0$. By solving these equations, we obtain

$$\mathbf{v} = \mathbf{v}_0 + \eta(\mathbf{p}_v \times \mathbf{q}), \quad (8)$$

where \mathbf{v}_0 is an arbitrary vertical plane. Similarly, a horizontal plane $\mathbf{h} \equiv (h_1, h_2, h_3)$ can be represented by

$$\mathbf{h} = \mathbf{h}_0 + \rho(\mathbf{p}_h \times \mathbf{r}) \quad (9)$$

where \mathbf{r} is the horizontal direction of the projector, \mathbf{p}_h is one of the points on the axis of the pencil of planes that include the horizontal planes, and \mathbf{h}_0 is an arbitrary horizontal plane. The plane that goes through the optical center of the projector and is parallel with the projector's image plane is called the focal plane of the projector, and we use \mathbf{s} to represent this plane. Plane \mathbf{s} is shared by the two pencils of planes. Thus, by defining $\mathbf{v}' \equiv \mathbf{p}_v \times \mathbf{q}$, and $\mathbf{h}' \equiv \mathbf{p}_h \times \mathbf{r}$, vertical planes and horizontal planes can be represented by the following equations (see Fig. 6(left)):

$$\mathbf{v} = \mathbf{s} + \eta\mathbf{v}', \quad \mathbf{h} = \mathbf{s} + \rho\mathbf{h}'. \quad (10)$$

We also use the 1D parameter to represent plane itself (i.e., ρ is used to represent plane \mathbf{v}).

5.3. Solution

If a grid point (i.e., an intersection point between vertical and horizontal curves, some of which are shown in Fig.5(middle)) is detected, and the vertical and horizontal curves are respectively included by a vertical plane \mathbf{v} and horizontal plane \mathbf{h} , then the following equation is fulfilled:

$$\mathbf{u} \cdot (\mathbf{v} - \mathbf{h}) = 0 \quad (11)$$

where $\mathbf{u} \equiv (u, v, 1)$ represents the 2D position of the detected grid point (u, v) in the normalized camera coordinates (see Fig.6(left)). About intuitive meanings of this

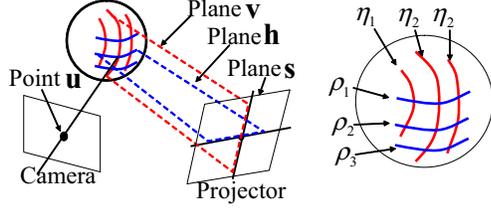


Figure 6. Left: Constraints of planes. Right: Dependencies between planes. Each of the parameters of the red planes (i.e., η_1, η_2 , and η_3) linearly depends on the parameters of the blue planes (ρ_1, ρ_2 , and ρ_3), if the blue planes are fixed and the red planes are estimated.

equation, refer to the work of Bouguet[2]. From Eqs. (10) and (11), we obtain

$$\eta(\mathbf{u} \cdot \mathbf{v}') = \rho(\mathbf{u} \cdot \mathbf{h}'). \quad (12)$$

Each vertical line detected from the image is associated with a vertical plane. Let the index of a detected vertical line be i the corresponding vertical plane be η_i . Similarly, let the index of a detected horizontal line be j , and the corresponding plane be ρ_j . If those curves have a grid point $\mathbf{u}_{i,j}$, then $\eta_i(\mathbf{u}_{i,j} \cdot \mathbf{v}') = \rho_j(\mathbf{u}_{i,j} \cdot \mathbf{h}')$. By defining constants $F_{i,j} \equiv \mathbf{u}_{i,j} \cdot \mathbf{v}'$ and $G_{i,j} \equiv \mathbf{u}_{i,j} \cdot \mathbf{h}'$, we obtain

$$F_{i,j}\eta_i = G_{i,j}\rho_j, \quad (13)$$

which means that the 1D parameters of the vertical and horizontal planes have a simple proportional relationship.

By accumulating Eq. (13) for all the grid points, L simultaneous equations with $(M + N)$ variables ($\eta_i, \rho_j, 1 \leq i \leq M, 1 \leq j \leq N$) are obtained, where M and N are the numbers of detected vertical and horizontal curves, respectively.

Let k be an index of a grid point, and let $i(k)$ and $j(k)$ be the indices of the vertical and horizontal planes that go through grid point k . Let \mathbf{T} be a $L \times M$ matrix whose (p, q) element is $F_{i(p),j(p)}$ if $q = i(p)$, and otherwise 0. Let \mathbf{R} be a $L \times N$ matrix whose (p, q) element is $G_{i(p),j(p)}$ if $q = j(p)$, and otherwise 0. Then, by defining $\vec{\eta} \equiv (\eta_1, \dots, \eta_M)^\top$ and $\vec{\rho} \equiv (\rho_1, \dots, \rho_N)^\top$, simultaneous Eq. (13) are represented by

$$\mathbf{T}\vec{\eta} = \mathbf{R}\vec{\rho}. \quad (14)$$

To solve Eq. (14) by a least squares method, $\|\mathbf{T}\vec{\eta} - \mathbf{R}\vec{\rho}\|^2$ should be minimized. This can be achieved by calculating the eigenvector associated to the minimum eigenvalue of symmetric $(M + N) \times (M + N)$ matrix $[\mathbf{T} | -\mathbf{R}]^\top [\mathbf{T} | -\mathbf{R}]$. There are efficient numerical algorithms for this problem.

It is possible to reduce variables from Eq. (14). Solution of least squares method is obtained by

$$\min_{\vec{\eta}, \vec{\rho}} \|\mathbf{T}\vec{\eta} - \mathbf{R}\vec{\rho}\|^2 = \min_{\vec{\rho}} (\min_{\vec{\eta}} \|\mathbf{T}\vec{\eta} - \mathbf{R}\vec{\rho}\|^2). \quad (15)$$

The minimum value $\min_{\vec{\eta}} \|\mathbf{T}\vec{\eta} - \mathbf{R}\vec{\rho}\|^2$ with respect to $\vec{\eta}$

is achieved when $\vec{\eta} = \mathbf{T}^\dagger \mathbf{R}\vec{\rho}$, where $\mathbf{T}^\dagger \equiv (\mathbf{T}^\top \mathbf{T})^{-1} \mathbf{T}^\top$ is a pseudo inverse of \mathbf{T} . This equation means that each vertical plane η_i linearly depends on horizontal planes $\vec{\rho}$ if $\vec{\rho}$ is fixed. (See Fig. 6(right)).

Since (r, c) element of $\mathbf{T}^\top \mathbf{T}$ is

$$(\mathbf{T}^\top \mathbf{T})_{r,c} = \begin{cases} \sum_s T_{s,r}^2 & \text{if } r = c \\ 0 & \text{otherwise} \end{cases}, \quad (16)$$

$\mathbf{T}^\top \mathbf{T}$ is a diagonal matrix (equal of Eq. (16) comes from the fact that each row of $\mathbf{T}_{s,c}$ has only one non-zero element). Thus, $(\mathbf{T}^\top \mathbf{T})^{-1}$ can be directly calculated and \mathbf{T}^\dagger is obtained by simple multiplication. By substituting $\vec{\eta} = \mathbf{T}^\dagger \mathbf{R}\vec{\rho}$ into Eq. (15), we obtain

$$\min_{\vec{\rho}} \|\mathbf{T}\vec{\eta} - \mathbf{R}\vec{\rho}\|^2 = \min_{\vec{\rho}} (\|(\mathbf{T}\mathbf{T}^\dagger \mathbf{R} - \mathbf{R})\vec{\rho}\|^2). \quad (17)$$

This means that the optimal solution of horizontal planes $\vec{\rho}$, which we term by $\hat{\rho}$, is calculated as the eigenvector associated to the minimum eigenvalue of matrix $(\mathbf{T}\mathbf{T}^\dagger \mathbf{R} - \mathbf{R})^\top (\mathbf{T}\mathbf{T}^\dagger \mathbf{R} - \mathbf{R})$. Then the optimal solution $\hat{\eta}$ is obtained by $\hat{\eta} = \mathbf{T}^\dagger \mathbf{R}\hat{\rho}$. Since $(\mathbf{T}\mathbf{T}^\dagger \mathbf{R} - \mathbf{R})^\top (\mathbf{T}\mathbf{T}^\dagger \mathbf{R} - \mathbf{R})$ is a $N \times N$ symmetric matrix, the computational complexity can be further reduced.

5.4. Solving ambiguity

Eigenvectors always has an ambiguity of scaling. Actually, if Eq. (14) is fulfilled, $\vec{\eta}$ and $\vec{\rho}$ can be replaced by $c\vec{\eta}$ and $c\vec{\rho}$, respectively. Inversely, if all the vertical and horizontal lines extracted from the image are connected by intersections, then, this equation does not have indeterminacy except for scaling. So, 3D reconstruction is performed for each of the connected groups of the curves. The value of c cannot be determined from Eq. (14), however, if the true value of c is \bar{c} , then the vertical and horizontal planes of solutions $\bar{c}\vec{\eta}$ and $\bar{c}\vec{\rho}$ should be coincide with the given grid pattern which is actually projected. In this study, 1-D search [10] with respect to c to obtain \bar{c} is used to solve the indeterminacy of scaling.

For simplicity, we calculate the similarity between a solution of horizontal plane calculated from grid points and the horizontal lines of the given grid pattern that is actually emitted. Let P be the number of the horizontal lines of the given grid and let $\mathbf{g}_l, 1 \leq l \leq P$ be the horizontal planes that are associated to the lines. These vectors can be calculated in advance. Then, the matching error function to compare $c\rho_j, 1 \leq j \leq N$ and $\mathbf{g}_l, 1 \leq l \leq P$ can be defined as

$$E(c) = \sum_{1 \leq i \leq N} \min_{1 \leq l \leq P} e(\mathbf{g}_l, \mathbf{s} + c\rho_j \mathbf{h}'), \quad (18)$$

where $e(\mathbf{a}, \mathbf{b})$ is a function that outputs the square of the angle between \mathbf{a} and \mathbf{b} . By minimizing $E(c)$ with respect to c , \bar{c} is obtained. Since $E(c)$ has numerous local minima, it is difficult to apply numerical optimization methods. In actual cases, the range of c is limited by the given horizontal planes, thus, full 1D search of the range by intervals of required precision is sufficient.

To increase robustness of the search of solution, we use ID information of de Bruijn sequence for matching. The method is simple. For the calculation of similarity function (18), comparison is only done if the IDs of the given grid line g_l and the detected line $s + c\rho_j\mathbf{h}'$ are the same. Thus, $e(\mathbf{a}, \mathbf{b})$ can be redefined to be a function that outputs the square of the angle between \mathbf{a} and \mathbf{b} if $B(\mathbf{a}) = B(\mathbf{b})$, and outputs 0 otherwise, where $B(\mathbf{a})$ means the ID of de Bruijn sequence associated to the plane represented by \mathbf{a} . This reduces the matching candidates (by about 1/8 in the case of the implementation of this paper) and the stability of search of the solution is much stabilized. Note that, since the IDs are just used for additional information for 1D optimization, error of identification is not serious compared to previous color coding methods [9, 15, 19].

6. Experiments

6.1. Comparison and evaluation

To confirm the advantages of our method, we reconstructed an object with texture and an object with occlusion. In the proposed method, de Bruijn sequence of 2 colors (blue and green) with the code length of 3 was used in vertical and horizontal lines. For comparison, color coding method using only vertical color stripes [19] was also applied to the objects, in which the color pattern was constructed based on de Bruijn sequence.

The result is shown at Fig. 7. In the color coding method, 3D reconstruction was successfully achieved by using both the geometrical information obtained from the grid points and the coded IDs, even if ambiguity remains in the IDs for deciding unique correspondences. Moreover, in the proposed method, extraction of edges were less affected by the textures because of using only two colors. In addition, the numbers of reconstructed points in the proposed method, and the color coding method was 43852(book)/24796(cup) and 12321/8319, respectively. Thus, the proposed method could reconstruct more points than the other by the factor of 3.3 due to using dense lines. Standard deviation of the proposed method and the color coding method was 2.09mm and 4.15mm, respectively. Note that for calculating the standard deviation of the color coding method, it required manual process to remove many outliers, whereas nothing was required for our method. From the results, we can confirm that the proposed method provides more dense and stable reconstruction than the previous color coding method.

6.2. Exploding balloon and breaking ceramic dish

Next, we conducted a dense 3D shape reconstruction using an object which changes its shape at the moment to prove the ability of high frame-rate scanning. For the purpose, we captured two objects, such as exploding balloon with 1000fps and shutter speed 1/20000 and breaking dish with 300fps and shutter speed 1/3200. In terms of balloon, it explodes so quick and only a several frames can be cap-

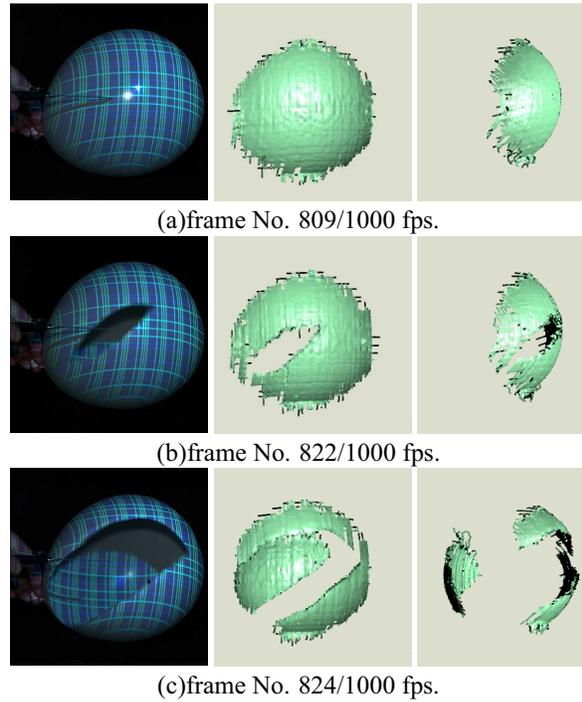


Figure 8. Results of a exploding balloon.

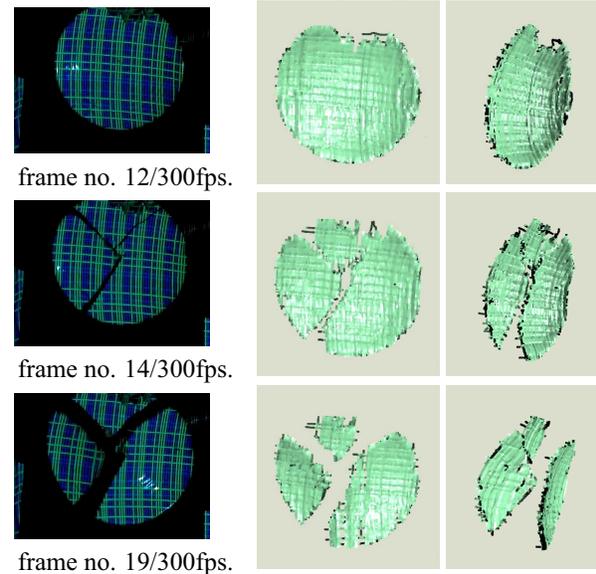


Figure 9. Results of a breaking dish.

tured at the moment of exploding as shown in Fig. 8. In terms of dish, we can learn that how cracks are propagated on the dish with 3D information, thanks to our dense reconstruction method.

7. Conclusion

In this paper, a one-shot active stereo system that can reconstruct dense 3D shapes was proposed. In the proposed method, a grid pattern is emitted from a video projector and the pattern projected onto the scene is captured by a cam-

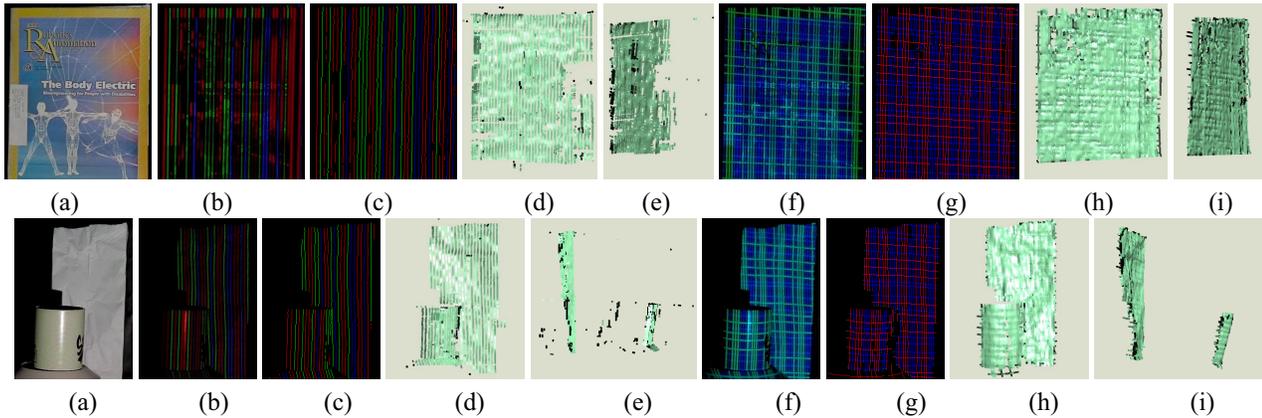


Figure 7. Comparison between color coding method [19] and our method: (a) target object, (b) captured scene with color coding method, (c) detect curves (d)(e) reconstructed shape, (f) captured scene with our method, (g) detected curves, and (h)(i) reconstructed shape.

era. To detect dense lines of the grid pattern stably, pixels are classified based on belief propagation, and the grid lines are extracted using boundaries of the classes in sub-pixel accuracy. Then, the extracted lines are labeled by IDs using de Bruijn sequence. In the paper, geometrical constraints caused by the projected grid pattern were formulated using pencils of planes. By using the structure, simple proportional relationships between parameters of planes were shown, and an efficient method to reconstruct 3D information up to 1-DOF indeterminacy using these constraints was presented. The remaining indeterminacy can be stably solved by using the IDs of de Bruijn sequence. Note that the IDs may be non-unique, because they are used with information of geometrical constraints of grids. Thus, only two colors are sufficient for the grid pattern, which helps to achieve stable extraction of the pattern. To show validity of the proposed method, reconstruction of sequences of fast moving objects such as a bursting balloon or a breaking dish was demonstrated.

Acknowledgment

This work was supported in part by SCOPE No.072103013 and Grant-in-Aid for Scientific Research No.19700098 and 21700183 in Japan.

References

- [1] J. Battle, 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
- [2] J.-Y. Bouguet, M. Weber, and P. Perona. What do planar shadows tell about scene geometry? *CVPR*, 01:514–520, 1999. 2, 4, 5, 6
- [3] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz. Space-time stereo: A unifying framework for depth from triangulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(2):296–302, Feb. 2005. 1, 2
- [4] A. Ecker, K. N. Kutulakos, and A. D. Jepson. Shape from planar curves: A linear escape from flatland. In *CVPR*, pages 1–8, 2007. 4
- [5] P. Felzenszwalb and D. Huttenlocher. Efficient belief propagation for early vision. *IJCV*, 70:41–54, 2006. 3
- [6] 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. 2
- [7] 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. 1, 2
- [8] S. Inokuchi, K. Sato, and F. Matsuda. Range imaging system for 3-D object recognition. In *ICPR*, pages 806–808, 1984. 2
- [9] 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. 2, 4, 7
- [10] 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. 2, 4, 6
- [11] H. Kawasaki and R. Furukawa. Shape reconstruction from cast shadows using coplanarities and metric constraints. In *ACCVLNCSS 4843*, volume II, pages 847–857, 2007. 2, 4
- [12] H. Kawasaki and R. Furukawa. Shape reconstruction and camera self-calibration using cast shadows and scene geometries. *IJCV*, 11 2008. 2, 4
- [13] T. P. Koninckx and L. V. Gool. Real-time range acquisition by adaptive structured light. *IEEE Trans. on PAMI*, 28(3):432–445, March 2006. 2
- [14] S. Narasimhan, S. Koppal, and S. Yamazaki. Temporal dithering of illumination for fast active vision. In *European Conference on Computer Vision*, volume 4, pages 830–844, October 2008. 2
- [15] J. Salvi, J. Battle, and E. M. Mouaddib. A robust-coded pattern projection for dynamic 3D scene measurement. *Pattern Recognition*, 19(11):1055–1065, 1998. 4, 7
- [16] P. Vuytsteke and A. Oosterlinck. Range image acquisition with a single binary-encoded light pattern. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(2):148–164, 1990. 2
- [17] T. Weise, B. Leibe, and L. V. Gool. Fast 3d scanning with automatic motion compensation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’07)*, pages 1–8, June 2007. 2
- [18] M. Young, E. Beeson, J. Davis, S. Rusinkiewicz, and R. Ramamoorthi. Viewpoint-coded structured light. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2007. 2
- [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. 4, 7, 8
- [20] L. Zhang, N. Snavely, B. Curless, and S. M. Seitz. Spacetime faces: High-resolution capture for modeling and animation. In *ACM Annual Conference on Computer Graphics*, pages 548–558, August 2004. 1, 2