

A Coded Structured Light Projection Method for High-Frame-Rate 3D Image Acquisition

Idaku Ishii
Hiroshima University
Japan

1. Introduction

Three-dimensional measurement technology has recently been used for various applications such as human modeling, cultural properties recording, machine inspection of industrial parts, and robot vision. The light-section method and coded structured light projection method are well-known active measurement methods that can accurately obtain three-dimensional shapes by projecting light patterns on the measurement space. These active measurement methods have been applied to practical systems in real scenes because they are robust regarding texture patterns on the surfaces of the objects to be observed, and have advantages in calculation time and accuracy. Following recent improvements in integration technology, three-dimensional image measurement systems that can operate at a rate of 30 fps or more have already been developed [1, 2]. In many application fields, dynamic analysis tools to observe dynamic changes in high-speed phenomena in three-dimensional shapes at high frame rates are required. Off-line high-speed cameras that can capture images at 1000 fps or more have already been put into practical use for analyzing high-speed phenomena; however, many of them can only record high-speed phenomena as two-dimensional image sequences.

In this chapter, we propose a spatio-temporal selection type coded structured light projection method for three-dimensional shape acquisition at a high frame rate. Section 2 describes the basic principle and related work on coded structured light projection methods. Section 3 describes our proposed coded structured light projection method that can alternate a temporal encoding and a spatial encoding adaptively according to the temporal changes of image intensities so as to accurately obtain a three-dimensional shape of a moving object. In section 4, several experiments were performed for several moving objects, and our proposed algorithm was evaluated by capturing their three-dimensional shapes at 1000 fps on a verification system comprising an off-line high-speed camera and a high-speed DLP projector.

2. Coded structured light projection method

Our proposed three-dimensional measurement method can be described in terms of the coded structured light projection method proposed by Posdamer et al. [3]. In this section, we describe its basic principles and the related previous coded structured light projection methods.

2.1 Basic principle

In the coded structured light projection method, a projector projects multiple black and white 'zebra' light patterns whose widths are different onto the objects to be observed, as shown in

Figure 1. A three-dimensional image is measured by capturing the projection images on the objects using a camera whose angle of view is different from that of the projector. When n patterns are projected on the objects, the measurement space is divided into 2^n vertical pieces, corresponding to the black and white areas of the zebra light patterns. Thereafter, we can obtain the n -bit data at each pixel in the projection image, corresponding to the presence of the light patterns. The n bit data is called a space code, and it indicates the projection direction. Based on the relationship between such a space code and the measurement directions that are determined by pixel positions, we can calculate depth information at all pixels of an image using triangulation.

2.2 Related work

Posdamer et al. [3] have used multiple black and white light patterns with a pure binary code as shown in Figure 1. In this case, serious encoding errors may occur, even when there is a small amount of noise, because the brightness boundaries of the multiple projection patterns with a pure binary code exist at the same positions. To solve this problem, Inokuchi et al. [4] proposed a technique for minimizing encoding errors using boundaries introduced in the form of multiple light patterns with a gray code. Bergmann [5] proposed an improved three-dimensional measurement method that combines the gray code pattern projection method and a phase shift method; in this method, the number of projection patterns increases. Caspi et al. [6] proposed an improved gray code pattern projection method that can reduce the number of projection patterns using color patterns. In these methods, three-dimensional shapes can be measured as high-resolution three-dimensional images because depth information is calculated at every pixel; however, it is difficult to accurately measure the three-dimensional shapes of moving objects because multiple light patterns are projected at different timings.

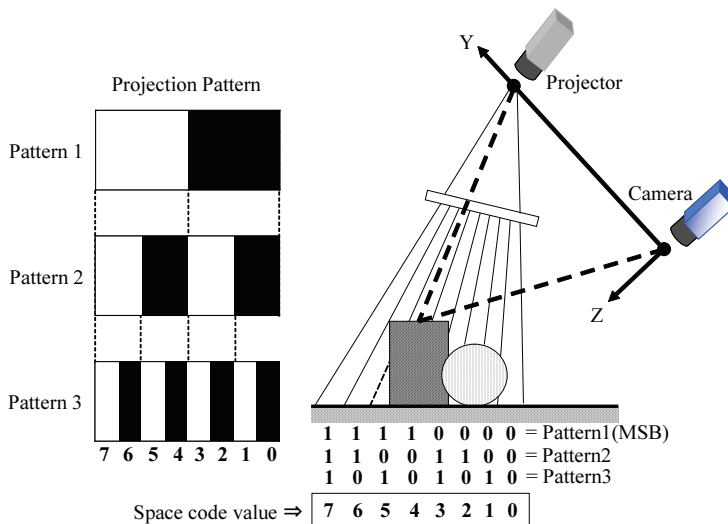


Fig. 1. Coded structured light projection method.

On the other hand, several three-dimensional measurement methods that use only a single projection pattern for spatial encoding have also been proposed. Maruyama et al. [7] proposed a method that can measure three-dimensional shapes by encoding the spatial codes on the basis of slits when a light pattern with multiple slits of different lengths is projected on the object to be observed. Durdle [8] proposed a measurement method based on a single light pattern that periodically arranges multiple slits with three grayscale levels. These methods have a disadvantage that some ambiguity exists in spatial encoding when there are pixels whose brightness are the same as that of its spatial neighborhood. To improve this ambiguity in spatial encoding, several methods introduced color slit pattern projections based on de Bruijn sequences as a robust coded pattern projection [9, 10]; this spatial encoding depends on the color surface reflectance properties of the object to be observed. These spatial encoding methods can measure the three-dimensional shapes of moving objects by projecting only a single light pattern; however, their spatial resolution is not as accurate as those obtained with the methods that use multiple light patterns because most of them assume local surface smoothness of objects in spatial encoding with a single light pattern projection.

3. Spatio-temporal selection type coded structured light projection method

3.1 Concept

As described in the previous section, the coded structured light projection method using multiple light patterns enables highly accurate three-dimensional measurement of static objects. The projection method using a single light pattern is robust when the object to be observed is moving rapidly. In this chapter, we propose a spatio-temporal selection type coded structured light projection method that attempts to combine the advantages of both temporal encoding methods and spatial encoding methods, that is, high accuracy in the case of a static object and robustness in the case of a moving object. The main features of our proposed method are in the following:

- Projection of multiple coded structured light patterns that enable both temporal encoding along the time axis and spatial encoding in the spatial domain.
- Adaptive selection of encoding types in every local image region by calculating the image features that are dependent on the measured object's motion.

Consequently, temporal encoding using multiple light patterns is selected so as to allow accurate three-dimensional measurement in the case in which there is no motion, while spatial encoding using a single light pattern is selected for robust three-dimensional measurement in the case in which the brightness changes dynamically. The concept of our proposed coded structured light projection method is shown in Figure 2.

3.2 Proposed coded structured light pattern

In the spatio-temporal selection type coded structured light projection method, when the size of a projected binary image $I(x, y, t)$ is given by $I_x \times I_y$ pixels and the space code is n bits, $I(x, y, t)$ can be expressed by the following equation:

$$I(x, y, t) = G \left(\left\lfloor \frac{x}{m} + t \right\rfloor \bmod n, y \right), \quad (1)$$

where $\lfloor x \rfloor$ is the greatest integer less than or equal to x , and m is the unit width of a light pattern in the x direction. $G(k, y)$ represents n types of one-dimensional gray code patterns

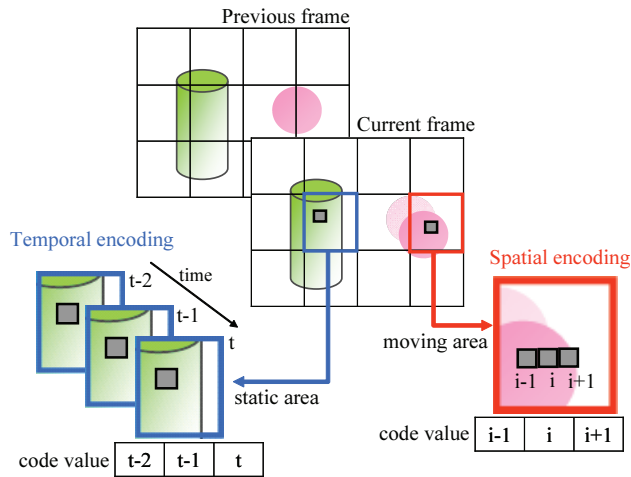


Fig. 2. Concept of our proposed algorithm.

in the y direction ($0 \leq k \leq n - 1$), which can minimize encoding errors on code boundaries. $G(k, y)$ is given by:

$$G(k, y) = \left\lfloor \frac{y \times 2^k}{I_y} + \frac{1}{2} \right\rfloor \bmod 2. \quad (2)$$

The coded pattern defined by Eq.(1) has the following properties:

- The coded pattern has a periodic branched pattern based on the gray code.
- The coded pattern shifts in the x direction over time.

These features of the coded pattern enable spatio-temporal selection type encoding that can select not only encoding along the time axis, but also encoding in the spatial domain.

As an example of the coded patterns, the coded patterns at time $t = 0, 1, 2$, and 3 are shown in Figure 3 when the number of bits of the space codes is set at $n = 8$. The numbers $k (= 0, \dots, 7)$ at the bottom of each image indicate that the one-dimensional pattern in the y direction is set to $G(k, y)$. In Figure 3, the space code can be generated only using spatial neighborhood information because eight gray code patterns from $G(0, y)$ to $G(7, y)$ are periodically arranged in a single image. From Eq.(1), the projection pattern also shifts to the left after a certain period of time; the gray code pattern on the same pixel changes $G(0, y), G(1, y), G(2, y), \dots$, every unit time. Similarly, a space code can be obtained at each pixel along the time axis.

3.3 Spatio-temporal selection type encoding algorithm

3.3.1 Calibration between a projection pattern and a captured image

When the shifting light pattern is projected using a projector to measure the three-dimensional shape of objects; the projection results are captured as an image by a camera whose angle of view is different from that of the projector. To calculate space codes using this captured image,

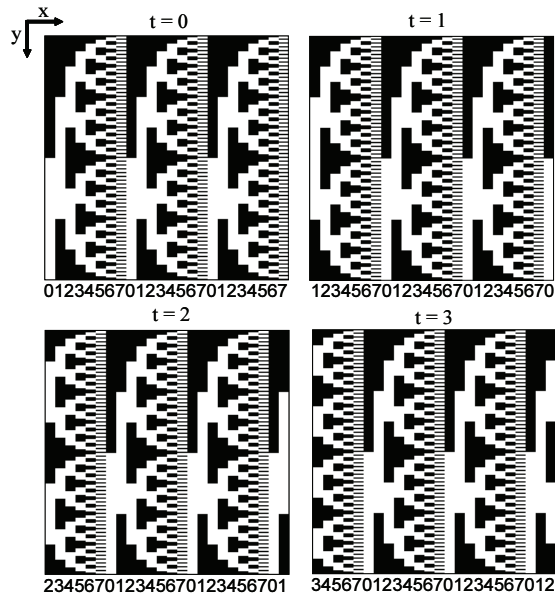


Fig. 3. Coded structured light patterns for spatio-temporal selection.

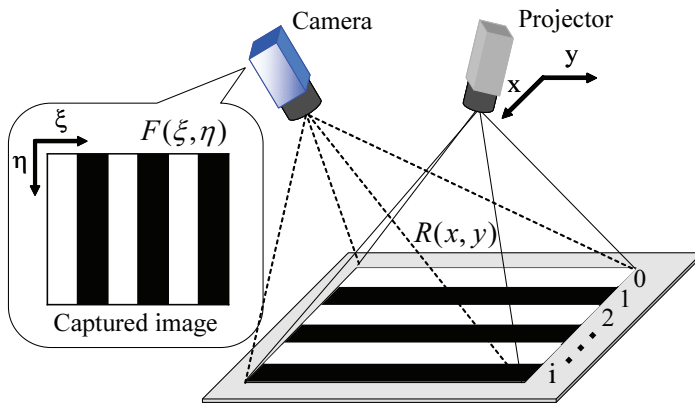


Fig. 4. Spatial relationship between a projector and a camera.

the type of gray code patterns expressed in Eq. (1), $G(0, y), \dots, G(7, y)$, that is projected at each pixel on the captured image must be identified after the captured image has been binarized.

Figure 4 shows the spatial relationship between the xy -coordinate system of a projector and the $\xi\eta$ -coordinate system of a camera. It is assumed that the x -axis and ξ -axis are parallel for light patterns projected from the projector and images captured by the camera.

Here, a region corresponding to a gray code pattern $G(k, y)$ is assumed to be a rectangular area in the captured image. When the number of rectangular areas is assumed to be i , the i -th rectangular area r_i on the projector coordinate system can be defined as follows:

$$r_i = \left\{ (x, y) \mid \left\lfloor \frac{x}{m} \right\rfloor = i \right\}. \quad (3)$$

Here, a rectangular area r_i is assumed to be projected onto a rectangular area r'_i in the camera coordinate system. To make a correspondence between r_i and r'_i , the following light pattern $R(x, y)$ is projected as a reference pattern that divides the measurement space on the captured image into multiple rectangular areas:

$$R(x, y) = \left\lfloor \frac{x}{m} + 1 \right\rfloor \bmod 2. \quad (4)$$

$F(\xi, \eta)$ is the image captured when the reference light pattern $R(x, y)$ is projected; the region number at pixel (ξ, η) can be matched with that at $h(\xi, \eta)$. $h(\xi, \eta)$ is defined as the number of switching obtained by counting the changes in brightness until reaching pixel (ξ, η) when scanning $F(\xi, \eta)$ in the positive direction of the ξ -axis. A rectangular area r'_i on the camera coordinate system can be assigned as follows:

$$r'_i = \{(\xi, \eta) \mid h(\xi, \eta) = i\}. \quad (5)$$

In the space encoding discussed below, the initially given map $h(\xi, \eta)$ can uniquely specify the number of a rectangular area i at a pixel (ξ, η) in the captured image via Eq. (5). Thus, we can judge which gray code pattern of $G(0, y), \dots, G(7, y)$ is projected on pixel (ξ, η) at time t .

3.3.2 Space encoding types

Next, we put forward a method of space code calculation that can select coded patterns temporally and spatially according to brightness changes in the captured images, which are strongly dependent upon an object's motion in the images. Figure 5 shows the types of gray code patterns that are projected onto a part of the given line for eight frames when the light patterns defined by Eqs. (1) and (2) are projected. In the figure, the vertical axis represents time, and the horizontal axis represents the number of rectangular areas. We show an example of spatio-temporal selection type encoding at a given pixel (ξ, η) . There are four bits in the time direction and three bits in the space direction; they are referred to to obtain space code values at pixel (ξ, η) . Thus, a space code value at pixel (ξ, η) in a rectangular area i uses not only temporal encoding or spatial encoding in a single direction, but also spatio-temporal selection type encoding (which refers image information both temporally and spatially).

In this study, we introduce n selectable space code values ${}^p X(\xi, \eta, t)$ at pixel (ξ, η) at time t , whose number of referred bits in time and space (excluding the specified pixel) are $(n - 1, 0), (n - 2, 1), \dots, (0, n - 1)$, respectively. Here, n is the number of bits of space code, and $p (= 1, \dots, n)$ is an alternative parameter for space coding; the space encoding is close to the space pattern selection when p is small, whereas it is close to the temporal one when p is large.

Here, $g(\xi, \eta, t)$ is a binarized image obtained from a camera at time t ; it corresponds to coded structured light patterns. The binarized image $g(\xi, \eta, t)$ belonging to a rectangular area i

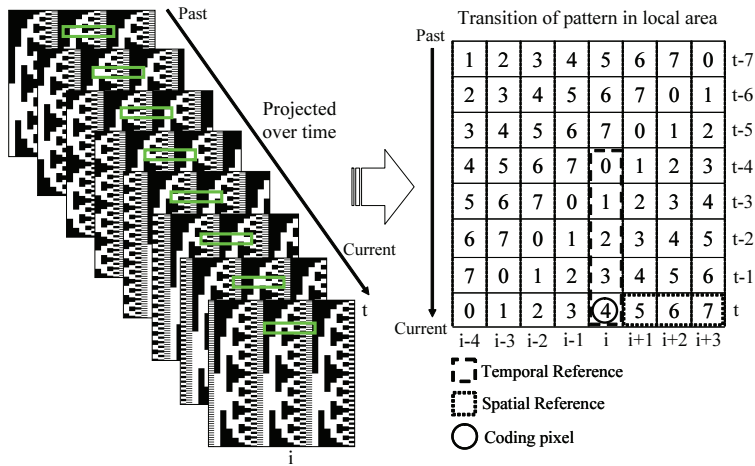


Fig. 5. Spatio-temporal selection type encoding ($p = 4$).

is abbreviated as ${}^i g_t$. The space code ${}^p X(\zeta, \eta, t)$ is abbreviated as ${}^p X_t$. Eqs. (6)–(13) are enumerated as examples of selectable space code values when $n = 8$.

$${}^1 X_t = {}^i g_{t-4} {}^i g_{t-3} {}^i g_{t-2} {}^i g_{t-1} {}^i g_t {}^i g_{t-7} {}^i g_{t-6} {}^i g_{t-5}, \tag{6}$$

$${}^2 X_t = {}^i g_{t-4} {}^i g_{t-3} {}^i g_{t-2} {}^i g_{t-1} {}^i g_t {}^{i+1} g'_t {}^i g_{t-6} {}^i g_{t-5}, \tag{7}$$

$${}^3 X_t = {}^i g_{t-4} {}^i g_{t-3} {}^i g_{t-2} {}^i g_{t-1} {}^i g_t {}^{i+1} g'_t {}^{i+2} g'_t {}^i g_{t-5}, \tag{8}$$

$${}^4 X_t = {}^i g_{t-4} {}^i g_{t-3} {}^i g_{t-2} {}^i g_{t-1} {}^i g_t {}^{i+1} g'_t {}^{i+2} g'_t {}^{i+3} g'_t, \tag{9}$$

$${}^5 X_t = {}^{i-4} g'_t {}^i g_{t-3} {}^i g_{t-2} {}^i g_{t-1} {}^i g_t {}^{i+1} g'_t {}^{i+2} g'_t {}^{i+3} g'_t, \tag{10}$$

$${}^6 X_t = {}^{i-4} g'_t {}^{i-3} g'_t {}^i g_{t-2} {}^i g_{t-1} {}^i g_t {}^{i+1} g'_t {}^{i+2} g'_t {}^{i+3} g'_t, \tag{11}$$

$${}^7 X_t = {}^{i-4} g'_t {}^{i-3} g'_t {}^{i-2} g'_t {}^i g_{t-1} {}^i g_t {}^{i+1} g'_t {}^{i+2} g'_t {}^{i+3} g'_t, \tag{12}$$

$${}^8 X_t = {}^{i-4} g'_t {}^{i-3} g'_t {}^{i-2} g'_t {}^{i-1} g'_t {}^i g_t {}^{i+1} g'_t {}^{i+2} g'_t {}^{i+3} g'_t, \tag{13}$$

where ${}^i g'_t = {}^i g'_t(\zeta, \eta, t)$ is a value obtained via spatial neighborhood processing, and which is set to 0 or 1. The value 1 is taken when the numbers of 0, 1 of the binarized image $g(\zeta, \eta, t)$ are counted in the ζ direction for the nearest i -th rectangular area to pixel (ζ, η) .

Figure 6 shows eight space encoding types as defined by Eqs. (6)–(13). The example in Figure 5 is a case of $p = 4$: four bits are referred to in the time direction and four bits are referred to in the spatial neighborhood. When the objects to be observed are not in motion, space encoding that refers only to information along the time axis ($p = 1$) is effective, that is, it is equivalent to the conventional gray code pattern projection method [4]. In the case of $p = 8$, the space encoding type, it refers only to information in a single projected image; this encoding type is effective in measuring the three-dimensional shapes of moving objects because its accuracy is completely independent of the motion of objects. However, there is a disadvantage in that measurement errors increase when undulated shapes are measured because the space encoding measurement when $p = 8$ assumes spatial smoothness for the shape.

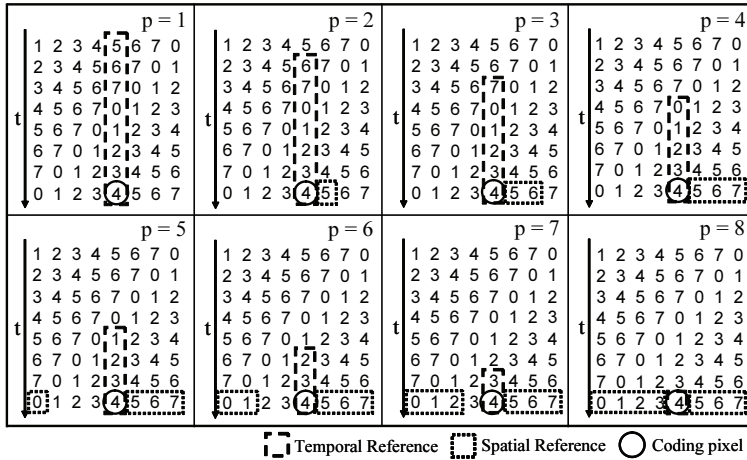


Fig. 6. Types of encoding.

3.3.3 Adaptive selection of encoding types

Considering the brightness changes in space code images, we introduce a criterion for the adaptive selection of encoding types defined by Eqs. (6)–(13). In the introduced criterion, we use the frame differencing feature obtained by differentiating a space code image $T(\xi, \eta, t)$ at time t from a space code image $S(\xi, \eta, t - 1)$ at time $t - 1$. $T(\xi, \eta, t)$ is the space code image that refers only to information along the time axis, and $S(\xi, \eta, t - 1)$ is the space code image in a previous frame selected using the following equation:

$$D(\xi', \eta', t) = \sum_{(\xi, \eta) \in q(\xi', \eta')} T(\xi, \eta, t) \oplus S(\xi, \eta, t - 1), \tag{14}$$

where \oplus refers to the logical exclusive OR operation, and $q(\xi', \eta')$ refers to the ξ' -th and η' -th $s \times s$ -pixel block area in the ξ and η directions, respectively. $T(\xi, \eta, t_0)$ is initially set to $S(\xi, \eta, t_0 - 1)$ at the start time t_0 to generate the space codes. After differentiation, the criterion $D(\xi', \eta', t)$ is provided for every divided square image region (each unit of which is $s \times s$ pixels), by calculating the summation of the differentiation result in Eq. (14). $D(\xi', \eta', t)$ corresponds to the number of pixels where the values of $T(\xi, \eta, t)$ are different from those of $S(\xi, \eta, t - 1)$ in the block area $q(\xi', \eta')$.

When moving objects are measured, the space code image $T(\xi, \eta, t)$ is not always encoded correctly because it refers only to the information along the time axis, which requires multiple images at different frames. However, the space code image $S(\xi, \eta, t)$ generated by the spatio-temporal selection type encoding defined by Eq. (15) is robust to errors caused by moving objects. Thus, the criterion $D(\xi', \eta', t)$ can be defined with a frame differencing calculation for every block area to detect the coding errors caused by motion.

Based on the criterion $D(\xi', \eta', t)$, the spatio-temporal encoding type $p = p(\xi', \eta', t)$ is selected from the encoding types ${}^p X_t$, which are defined by Eqs. (6)–(13) for every block area $q(\xi', \eta')$ at time t . In this study, a spatio-temporal selection type space code image $S(\xi, \eta, t)$ at various

times t is determined using the calculated encoding types for all the block areas as follows:

$$p(\xi', \eta', t) = \begin{cases} n & (D(\xi', \eta', t) = s^2) \\ \left\lfloor \frac{nD(\xi', \eta', t)}{s^2} \right\rfloor + 1 & (\text{otherwise}) \end{cases} \quad (15)$$

Figure 7 shows an example of our newly introduced criterion for spatio-temporal encoding

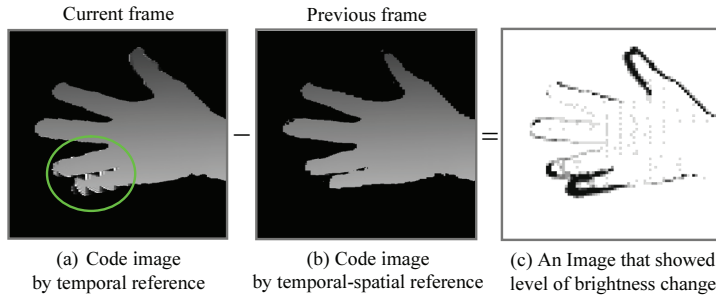


Fig. 7. Criterion for spatial-temporal encoding.

when a human extends and contracts his fingers. In the figure, (a) a space code image $T(\xi, \eta, t)$ coded using temporal reference, (b) a space code image $S(\xi, \eta, t - 1)$ coded using temporal-spatial reference in a previous frame, and (c) selected spatial encoding types $p(\xi', \eta', t)$, are shown. In (c), the value of p increases when the gray-level tone becomes darker. It can be observed that the number of the spatio-encoding type, p , increases if code errors are caused by the motion of fingers, whereas p decreases in the area in which motion is slight. This fact indicates that space codes that are robust to motion are selected for the dynamically changing scenes, and space codes that enable accurate shape measurement are selected for the static scenes. Thus, the spatio-temporal encoding type can be selected both adaptively and according to the defined criterion based on frame differencing features for space code images.

3.4 Correction of misalignment in the projected image

Next, we consider a misalignment problem in the projected image, and introduce a correction method to reduce misalignment errors in space code images. Figure 8 shows a framework for our method of correcting the misalignment of the projected image. When ideally projecting, as shown in (a), the projected black and white zebra patterns are accurately matched with the initially assigned rectangular areas r'_i , and we can always observe the same value of 0 or 1 at all the pixels on a certain line segment in the ξ direction in the same rectangular area when the projected zebra pattern has a width of several pixels. However, when there is a certain displacement between the projected zebra patterns and the initially assigned rectangular areas, as shown in (b), there are cases in which there are both black and white pixels on a certain line segment, even in the same rectangular area. This ambiguity may generate errors, especially around the edge boundaries of the black and white zebra patterns, when computing space code images in the coded structure light projection method.

To avoid this ambiguity (caused by the displacement between the projected zebra patterns and the initially assigned rectangular areas), the pixel values of 0 or 1 in the same rectangular

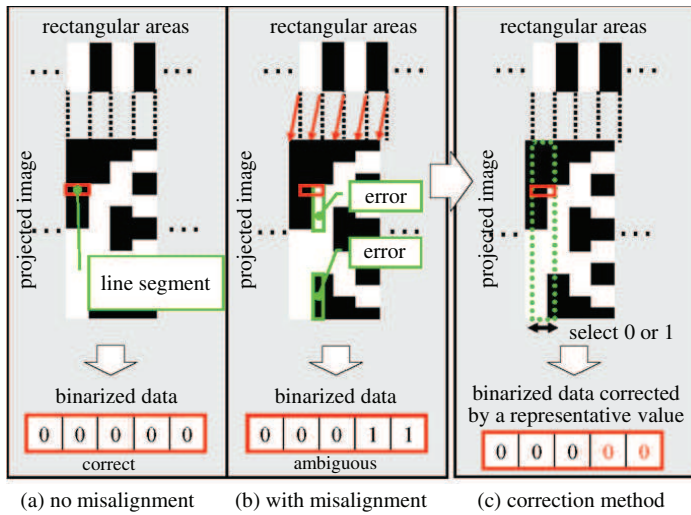


Fig. 8. Correction of misalignment in the projected image.

area are calibrated by replacing with a representative value on a certain line segment in the ζ direction; the representative value on a certain line segment in the ζ direction is determined to have the pixel value of 0 or 1 when its number surpasses that of its opposite value on the given line segment in the same rectangular area, as shown in (c). Thus, we can reduce the influence of the misalignment of the projected image and correct the correspondence between the projected zebra patterns and the initially assigned rectangular areas.

As an example, Figure 9 shows the corrected space code images when a human hand is observed using our structured light projection method. Before correction, many slit-like noises can be observed around the edge boundaries of the black and white stripe patterns. By introducing the correction process, we can reduce this slit-like noise and obtain the correct space code images, which correspond to the fact that the hand shape smoothly varies in space.

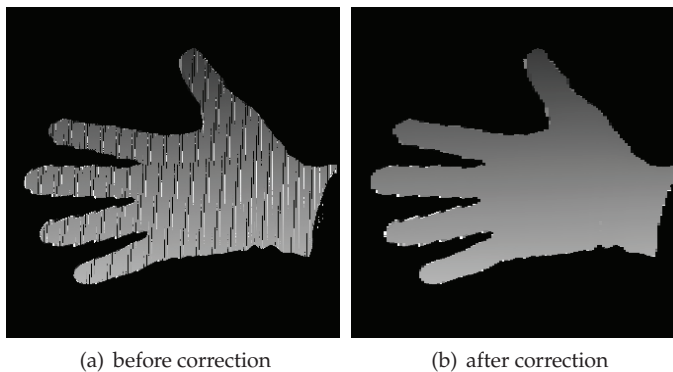


Fig. 9. Corrected space code images for a human hand.

3.5 Calculation of three-dimensional information

To obtain three-dimensional information from the space code image $S(\xi, \eta, t)$, it is necessary to transform $S(\xi, \eta, t)$ using the relationship between the location of the camera and the projection; this requirement was common in the previously-reported structured light pattern projection methods. In this chapter, the space code image is transformed into three-dimensional information at each pixel using the method described in [4].

The three-dimensional coordinate $\mathbf{X}(t)$ is obtained from the pixel position (ξ, η) on the camera. Its corresponding space code value $S(\xi, \eta, t)$ is obtained by solving the following simultaneous equation with a 3×4 camera transform matrix \mathbf{C} and a 2×4 projector matrix \mathbf{P} ,

$$H_c \begin{pmatrix} \xi \\ \eta \\ 1 \end{pmatrix} = \mathbf{C} \begin{pmatrix} \mathbf{X}(t) \\ 1 \end{pmatrix}, \quad (16)$$

$$H_p \begin{pmatrix} S(\xi, \eta, t) \\ 1 \end{pmatrix} = \mathbf{P} \begin{pmatrix} \mathbf{X}(t) \\ 1 \end{pmatrix}, \quad (17)$$

where H_c and H_p are parameters, and the camera transform matrix \mathbf{C} and the projector matrix \mathbf{P} must be obtained by prior calibration. All of the pixels in the image are transformed using Eqs. (16) and (17). We can obtain a three-dimensional image as the result of our proposed coded structured pattern light projection method.

4. Experiments

4.1 Experimental system

To verify the three-dimensional measurement method proposed in the previous section, off-line experiments for three-dimensional shape measurement are performed using an off-line high-speed camera and a high-speed DLP projector. Figures 10 and 11 show its location configuration and an overview, respectively. We used an off-line high-speed video camera, FASTCAM-1024PCI (Photron Ltd., Tokyo, Japan), that can capture a 1024×1024 -pixel 10-bit gray-level image at 1000 fps. The high-speed DLP projector comprises a DMD (Direct Micro-mirror Device) Discovery 1100 controller board (Texas Instruments Inc., Dallas, USA), an ALP controller module (ViALUX GmbH, Chemnitz, Germany), and an LED

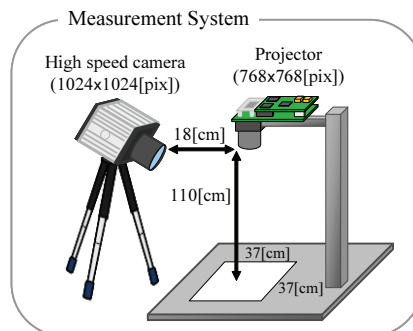


Fig. 10. Configuration of experimental system.

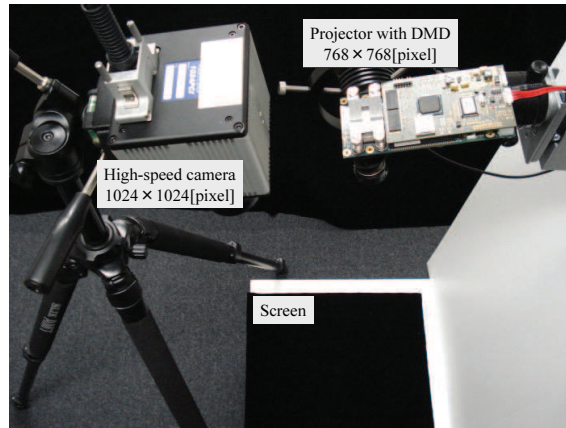


Fig. 11. Overview of experimental system.

Integrated Optical Module (ViALUX GmbH, Chemnitz, Germany). The projector can project 1024×768 -pixel binary images at a frame rate of 1000 fps or more.

In the experimental setup, the camera is located at a distance of approximately 18 cm from the projector; and a background screen is placed at a distance of approximately 110 cm from the camera and projector. Light patterns from the projector are projected over an area of $37 \text{ cm} \times 37 \text{ cm}$ on the screen. Here, the camera is positioned such that the rectangular areas in the x direction are not distorted, even if the heights of the measured objects vary. In the experiments, a 768×768 -pixel light pattern defined by Eq. (1) is projected at 1000 fps. The number of bits of space codes and the width of rectangular areas are set to $n = 8$ and $m = 4$, respectively. The high-speed camera is synchronized with the projector electrically, a 1024×1024 -pixel image is captured at 1000 fps so that the projected light patterns are settled in the captured image. To reduce errors caused by illumination or surface properties, the image for space encoding is binarized by differentiating a pair of positive and negative projection images that are generated by projecting a consecutive pair of light patterns whose values of black and white are reversed. The size of the block area is set to 8×8 pixels ($s = 8$) to calculate the criterion $D(\xi', \eta', t)$ for the selection of encoding types.

4.2 Experimental results

To demonstrate the effectiveness of our proposed algorithm using the off-line high-speed camera and the high-speed DLP projector, we show three-dimensional shape measurement results for three types of moving objects in Figure 12; (a) a rotating plane, (b) a waving piece of paper, and (c) a moving human hand. First, we show the measured result for a rotating plane object; the measure object is a $8.5 \text{ cm} \times 10.0 \text{ cm}$ plate. It is rotated at approximately 5.5 revolutions per second (rps) by a DC motor. Figure 13 shows the images captured by a video camera at a frame rate of 30 fps on the left; the three-dimensional color-mapped images calculated using our proposed method are on the right. The image interval is set to 0.03 s. It can be observed that three-dimensional shape information can be obtained at a high frame rate even when the measured object moves too quickly for the human eye to see it.

Next, we show the measured result for a waving piece of paper; the measure object is an A4 paper of 297 mm × 210 mm size. The piece of paper is waved in the vertical direction. Figure 14 shows the images captured by a video camera at a frame rate of 30 fps on the left;

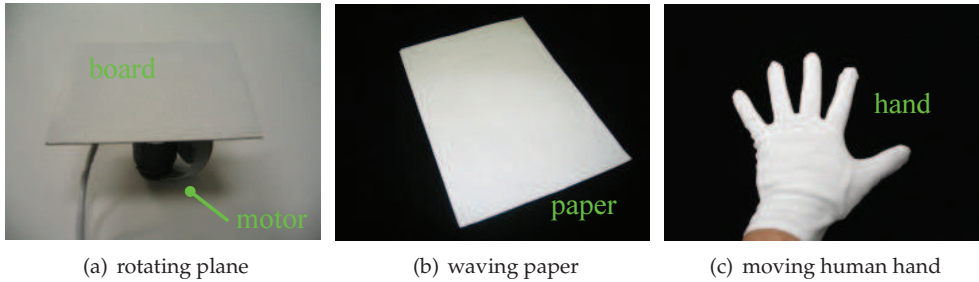


Fig. 12. Measurement objects.

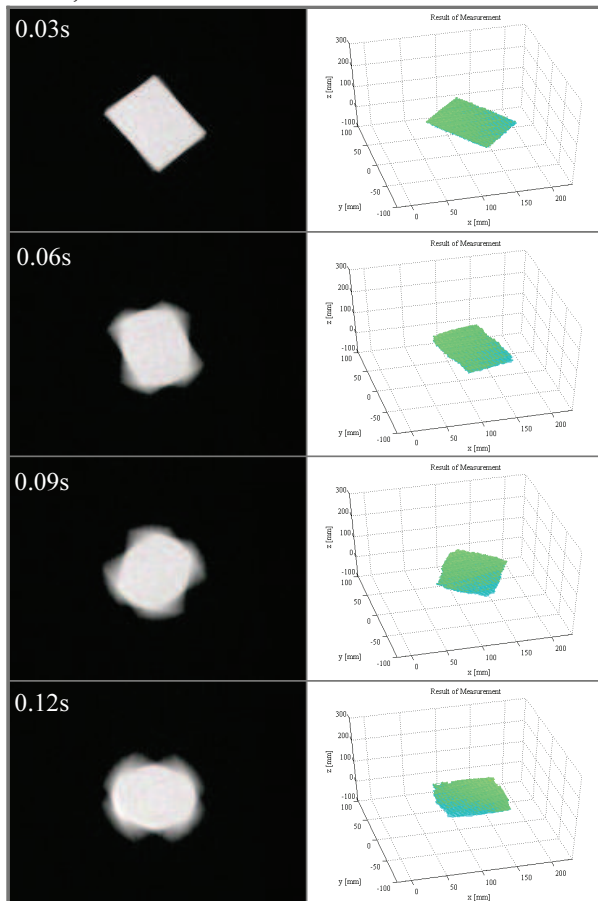


Fig. 13. Experimental result for a rotating plane.

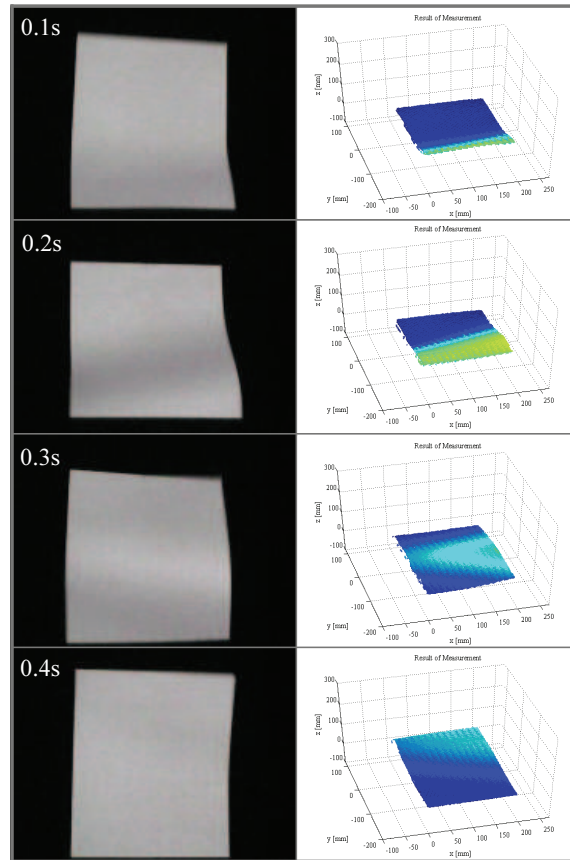


Fig. 14. Experimental result for a waving piece of paper.

the three-dimensional color-mapped images calculated using our proposed method are on the right. The image interval is set to 0.1 s. It can be observed that the temporal changes in the three-dimensional shape, which are generated by wave propagation, become visible in detail.

Finally, we show the measured result for a human hand whose fingers move quickly. Here, a human changes his fingers' shape from "paper" to "scissors" in rock-scissor-paper game with a duration of approximately 0.1 s. Figure 15 shows the images captured by a video camera at a frame rate of 30 fps on the left; the three-dimensional color-mapped images calculated using our proposed method are on the right. Their interval is set to 0.03 s. It can be observed that the three-dimensional shape of a human hand, a complicated shape, can be acquired, and that quick movements of the fingers or changes in the height of the back of the hand can also be accurately detected in the form of a three dimensional shape.

Figure 16 shows the examples of the selected spatial encoding types in these experiments. When the gray-level tone becomes darker, p increases, that is, the spatial encoding becomes more dominant. When the tone becomes brighter, p decreases, that is, the temporal encoding

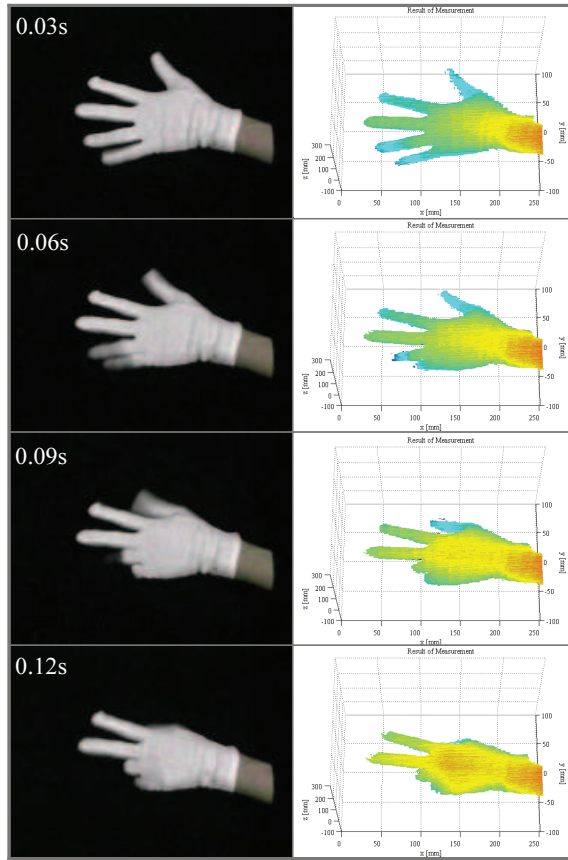


Fig. 15. Experimental result for a moving human hand

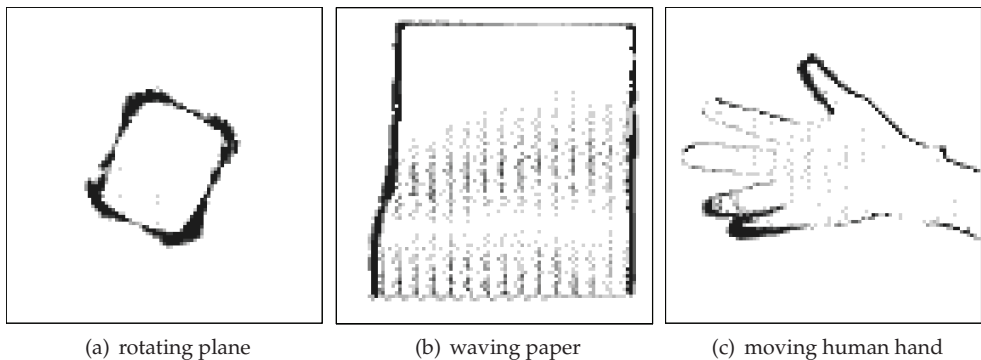


Fig. 16. Selected spatial encoding types spatio-temporal encoding types.

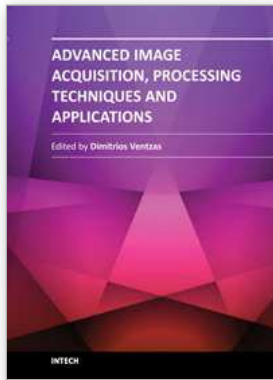
becomes more dominant. In all the snapshots, the dark pixels exist around the edges of the moving objects, because the depth changed primarily at the pixels around these spatial discontinuities (between the object and its background); temporal selection type encoding was mainly performed on the inner side of the object, where the depth largely did not change. In the snapshots for the waving piece of paper, it can be observed that the selected spatial encoding types varied spatio-temporally with the waving movement.

5. Conclusion

In this chapter, a spatio-temporal selection type coded structured light projection method is proposed for the acquisition of three-dimensional images at a high frame rate. The proposed method can select adaptive space encoding types according to the temporal changes in the code images. Our proposed method was verified off-line using a testbed that was composed of a high-speed camera and a high-speed projector. We evaluated its effectiveness by producing experimental results for various three-dimensional objects moving quickly, at a high frame rate such as 1000 fps, which is too fast for the human eye to observe their three-dimensional motion in detail.

6. References

- [1] S. Yoshimura, T. Sugiyama, K. Yonemoto, and K. Ueda. "A 48kframes/s CMOS image sensor for real-time 3-D sensing and motion detection," in *ISSCC Tech. Dig.*, 94–95, 2001.
- [2] Y. Oike, M. Ikeda, and K. Asada. "A CMOS image sensor for high-speed active range finding using column-parallel time domain ADC and position encoder," *IEEE Trans. Electron Devices*, 50(1), 152–158, 2003.
- [3] J.L. Posdamer and M.D. Altschuler. "Surface measurement by space-encoded projected beam systems," *Computer Graphics and Image Processing*, 18(1), 1–17, 1982.
- [4] S. Inokuchi, K. Sato, and F. Matsuda. "Range imaging system for 3-D object recognition," in *Proc. Int. Conf. Pattern Recog.*, 806–808, 1984.
- [5] D. Bergmann. "New approach for automatic surface reconstruction with coded light," in *Proc. Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*, SPIE 2572, 2–9, 1995
- [6] D. Caspi, N. Kiryati, and J. Shamir. "Range imaging with adaptive color structured light," *IEEE Trans. Pattern Anal. Mach. Intelli.*, 20 (5), 470–480, 1998.
- [7] M. Maruyama and S. Abe. "Range sensing by projecting multiple slits with random cuts," *IEEE Trans. Pattern Anal. Mach. Intelli.*, 15 (6), 647–651, 1993
- [8] N.G. Durdle, J. Thayoor, and V.J. Raso. "An improved structured light technique for surface reconstruction of the human trunk," in *Proc. IEEE Can. Conf. Elect. Comput. Eng.*, 874–877, 1998.
- [9] L. Zhang, B. Curless, and S.M. Seitz. "Rapid shape acquisition using color structured light and multi-pass dynamic programming," in *Proc. Int. Symp. 3D Data Processing Visualization and Transmission*, 24–36, 2002.
- [10] J. Salvi, J. Battle, and E. Mouaddib. "A robust-coded pattern projection for dynamic 3d scene measurement," *Int. J. Patt. Recog. Lett.*, 19, 1055–1065, 1998.



Advanced Image Acquisition, Processing Techniques and Applications I

Edited by Dr. Dimitrios Ventzas

ISBN 978-953-51-0342-4

Hard cover, 170 pages

Publisher InTech

Published online 14, March, 2012

Published in print edition March, 2012

"Advanced Image Acquisition, Processing Techniques and Applications" is the first book of a series that provides image processing principles and practical software implementation on a broad range of applications. The book integrates material from leading researchers on Applied Digital Image Acquisition and Processing. An important feature of the book is its emphasis on software tools and scientific computing in order to enhance results and arrive at problem solution.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Idaku Ishii (2012). A Coded Structured Light Projection Method for High-Frame-Rate 3D Image Acquisition, Advanced Image Acquisition, Processing Techniques and Applications I, Dr. Dimitrios Ventzas (Ed.), ISBN: 978-953-51-0342-4, InTech, Available from: <http://www.intechopen.com/books/advanced-image-acquisition-processing-techniques-and-applications-i/a-coded-structured-light-projection-method-for-high-frame-rate-3d-image-acquisition>

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.