

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Feature enhancement for a defocusing structured-light 3-D scanning system

Yu-Lun Liu
Tzung-Han Lin

Feature enhancement for a defocusing structured-light 3-D scanning system

Yu-Lun Liu^a and Tzung-Han Lin^{b,c,*}

^aNational Taiwan University of Science and Technology, Graduate Institute of Applied Science and Technology, Taipei, Taiwan

^bNational Taiwan University of Science and Technology, Graduate Institute of Color and Illumination Technology, Taipei, Taiwan

^cNational Taiwan University of Science and Technology, High Speed 3D Printing Center, Taipei, Taiwan

Abstract. Structured-light systems consisting of a camera and projector are powerful and cost-effective tools for three-dimensional (3-D) shape measurements. However, most commercial projectors are unable to generate distinct patterns due to defocusing and shallow focusing issues. We propose a hybrid method for enhancing the calibration and scanning features of the defocusing structured-light 3-D scanning system. Instead of using conventional sequential binary patterns, we replace the highest-level binary pattern by a high-order sinusoidal pattern. In our proposed system, a pan-tilt stage carrying a checkerboard is used to assist the simultaneous calibration of the camera and projector. Initially, the camera is calibrated to obtain the extrinsic positions of the stage. In addition, we utilize the multiplication of vertical and horizontal stripe patterns to enhance the corresponding features between the camera and projector. The projector is then calibrated using the extrinsic features determined from the calibrated camera. The experimental results show that the use of the high-order sinusoidal pattern significantly improves reprojection error. Our proposed method can easily be incorporated in the defocusing projector for scanning various types of objects. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.OE.57.6.064101](https://doi.org/10.1117/1.OE.57.6.064101)]

Keywords: structured light; out of focus; feature enhancement.

Paper 180158 received Jan. 29, 2018; accepted for publication May 15, 2018; published online Jun. 1, 2018.

1 Introduction

Optical three-dimensional (3-D) measuring systems have reached the mainstream for 3-D shape measurements, and they have been widely adopted in various fields.¹ The existing 3-D measurement techniques can be categorized according to several types, such as stereoscopy, laser triangulation, structured light or fringe projection, and time of flight (ToF). All of these methods have been extensively employed due to their noncontact and nondestructive features with regard to the physical probes of the coordinate measurement machine (CMM). For example, a ToF camera obtains the distance by utilizing the traveling time required for a light signal to pass between the camera and subjects. The other methods, such as laser triangulation, stereoscopy, and structured light, utilize projective geometry to resolve distance measurements. The stereoscopy method usually utilizes the disparities of rectified stereo images to obtain 3-D shapes. Parallel binocular camera configurations, as well as stereo image after rectification, induce parallel epipolar lines for accelerating the feature matching process in the stereoscopy method. However, robustly determining the correspondences between two images, particularly in subpixel level, is still a challenging task. To simplify this task, the structured-light system, as well as the laser triangulation method, actively projects multiple stripes on a calibration object, and the cast stripes are correlated to the given stripes. Moreover, precise subpixel localizations of stripes can be obtained using super-resolution algorithms.²

The structured-light system is considered to be one of the most efficient techniques for recovering 3-D points of objects. In the last few decades, there were several coded structured-light systems revealed for a variety of applications. For example, Salvi et al.³ presented a definitive classification and comparison review of existing structured-light techniques. A comprehensive tutorial for assisting beginners in constructing various types of 3-D scanners was first provided by Lanman and Taubin.⁴ Since the 3-D scanning technique is interdisciplinary, not only the hardware design but also the software algorithms are highly dependent on the particular system. Gorthi and Rastogi proposed an overview for state-of-the-art fringe projection techniques that mostly utilize phase-shifting algorithms.⁵ In addition, Geng⁶ focused on a comprehensive comparison of structured-light 3-D imaging systems, and several practical applications were illustrated. Among the existing structured-light systems, the time-multiplexing binary code, which produces distinct transitional boundaries between neighboring stripes, is one of the most commonly used techniques.

In our proposed system, binary-coded patterns are used and sequentially projected, in a manner similar to that of Valkenburg and McIvor.⁷ Moreover, our proposed scanning module consists of a monochromatic camera and a digital light projector (DLP) that are utilized for obtaining the 3-D shape of an object on a pan-tilt stage. For such a closed system, calibration using an external tool is difficult. Therefore, an automatic selfcalibration method is necessary.

2 Related Calibration and Measurement Systems

Calibration is the most critical procedure for 3-D shape measurement systems. Most of the noncontact measurement

*Address all correspondence to: Tzung-Han Lin, E-mail: thl@mail.ntust.edu.tw

methods rely on a digital camera for calibration and measurement. To calibrate the camera, a perspective projection of the pinhole model along with radial lens distortion is used to obtain the intrinsic parameters. The most commonly used camera calibration method is the method proposed by Tsai.⁸ To extend the flexibility of Tsai's method, Zhang⁹ proposed an algorithm based on homography transformations of multiple flat planes. Circular control points were further utilized by Heikkilä¹⁰ to improve the calibration accuracy. These calibration methods have developed into mature, well-established methods for precise 3-D computer vision applications.

In recent decades, several approaches have been developed for calibrating structured-light systems. Because the projector can function as an inverse camera, the projector has been commonly treated as a camera for calibration due to the mature techniques of the camera calibration. However, the calibration procedure must be modified due to the fact that the projector cannot directly measure the pixel coordinates of 3-D points in the same way as a camera. Therefore, to calibrate a structured-light system, several methods have been proposed, such as stereovision based on geometrical constraints¹¹ and plane constraints based on a calibrated camera.¹² An inverse camera calibration workflow, which adjusts the projected points so they are consistent with the physical cross corners, was proposed by Martynov et al.¹³ On the other hand, Gao et al.¹⁴ utilized an additional color cue to identify the real black-white and the cast red-blue checkerboards. Similarly, Ouellet et al.¹⁵ proposed a geometric calibration method based on circular dots, whereas Anwar et al.¹⁶ used virtual target images to calibrate the projector. Moreover, Orghidan et al.¹⁷ utilized the vanishing points to simplify the calibration procedure. Markerless-based projector calibration, which only uses a white plate, has also been proposed as a flexible solution.¹⁸

In their papers, the homography between the camera and a physical plane is initially determined. Then, the features projected from the projector are observed by the camera to determine their homography. The homography between the physical plane and projector is obtained by decomposing the first two homographies, and the calibration for the intrinsic and extrinsic parameters can be readily obtained. Nevertheless, in this method, the accuracy of the calibrated projector likely suffers from calibration errors of the camera.

Lens distortion is a nonlinear behavior due to optical refraction of lenses and the misalignment of assembled parts. Therefore, structured-light systems can suffer from lens distortions of both the camera and the projector. To overcome this nonlinear behavior, a local homography method was proposed by Moreno and Taubin.¹⁹ In this method, the transformation between the camera and the projector is decomposed into several linear transformations. Because the camera lens distortion is usually formulated as a high-order polynomial, the distortion effect can be readily removed by correcting the captured images.²⁰ Lens distortion of the projector can be corrected by applying an inverse distorted image to the projector.^{21,22} Therefore, the calibration accuracy can be considerably improved by eliminating the lens distortions in a structured-light system.

Unlike lens-distortion compensation, a look-up table (LUT), which records the transformation between the physical 3-D ground truth and the observed stripes, is also

a feasible solution for achieving precise 3-D measurement.²³ Since LUT can be regressed by high-order surface equations or a volumetric function, the distortion behavior in the camera and projector is precisely recorded. Therefore, without knowing the intrinsic and extrinsic parameters of both components, the structured-light system is still able to obtain accurate 3-D points.

Most existing methods require a good focusing projector to function as an inverse pinhole camera. However, the depth of focus is limited by the finite size of the aperture. For an out-of-focus projector, the projected features are difficult to identify. As a result, the system may suffer from ambiguous cast stripe boundaries, which can substantially enlarge the variance of the measurement results. To overcome this problem, an out-of-focus projector calibration, which can produce consistent results under various defocusing degrees, was proposed by Li et al.²⁴

A complete 3-D model usually requires several scans from multiple directions in addition to 3-D shape registration or extrinsic calibration. The iterative closest-point (ICP) method, which performs closed-form rigid-body transformations, is commonly utilized for 3-D shape registration.²⁵ The ICP method needs a good, close initial position to search for corresponding 3-D features. Barone et al.²⁶ utilized an external tracking device to achieve global registration. Therefore, a large object, such as a statue, can be completely integrated from multiple scans. For a specific distribution of the relative scanning positions, such as an object on a turntable, Pang et al.²⁷ utilized a global registration method. Similarly, our system uses a pan-tilt stage for carrying the scanned object, and the rotating axes are calibrated by retrieving the extrinsic parameters for each position on a checkerboard.

Finding correspondences is another critical task for a structured-light system. A variety of coded patterns have been proposed in which the projected pattern carries unique information of the position with respect to the projector coordinates. This camera-obtained information can be used to determine the correspondence between the camera and projector.²⁸ Generally speaking, the epipolar constraint, which allows the correspondence to be recognized in a specific region, is the most popular constraint.²⁹ Although the projected coded features should be found in a visible region, the reflected features sometimes interfere with the geometrical shapes and surface properties. As a result, the classification of some pixels or identification of coded features may fail. Therefore, several studies were proposed to reduce noise by colorimetric lights,³⁰ to improve the pixel classification³¹ and to remove interreflections of surface scattering.³²

Binocular vision is another common method for the structured-light systems consisting of two cameras and a projector.³³ For high-speed data acquisition, the two cameras are usually synchronized. In this configuration, the projector provides distinct coded features that are observed by both cameras. The corresponding features are then obtained under epipolar constraints. Therefore, the 3-D points can be readily obtained by the direct linear triangulation method under the known parameters of a calibrated stereo camera. The projector becomes a feature generator, and geometrical calibration of the projector is not always necessary.^{34,35} Based on the calibrated stereo camera, 3-D equations of the stripes from the projector can be further estimated for consistent feature matching.^{36,37} This configuration can

also be considered as two individual structured-light systems that are able to reduce the shadow region.³⁸ The use of two different focal length cameras will allow for the acquisition of dense, multiple spatial-resolution images.³⁹

3 Proposed Method

In a structured-light 3-D scanning system, the projector is usually used to generate binary stripe patterns for measuring 3-D objects. However, the performance of the projector will dominate the measurement accuracy. This study considered a robust method for overcoming potential problems caused by the projector optics mechanism, such as the small depth of focus, nonuniform brightness distribution, low resolution, and low modulation transfer function (MTF). Specifically, in our proposed system, the projector was used to generate both binary stripe patterns and a sinusoidal pattern for producing features and measuring 3-D objects, which reduced errors obtained with conventional projector optics.

3.1 Camera Model

The mathematical model of a physical camera is commonly treated as an ideal pinhole model. The mathematical formulation of a pinhole camera is the projective operation of

$$\mathbf{x} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{X}, \quad (1)$$

where \mathbf{X} denotes a point of 3-D scene in the world coordinate system, and \mathbf{x} indicates where \mathbf{X} is projected onto the image plane. The extrinsic parameter $[\mathbf{R}|\mathbf{t}]$ consists of a rotation and a translation matrix and describes where the camera is in the world coordinate system. Therefore, the 3-D points in camera coordinates can be obtained by applying the $[\mathbf{R}|\mathbf{t}]$ matrix on world coordinate \mathbf{X} . Finally, \mathbf{K} is the intrinsic parameter representing the characteristics of the camera such as the focal length, optical center, and aspect ratio. Therefore, \mathbf{x} on the two-dimensional (2-D) image can be obtained after applying \mathbf{K} to the 3-D points in camera coordinates. In practice, the behavior of a camera does not perfectly fit the ideal pinhole model due to nonlinear optical distortion. The relationship between the measured distorted point (u_d, v_d) and the ideal point (u, v) in an image is usually governed by a high-order polynomial function as follows:

$$\begin{aligned} u_d &= (1 + a_0r^2 + a_1r^4 + a_2r^6)u + s_0r^2 \\ &\quad + (p_0 + p_2r^2)(r^2 + 2u^2) \\ v_d &= (1 + a_0r^2 + a_1r^4 + a_2r^6)v + s_1r^2 \\ &\quad + (p_1 + p_3r^2)(r^2 + 2v^2) \\ r^2 &= u^2 + v^2. \end{aligned} \quad (2)$$

Here, (a_0, a_1, a_2) , (s_0, s_1) , and (p_0, p_1) represent the radial, prism, and tangential distortion coefficients, respectively.

3.2 Camera Calibration

The command configuration of the structured-light 3-D system is shown in Fig. 1. The camera and projector were rigidly fixed on a rig. In practice, we placed a flat checkerboard on a pan-tilt turntable to automatically collect the calibration

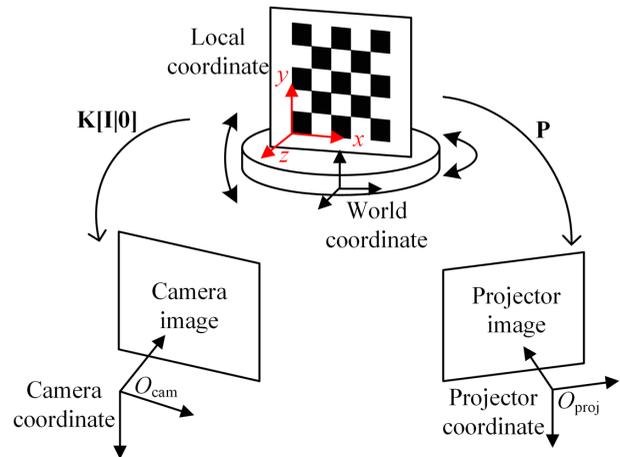


Fig. 1 Configuration of a structured-light 3-D system.

features. There were four coordinate systems used in this configuration, namely the camera, projector, world, and local coordinates.

The local coordinate represented the coordinate on each checkerboard, which can be determined by the corresponding extrinsic parameter of the camera. To estimate the relationship between the local coordinates and extrinsic camera parameters, we initially calibrated the camera to obtain the intrinsic, extrinsic, and distortion parameters of the camera based on Zhang's method.⁹ There were a total of 18×13 corner features generated in a frame. The calibration procedure typically requires a number of pose images to obtain the best calibration result. In our implementation, at least 30 regularly distributed pose images were used.

3.3 Projector Calibration

To complete the projector calibration, we utilized the centroid features of the projected image plane and estimated the centroid features on a real checkerboard. Based on this framework, the projector calibration procedure could be considered to be exactly the same as the camera calibration procedure. Since the extrinsic parameters from the calibration patterns were determined in the camera calibration, the transformation between each local coordinate and camera coordinate has been established. Nevertheless, the projection matrix P , which is a transformation from camera coordinates to projector coordinates, is still unknown. In general, the mathematical model of a projector is similar to that of a camera. However, to remove the distortion effects for the projector, estimation of the intrinsic parameters is necessary. In a structured-light 3-D system, we cannot directly measure the features on the projector coordinate. Therefore, we collected the centroid features of both the vertical and horizontal bright stripes from the highest-level patterns. In particular, we collected the projected centroid features observed by the camera using a second "camera calibration" to obtain the intrinsic, extrinsic, and distortion parameters of the projector. The projected centroid features were then converted into local coordinates using a homography transformation. Thus, the coordinate values of the projected centroid features on a real checkerboard were generated, establishing the correspondence between the projector image and the real checkerboard.

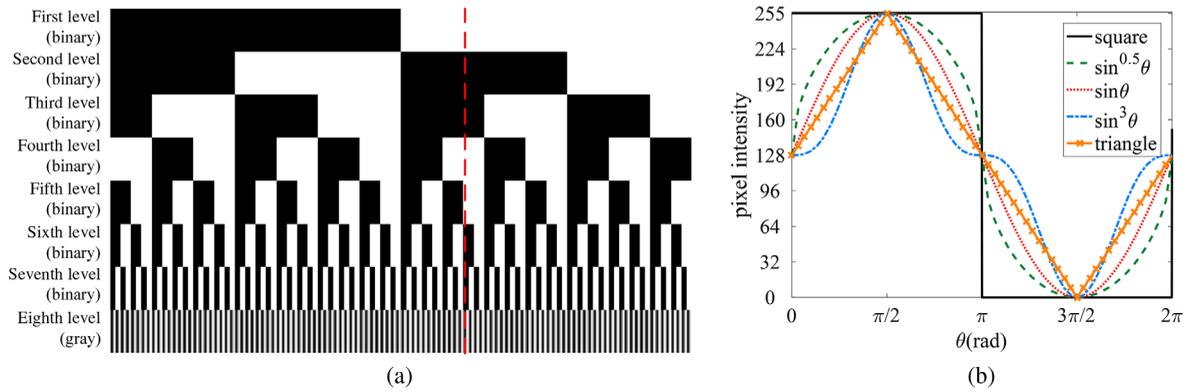


Fig. 2 (a) The eighth-level patterns and (b) different wave functions of the eighth-level pattern.

3.3.1 Centroid feature generation

The centroid features indicate the center of both the vertical and horizontal bright stripes of the highest-level pattern. However, the observed pixel intensities of the projected binary pattern will differ from those of the given binary pattern due to material reflectivity and the tone reproduction of both the camera and projector. Figure 2(a) shows an example of our given patterns. In practice, the eighth-level pattern was the highest level used for calibration. Moreover, to ensure the centroid features were distinct, the pixel intensity distribution of the eighth-level pattern was a critical component of our framework. Unlike the commonly used binary pattern, the sinusoidal pattern induced locally concentrated bright peaks. Since the centroid features are the local maximums of two overlaid eighth-level images, their positions could be determined by simply multiplying the two images. The projection of a given sinusoidal pattern on a physical surface has a similar effect to a Gaussian operation. Therefore, sinusoidal waves were preferred. Different types of waves applied to the eighth-level pattern are shown in Fig. 2(b). Higher order sinusoidal waves should generate more concentrated peaks.

3.3.2 Centroid feature on the local coordinate system

In our structured-light 3-D system, the stripe patterns were sequentially projected. The most critical pattern for the projector calibration was the highest-level pattern. We already have the given centroid features on the projector image from the camera calibration. The next problem was to determine their physical positions on the local coordinate system. The procedure was performed in two steps. First, all candidate centroid features were determined, and then, the centroid features were transformed into local coordinates. Figure 3 shows an example of centroid feature generation using eighth-level patterns. To obtain distinct centroid features, we multiplied the vertical and horizontal stripes of the eighth-level pattern and performed a Hadamard product operation. Figures 3(a) and 3(b) show the vertical and horizontal stripes, respectively, and Fig. 3(c) shows the result after multiplying the two images. We used multiplication for the two images to emphasize the local intensity of each centroid feature, which was helpful in determining distinct position at the subpixel level. In the calibration procedure, a printed checkerboard was used, and all candidate centroid features were determined using the white squares.

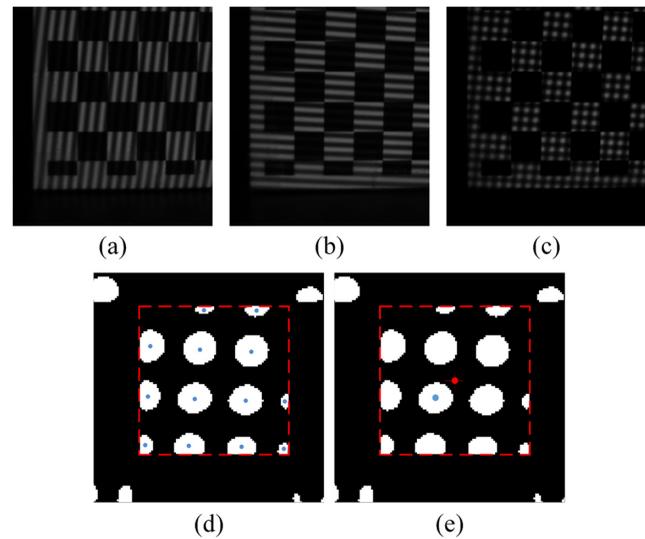


Fig. 3 Determination of centroid features on camera images: (a) vertical stripes (eighth-level), (b) horizontal stripes (eighth-level), (c) multiplied patterns (eighth-level), (d) all candidate centroid features on the white checker regions, and (e) the centroid feature closest to the center of the checkerboard is selected.

However, several candidate centroid features at the checker boundaries suffered from destroyed spot shapes, as shown in Fig. 3(d). To remove these outliers, only the centroid feature closest to the center of the white block was selected, as shown in Fig. 3(e).

To convert the image coordinate to the local checkerboard coordinate, one additional conversion was required. In the camera calibration, the checkerboard was rotated by the turntable to obtain multiple poses. Each pose represented an extrinsic parameter of the camera. Using Eq. (1), the 3-D point \mathbf{X} in checkerboard coordinates was converted into the 2-D point \mathbf{x} in image coordinates. However, to obtain \mathbf{X} from a known \mathbf{x} , the reverse conversion was required. Since our desired centroid features were given in local coordinates on the surface of checkerboard, the z -component of \mathbf{X} was zero. The $\mathbf{K}[\mathbf{R}|\mathbf{t}]$ matrix in Eq. (1) was reduced to a 3×3 matrix by removing the z -component (third column) of the $\mathbf{K}[\mathbf{R}|\mathbf{t}]$ matrix. Therefore, the conversion between the image coordinate and the surface of the checkerboard became a homography transformation. Thus, the centroid features were determined in local coordinates for every pose.

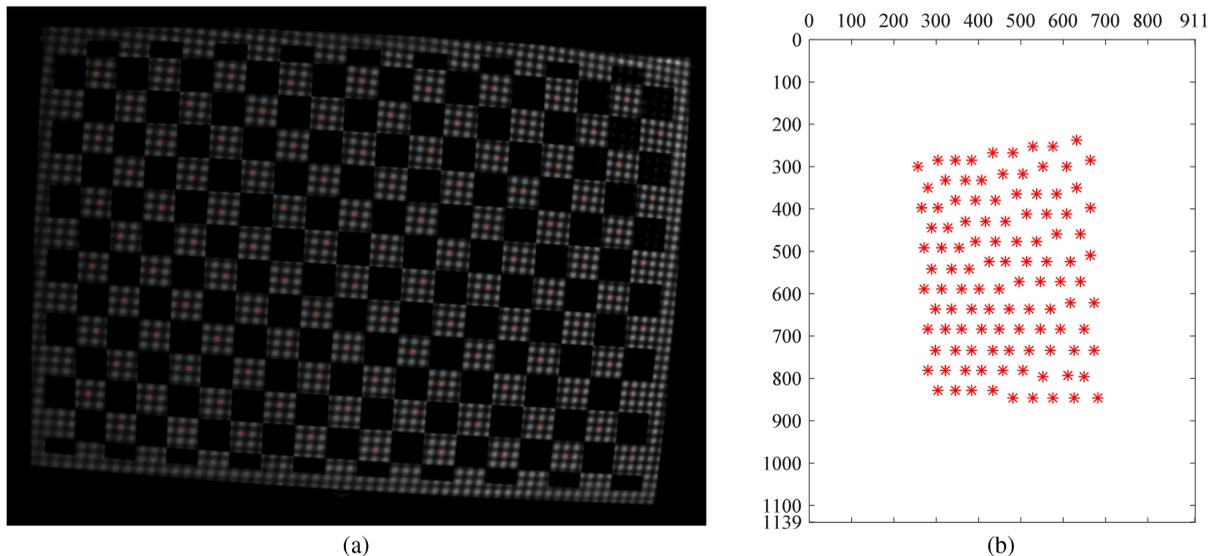


Fig. 4 The corresponding features on the projector image: (a) the centroid features in one camera image and (b) corresponding features on the projector image after the decoding procedure.

3.3.3 Corresponding projector image features to the centroid features

Though the centroid features were determined in image coordinates, their corresponding features on the projector image are still unknown. To determine those features, a decoding procedure for each centroid feature was required. The decode procedure involved finding the matching codes by comparing the observed camera image with the given projector image. Since the 2-D feature has two degrees of freedom, constraints for the decoding procedure in both directions were required. Therefore, vertical and horizontal stripe patterns were individually used for decoding the x - and y -components of the 2-D feature. Figure 4(a) shows the centroid features in one camera image, and Fig. 4(b) shows the determined corresponding features on the projector image.

3.3.4 Projector calibration

After obtaining the centroid features on the local coordinates and their corresponding features on the projector image, the projector calibration, which is exactly the same procedure as the camera calibration, was readily performed. The flowchart of the projector calibration is provided in Fig. 5. In practice, the turntable rotated the checkerboard to 30 different poses. In each pose, a total of 33 images were taken, including one white pattern, 16 vertical patterns, and 16 horizontal stripe patterns. The white pattern images in all poses were intentionally used for calibrating the camera because the calibrated camera was used to induce the extrinsic parameters for determining the position of each checkerboard. These 16 vertical stripe patterns, including eight-level pattern images and their inverse image, were used for robustly determining the transition boundary of the stripes due to the non-uniform brightness distribution of the acquired images. Due to the limitation of the projector resolution, the highest level used in this study was eight. There are two reasons for using the vertical and horizontal stripe patterns. The first reason for using the two stripe patterns is to generate the centroid features of the eighth-level images. The second reason is to

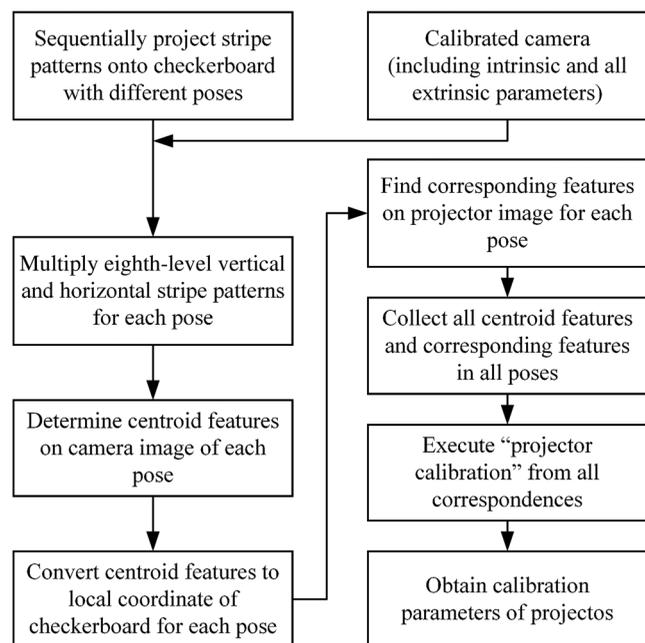


Fig. 5 The flowchart of the projector calibration procedure.

assist in finding the correspondence features on the projector image. After the projector calibration, the intrinsic and distortion parameters were obtained in order to remove the distortion effect of the projector. The extrinsic parameters of project were also used for recovering the relationship corresponding to the camera coordinate.

3.4 Three-Dimensional Reconstruction

The direct-triangulation method was implemented in the 3-D reconstruction procedure. Before determining the 3-D points, the corresponding points between the camera image and projector image must be determined. In practice, we projected the vertical stripe patterns and determined the centroid

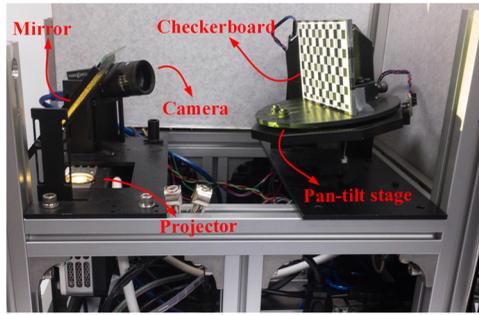


Fig. 6 The system configuration of our structured-light scanner.

lines in both the camera and projector images. All vertical centroid lines were determined at the eighth-level and encoded using hierarchical binary patterns. Thus, the line correspondences between the camera and projector images were obtained by comparing the encoding at all level patterns. For the horizontal constraints, we utilized a fundamental matrix to convert all features of the camera image into epipolar lines on the projector image. The fundamental matrix was obtained using the corresponding features from the camera and projector images. The intersections of the epipolar lines and centroid lines were collected as matching features. As a result, the 3-D coordinates of all corresponding features were determined using the direct-triangulation method.

4 Experimental Results

4.1 System Configuration

The mechanism of our structured-light system is shown in Fig. 6. A 3M-pixel camera manufactured by Pointgrey (FL3-U3-32S2M-CS) was used with an additional DLP (VIVITEK UMI-Q5), which has a native resolution of 912×1140 pixels. The inclined angle between the camera and projector was roughly 20 deg. To shorten the distance from the scanned objects, a mirror was inserted between the projector and the work zone. In the work zone, a pan-tilt stage driven by a step motor was used to automatically generate various poses of the checkerboard. In practice, the encoder for the motor was not required during calibration.

4.2 Comparison of the Cast Pattern in the Projector Calibration

To generate distinct centroid features, several patterns for the highest level are considered, as discussed in Sec. 3.3.1.

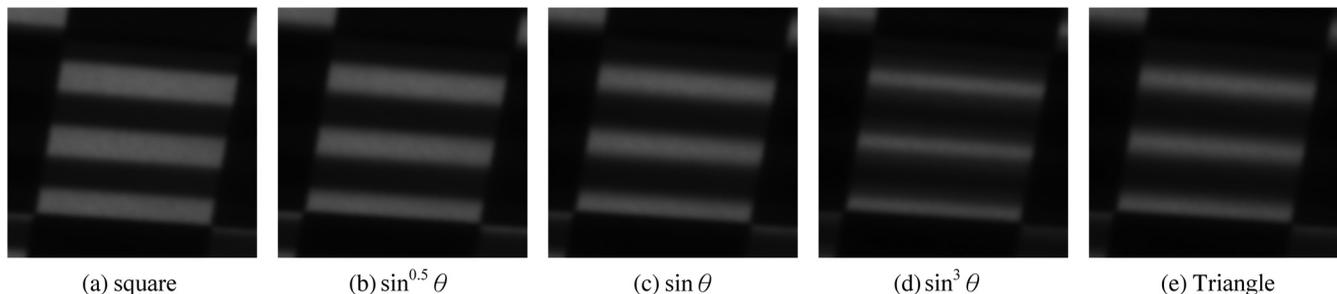


Fig. 7 Five different patterns cast on a white checker.

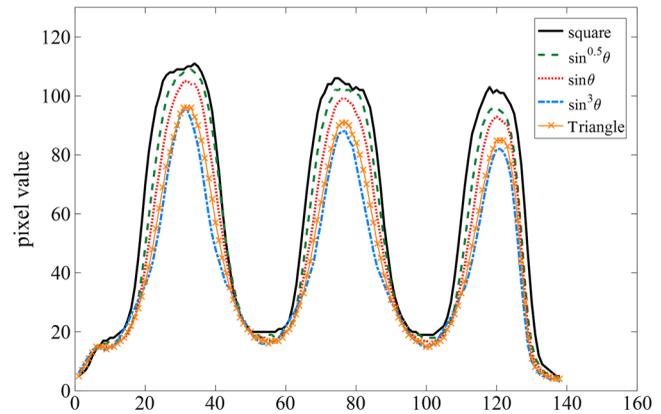


Fig. 8 Observed pixel intensity on a white checker.

The camera can be expected to observe more distinct features when projector casts a sharper pattern. Figure 7 shows the close up images of five different high-level patterns, i.e., level 8. The intensity distribution observed by the camera is known to be a combination of three factors: the real emission distribution of the projection, the material reflection property, and the light response of the camera. From the result shown in Fig. 8, the highest order pattern considered, i.e., $\sin^3 \theta$, had more distinct peaks, but induced a dimmer pattern.

To evaluate the calibration performance under different patterns, the camera calibration was initially performed based on Zhang's method.⁹ The projector was then calibrated based on the known extrinsic parameters of the camera. The reprojection errors of the projector in the case of several patterns are listed in Table 1. The experimental results showed that the higher order pattern had a smaller and more concentrated error. Based on the framework, the root mean square (RMS) error of the $\sin^3 \theta$ pattern was as small as 0.15 pixels.

4.3 Projector Calibration under Shallow Focus

The reason we obtained a stable result is that we do not rely on the highest-level pattern for decoding the correspondences. Since the resolution of the projector in our structured-light system is 912×1140 pixels, the projector was able to generate 2^{10} and 2^{11} vertical and horizontal patterns, respectively. However, for most commercial projects, the depth of focus (DoF) is usually short because a relative large aperture is used to increase the radiant flux. Therefore, blur patterns are expected in the work zone. In addition,

Table 1 Calibration error for different eighth-level patterns.

Projected pattern	Camera calibration error only (pixel) compared to the 3M-pixel image			Projector calibration error (pixel) compared to the 912 × 1140 pixel image		
	Average distance error	Deviation of distance error	RMS error	Average distance error	Deviation of distance error	RMS error
Square pattern	0.25	0.21	0.32	0.20	0.102	0.22
$\sin^{0.5} \theta$ pattern				0.15	0.082	0.17
$\sin \theta$ pattern				0.15	0.085	0.18
$\sin^3 \theta$ pattern				0.13	0.072	0.15
Triangle pattern				0.15	0.086	0.18

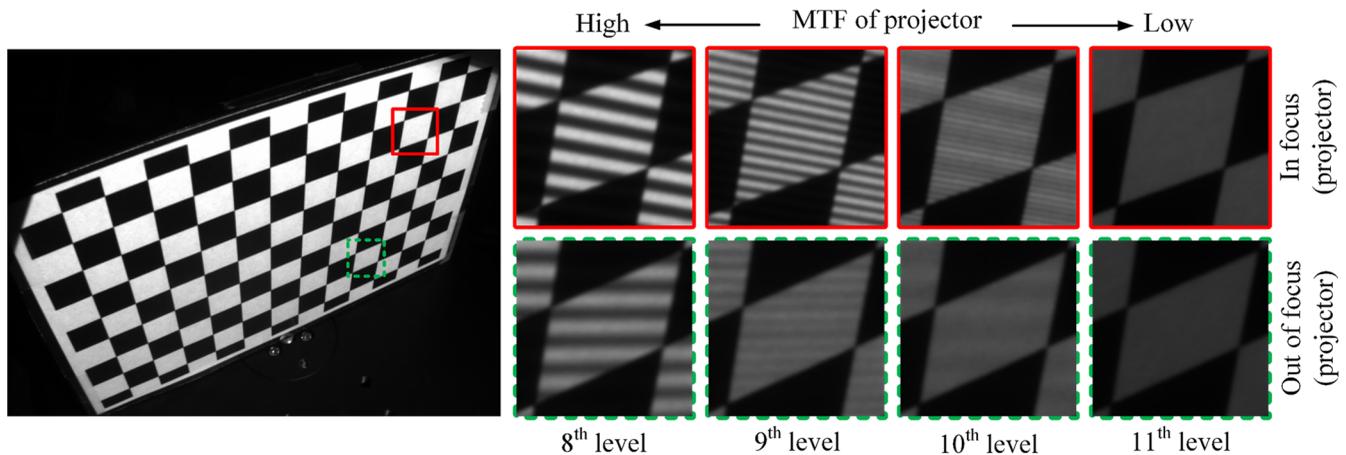


Fig. 9 Captured images with different levels of horizontal patterns. The first row figures indicate that projectors with a lower MTF will yield a blurred pattern. The second row figures show that, when the object is out of focus, the pattern becomes blurred as well.

low MTF in the projector is a critical factor in the generation of a blurred image. For a prime lens camera, the DoF can be extended by adjusting aperture size, and the camera does not usually suffer from low MTF. Therefore, the images in the work zone are unlikely to be blurred as a result of proper camera settings. In contrast, the DoF and MTF of the projector are occasionally limited. Figure 9 shows a calibration board at a wide depth. In this condition, the low MTF of the projector induces blurred patterns, even for in-focus regions, as shown in the top row in Fig. 9. Another factor inducing blurred patterns is a result of a shallow focus projector, as shown in the bottom row of Fig. 9. In general, a higher-level pattern should induce more accurate correspondences. However, the use of high-level patterns may suffer from uncertainty due to the physical resolution limitations of both the camera and projector. To avoid this uncertainty, the cubic sinusoidal pattern is considered instead of the binary pattern in the projector calibration. The experiment in Table 2 shows the calibration results in different three-level patterns. Unlike Moreno and Taubin,¹⁹ our proposed method was able to determine distinct features. Moreover, the calibration error in the projector was as small as 0.14 pixels. Our highest-level pattern was a sinusoidal wave pattern, and the remaining levels were binary patterns. Thus, the centroid features from the multiplication of vertical and

Table 2 Reprojection error (RMS) of the projector.

Highest level ($\sin^3 \theta$ pattern)	Taubin ¹⁹ (pixel)	Proposed method (pixel)
7	20.57	0.16
8	13.41	0.15
9	25.28	3.02

horizontal patterns will be distinct even if the projective patterns are out of focus.

4.4 Scan Benchmark

To evaluate the performance of the calibration beyond its reprojection errors, we scanned a qualified ceramic sphere having a diameter of 19.9824 mm (Mitutoyo CMM Masterball 06ABM944D). To suppress the reflection due to its glossy surface, we sprayed a very thin and uniform layer of paint on the sphere. In our system, the pan-tilt stage was installed in the center of the work zone only for automatically generating various poses of the calibration checkboard. Extreme corners of the work zone may suffer

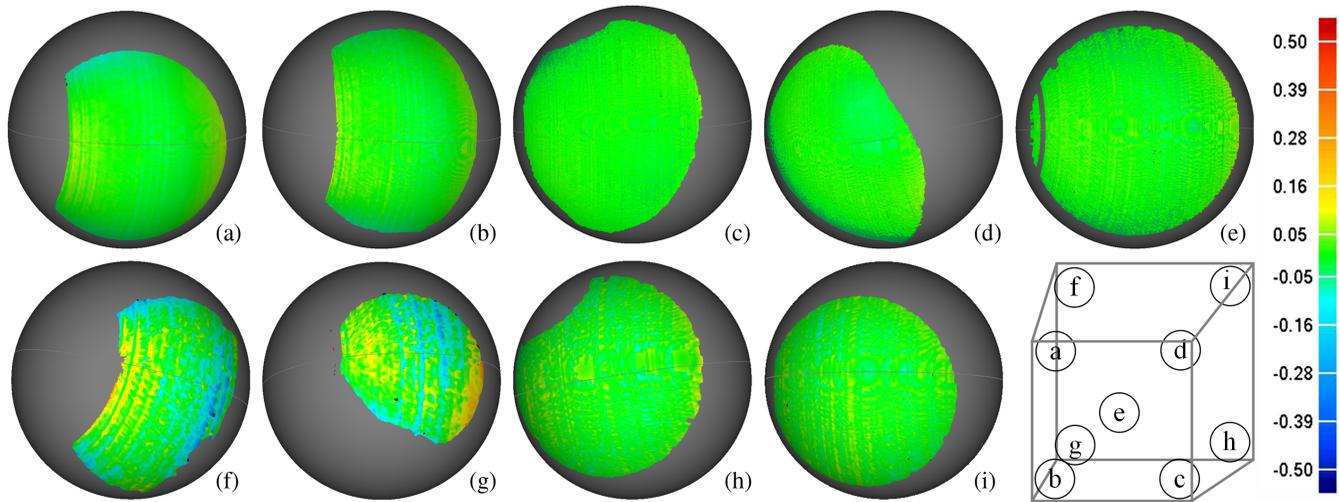


Fig. 10 Comparison of the error distribution for various scanned positions in the work zone.

from a lack of valid features. Nevertheless, the scan accuracy was still qualified. Figure 10 shows the error distributions of the scanned spheres and their fitting spheres at different positions of the work zone, including the center and eight extreme corners. In each position, the scanned 3-D points were collected to estimate their fitting spheres. The overall mean error and standard deviation were 46.6 and 52.6 μm , respectively. For the center of the work zone, the mean error

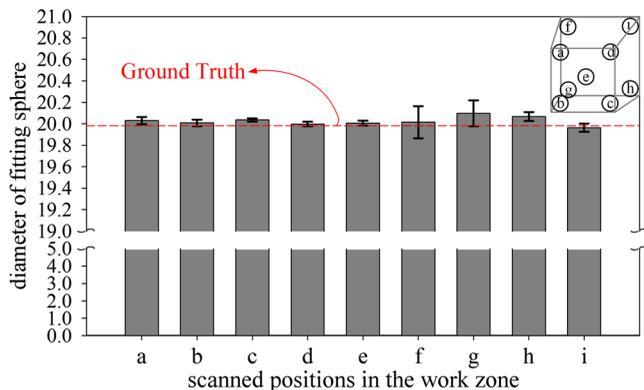


Fig. 11 The verification of the scanned spheres.

and standard deviation were as small as 23.9 and 23.6 μm , respectively. In Fig. 10, positions *f* and *g*, which are far from the projector, suffer from blurred images due to the out-of-focus patterns of the projector. As a result, their estimated dimension errors reached 115 μm , as shown in Fig. 11.

To compare the error distribution with previous work, we scanned a cubic block based on two calibration data in Table 2. The block has a width of 70 mm, and all scanned points on two perpendicular surfaces of the block are collected to visualize the overall error distribution, as shown in Fig. 12. As mentioned, the projector under shallow focus will induce poor scan result. Figures 12(a) and 12(b) represent our proposed method and Moreno and Taubin's,¹⁹ respectively. Comparing to Moreno and Taubin's method, our proposed method has wide depth of field. As a result, the overall error within the work zone is relatively small.

To verify the scanning of 3-D objects, we tested various materials, including gypsum, plastic with metal coding, rubber, and earthenware, as shown in Fig. 13. Some of the objects had rich colors. The gypsum sculpture surface has the property of Lambertian reflectance. Therefore, the apparent brightness of the reflective patterns was uniform regardless of observed angle of view. The observed features from the centroids of the high-level strips were convincing for

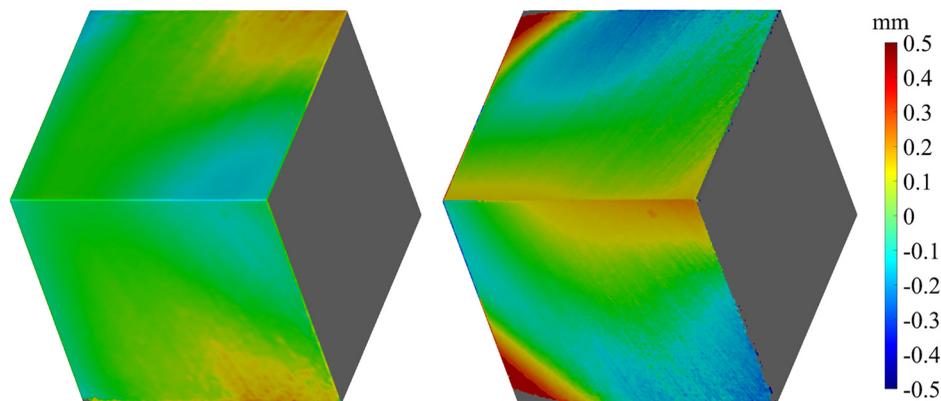


Fig. 12 Pair comparison of error distribution on the block.



Fig. 13 Scanning result for various objects, constructed of gypsum, plastic with metal coating, rubber, and earthenware, respectively.

determining the 3-D positions. For the plastic chicken with the metal-coated surface, the surface has severe and irregular reflectance. Consequently, the camera may encounter difficulty due to its lack of dynamic range. Therefore, fewer centroid features were therefore observed. Nevertheless, our proposed method adaptively adjusted the local threshold for each region. In the case of the colored rubber horse, its surface covers a large number of tiny concave structures. The observed brightness of the centroid feature in the image was affected by the lightness of the color on the horse figurine. If two different reflection rate colors are located at the bright period of a sinusoidal wave, the estimated position of the centroid feature will accommodate the combined intensities. This defect can be suppressed using a brightness

compensation mask by taking an additional image of a uniform white pattern. In the case of the earthenware object, the surface has dim, light, and glossy paints, as shown in the bottom row of Fig. 13. The transition between the different paints is gentle. To verify the reprojection error of our system, we further projected these 3-D points onto the camera and projector images, then measured the averaged distance to the corresponding features. The reprojection error is sometimes used to evaluate how confident the estimated 3-D position is. In Fig. 14, only one frontal range image of each 3-D object is retrieved for comparison. The reprojection errors of most of 3-D points are as small as 0.05 pixel. Our reconstruction successfully preserved the vivid 3-D features.

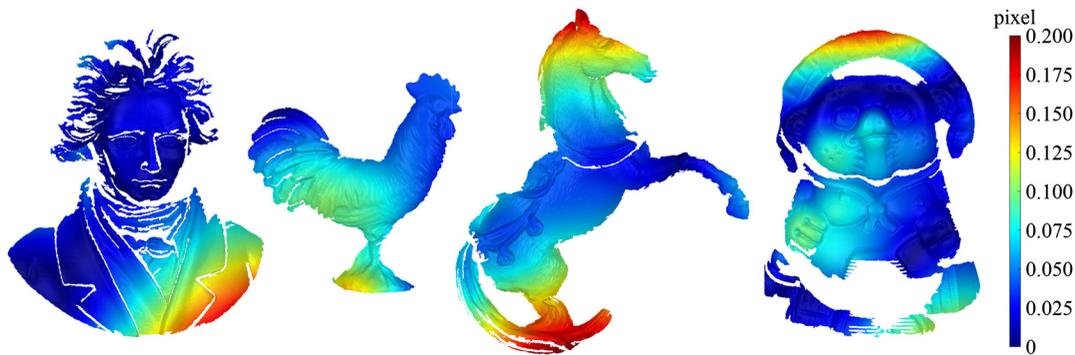


Fig. 14 Evaluation of reprojection error for various scanning result.

5 Conclusions

In this paper, a hybrid method primarily designed for the defocusing projector in a structured-light 3-D scanning system was proposed and implemented. The projected patterns consisted of the conventional sequential binary patterns and one additional sinusoidal pattern at the highest level. Because most commercial projectors usually suffer from shallow focus, our proposed method utilized the high-order sinusoidal pattern to enhance the corresponding features in the projector calibration. Calibration experimental results showed that the utilization of the $\sin^3 \theta$ pattern as the highest-level pattern significantly improved calibration error. Compared to the existing calibration method, the proposed method was shown to be more robust for the defocusing structured-light 3-D scanning system. The scan benchmark experiment demonstrated qualitative comparisons for scanning various 3-D objects. The results showed that the proposed method was capable of performing quality 3-D scans of various materials using a shallow focus projector.

Disclosures

There are no conflicts of interest to declare.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grant Nos. NSC102-2221-E-011-094-MY2 and MOST 103-2218-E-194-003.

References

1. F. Chen, G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," *Opt. Eng.* **39**(1), 10–22 (2000).
2. R. J. Valkenburg, A. M. McIvor, and P. W. Power, "An evaluation of sub-pixel feature localisation methods for precision measurement," *Proc. SPIE* **2350**, 229–238 (1994).
3. J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognit.* **37**(4), 827–849 (2004).
4. D. Lanman and G. Taubin, "Build your own 3D scanner: 3D photography for beginners," in *ACM SIGGRAPH* (2009).
5. S. S. Gorthi and P. Rastogi, "Fringe projection techniques: whither we are?" *Opt. Lasers Eng.* **48**(2), 133–140 (2010).
6. J. Geng, "Structured-light 3D surface imaging: a tutorial," *Adv. Opt. Photonics* **3**(2), 128–160 (2011).
7. R. J. Valkenburg and A. M. McIvor, "Accurate 3D measurement using a structured light system," *Image Vision Comput.* **16**, 99–110 (1998).
8. R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Rob. Autom.* **3**(4), 323–344 (1987).
9. Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *IEEE Int. Conf. on Computer Vision*, Vol. 1, pp. 666–673 (1999).
10. J. Heikkilä, "Geometric camera calibration using circular control points," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(10), 1066–1077 (2000).
11. D. Fofi, J. Salvi, and E. M. Mouaddib, "Uncalibrated reconstruction: an adaptation to structured light vision," *Pattern Recognit.* **36**(7), 1631–1644 (2003).
12. M. Kimura, M. Mochimaru, and T. Kanade, "Projector calibration using arbitrary planes and calibrated camera," in *IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 4, pp. 1–2 (2007).
13. I. Martynov, J. Kamarainen, and L. Lensu, "Projector calibration by inverse camera calibration," *Lect. Notes Comput. Sci.* **6688**, 536–544 (2011).
14. W. Gao, L. Wang, and Z. Hu, "Flexible method for structured light system calibration," *Opt. Eng.* **47**(8), 083602 (2008).
15. J. Ouellet, F. Rochette, and P. Hébert, "Geometric calibration of a structured light system using circular control points," in *Int. Symp. on 3D Data Processing, Visualization and Transmission*, pp. 183–190 (2008).
16. H. Anwar, I. Din, and K. Park, "Projector calibration for 3D scanning using virtual target images," *Int. J. Precis. Eng. Manuf.* **13**(1), 125–131 (2012).
17. R. Orghidan et al., "Structured light self-calibration with vanishing points," *Mach. Vision Appl.* **25**(2), 489–500 (2014).
18. J. Dráréni, S. Roy, and P. Sturm, "Methods for geometrical video projector calibration," *Mach. Vision Appl.* **23**(1), 79–89 (2012).
19. D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *Second Int. Conf. on 3D Imaging, Modeling, Processing, Visualization and Transmission*, pp. 464–471 (2012).
20. L. Huang, P. S. K. Chua, and A. Asundi, "Least-squares calibration method for fringe projection profilometry considering camera lens distortion," *Appl. Opt.* **49**(9), 1539–1548 (2010).
21. S. Ma et al., "Flexible structured-light-based three-dimensional profile reconstruction method considering lens projection-imaging distortion," *Appl. Opt.* **51**(13), 2419–2428 (2012).
22. K. Li, J. Bu, and D. Zhang, "Lens distortion elimination for improving measurement accuracy of fringe projection profilometry," *Opt. Lasers Eng.* **85**, 53–64 (2016).
23. H. Luo et al., "A simple calibration procedure for structured light system," *Opt. Lasers Eng.* **57**, 6–12 (2014).
24. B. Li, N. Karpinsky, and S. Zhang, "Novel calibration method for structured-light system with an out-of-focus projector," *Appl. Opt.* **53**(16), 3415–3426 (2014).
25. P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992).
26. S. Barone, A. Paoli, and A. V. Razonale, "3D reconstruction and restoration monitoring of sculptural artworks by a multi-sensor framework," *Sensors* **12**(12), 16785–16801 (2012).
27. X. Pang et al., "A tool-free calibration method for turntable-based 3D scanning systems," *IEEE Comput. Graphics Appl.* **36**(1), 52–61 (2016).
28. J. Batlle, E. Mouaddib, and J. Salvi, "Recent progress in coded structured light as a technique to solve the correspondence problem: a survey," *Pattern Recognit.* **31**(7), 963–982 (1998).
29. M. Maruyama and S. Abe, "Range sensing by projecting multiple slits with random cuts," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(6), 647–651 (1993).
30. D. Caspi, N. Kiryati, and J. Shamir, "Range imaging with adaptive color structured light," *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(5), 470–480 (1998).
31. Y. Xu and D. G. Aliaga, "Robust pixel classification for 3D modeling with structured light," in *Int. Conf. on Graphics Interface*, pp. 233–240 (2007).
32. M. Gupta et al., "A practical approach to 3D scanning in the presence of interreflections, subsurface scattering and defocus," *Int. J. Comput. Vision* **102**(1–3), 33–55 (2013).
33. Y. B. Mahdy, K. F. Hussain, and M. A. Abdel-Majid, "Projector calibration using passive stereo and triangulation," *Int. J. Future Comput. Commun.* **2**(5), 385–390 (2013).

34. C. Wen, K.-C. Chang, and M.-F. Yeh, "A low-cost, user-friendly, and real-time operating 3D camera," *IEEE Trans. Consum. Electron.* **51**(4), 1085–1091 (2005).
35. J. Sun et al., "Large 3D free surface measurement using a mobile coded light-based stereo vision system," *Sens. Actuators A* **132**(2), 460–471 (2006).
36. R. Yang et al., "Robust and accurate surface measurement using structured light," *IEEE Trans. Instrum. Meas.* **57**(6), 1275–1280 (2008).
37. R. Yang, S. Cheng, and Y. Chen, "Flexible and accurate implementation of a binocular structured light system," *Opt. Lasers Eng.* **46**(5), 373–379 (2008).
38. J. Lin, K. Jiang, and M. Chang, "A novel solution for camera occlusion in stereo vision technique," *Adv. Mech. Eng.* **5**, 253794 (2015).
39. S. Dong et al., "Calibration method for a structured light measurement system with two different focal length cameras," *Measurement* **73**, 462–472 (2015).

Yu-Lun Liu received his BS degree in aerospace engineering from Tamkang University in 2012, and his MS degree in mechanical engineering from the National Taiwan University of Science and Technology (NTUST), in 2014. Currently, he is pursuing his PhD in Graduate Institute of Applied Science and Technology at NTUST. His current research interests include computer vision and computer graphics.

Tzung-Han Lin received his PhD from the Department of Mechanical Engineering at National Taiwan University in 2006. He joined Industrial Technology and Research Institute as a senior engineer in 2007. He is currently an associate professor of the Graduate Institute of Color and Illumination Technology at NTUST. His research interests include 3-D data acquisition, 3-D reconstruction, computer vision, and computer graphics-related topics.