Novel calibration method for structured-light system with an out-of-focus projector

Beiwen Li, Nikolaus Karpinsky, and Song Zhang*

Department of Mechanical Engineering, Iowa State University, Ames, Iowa 50011, USA

*Corresponding author: song@iastate.edu

Received 13 February 2014; revised 13 April 2014; accepted 14 April 2014;
posted 15 April 2014 (Doc. ID 206416); published 23 May 2014

A structured-light system with a binary defocusing technique has the potential to have more extensive application due to its high speeds, gamma-calibration-free nature, and lack of rigid synchronization requirements between the camera and projector. However, the existing calibration methods fail to achieve high accuracy for a structured-light system with an out-of-focus projector. This paper proposes a method that can accurately calibrate a structured-light system even when the projector is not in focus, making it possible for high-accuracy and high-speed measurement with the binary defocusing method. Experiments demonstrate that our calibration approach performs consistently under different defocusing degrees, and a root-mean-square error of about 73 μm can be achieved with a calibration volume of 150(H) mm × 250(W) mm × 200(D)mm. © 2014 Optical Society of America

OCIS codes: (120.0120) Instrumentation, measurement, and metrology; (120.2650) Fringe analysis; (100.5070) Phase retrieval.

http://dx.doi.org/10.1364/AO.53.003415

1. Introduction

Three-dimensional (3D) shape measurement is an extensively studied field that enjoys wide applications in, for example, biomedical science, entertainment, and the manufacturing industry [1]. Researchers have been making great efforts in achieving 3D shape measurements with higher speed, higher resolution, and wider range. One of the crucial elements is to accurately calibrate each device (e.g., camera, projector) used in such a system.

The calibration of the camera has been quite extensively studied over a long period of time. The camera calibration was first performed with 3D calibration targets [2,3] that required high-precision manufacturing and higher accuracy measurements of the calibration targets, which is usually not easy to obtain. To simplify the calibration process, Tsai [4] has proved that two-dimensional (2D) calibration targets with rigid out-of-plane shifts are sufficient to achieve high-accuracy calibration without requiring complex 3D calibration targets. Zhang [5] proposed a flexible camera calibration method that further simplified the calibration process by allowing the use of a flat 2D target with arbitrary poses and orientations, albeit still requiring knowledge of the target geometry and the preselection of the corner points. Some recent advances of calibration techniques further improved the flexibility and the accuracy of calibration by using a not-measured or imperfect calibration target [6–9], or by using active targets [10,11].

The structured-light system calibration is more complicated since it involves the use of a projector. Over the years, researchers have developed a variety of approaches to calibrate the structured-light system. Attempts were first made to calibrate the system by obtaining the exact system parameters (position, orientation) of both devices (camera, projector) [12–14]. Then, to save the effort of the complex system setup required by those methods, some other methods [15–18] improved the flexibility by establishing equations that estimate the relationship between the depth and the phase value. Another

1559-128X/14/163415-12$15.00/0

popular calibration approach was to treat the projector as a device with the inverse optics of a camera, such as the Levenberg–Marquardt method [19], and thus the projector calibration can be as simple as a camera calibration. The enabling technology was developed by Zhang and Huang [20], which enabled the projector to "capture" images like a camera through projecting a sequence of fringe patterns to establish one-to-one mapping between the projector and the camera. Following Zhang and Huang's work, researchers have tried to improve the calibration accuracy by linear interpolation [21], bundle adjustment [22], or residual error compensation with planar constraints [23]. All the aforementioned techniques have proven to be successful in calibrating the structured-light system, but they all require the projector to be at least nearly focused. Therefore, they cannot be directly applied to calibrate the structured-light system with an out-of-focus projector.

Our recent efforts have been focusing on advancing the binary defocusing technology [24] because it has the merits of high speed [25], being gamma calibration free, and having no rigid requirement for precise synchronization. However, as aforementioned, none of the existing calibration methods can be directly applied to accurately calibrating our structured-light system in which the projector is substantially defocused. One attempt to calibrate the structured-light system with an out-of-focus projector was carried out by Merner *et al.* [26]. The method proposed by Merner *et al.* was able to achieve high-depth accuracy (± 50 μm), but the spatial (along x or y) accuracy was limited (i.e., a few millimeters). For measurement conditions only requiring high-depth accuracy, that method is good. However, for generic 3D shape measurement, x and y calibration accuracy is equally important.

This paper proposes to accurately calibrate the structured-light system with an out-of-focus projector. With the projector being out of focus, no one-to-one mapping between the camera pixel and the projector pixel can be established as in the prior study [20]. This paper will present the idea of virtually creating the one-to-one mapping between the camera pixel and the center point of the projector pixel in the phase domain. Meanwhile, by coinciding the world coordinate system with the camera lens coordinate system, we can employ a standard stereo system calibration method to accurately calibrate the structured-light system even when the projector is substantially out of focus. Our experimental results demonstrate that the novel calibration approach that we will present in this paper performs consistently over different amounts of defocusing, and can reach about 73 μm accuracy for a calibration volume of 150(H) mm × 250(W) mm × 200(D) mm.

Section 2 illustrates related basic principles of the proposed method including the phase-shifting method used, and the modeling of the out-of-focused projector. Section 3 presents the calibration procedures used and the calibration results obtained. Section 4 shows experimental results to verify the performance of the proposed calibration method, and Section 5 summarizes this paper.

2. Principles

A. N-Step Phase-Shifting Algorithm

Phase-shifting algorithms have gained great popularity in the area of optical metrology owing to its speed and accuracy. There are a variety of phase-shifting algorithms that have demonstrated their success in measurement including three-step, four-step, and five-step. In general, the more steps used, the better the accuracy that can be achieved. For an N-step phase-shifting algorithm, the kth projected fringe image can be mathematically represented as follows:

$$I_k(x,y) = I'(x,y) + I''(x,y)\cos(\phi(x,y) + 2k\pi/N), \quad (1)$$

where $I'(x,y)$ represents the average intensity, $I''(x,y)$ indicates the intensity modulation, and $\phi(x,y)$ denotes the phase to be solved for. The phase can be computed by

$$\phi(x,y) = \tan^{-1}\left[\frac{\sum_{k=1}^{N} I_k \sin(2k\pi/N)}{\sum_{k=1}^{N} I_k \cos(2k\pi/N)}\right]. \quad (2)$$

The nature of the arctangent function produces the wrapped phase $\phi(x,y)$ with a range from $-\pi$ to π; then a temporal or spatial phase unwrapping algorithm is needed to obtain a continuous phase map. The conventional spatial phase unwrapping algorithms only recover relative phase. In this research, the absolute phase is needed to establish the mapping between the camera pixel coordinates and projector pixel center coordinates, which will be introduced in Section 2.D. Here, we used nine-step phase-shifted fringe patterns ($N = 9$) for the narrowest fringe patterns (fringe period of 18 pixels), and two additional sets of three-step phase-shifted fringe patterns for wider fringe patterns (fringe pitch of 21 and 154 pixels). We then adopted the three-frequency temporal phase unwrapping algorithm introduced in [27] for absolute phase retrieval.

B. Camera Model

In our structured-light system, we used the standard pinhole camera model as shown in Fig. 1, where $(o^w; x^w, y^w, z^w)$ represents the world coordinate system, $(o^c; x^c, y^c, z^c)$ denotes the camera lens coordinate system, and $(o_0^c; u^c, v^c)$ is the camera pixel coordinate system. The relationship between a point on the object and its corresponding image pixel can be described as

$$s^c\mathbf{I}^c = \mathbf{A}^c[\mathbf{R}^c, \mathbf{t}^c]\mathbf{X}^w. \quad (3)$$

Here $\mathbf{I}^c = (u^c, v^c, 1)^T$ is an image point in the homogeneous image coordinate system, $\mathbf{X}^w = (x^w, y^w, z^w, 1)^T$ denotes the corresponding point on

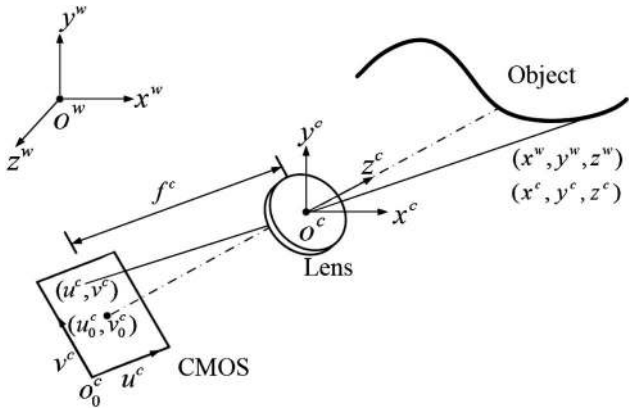


Fig. 1. Pinhole model of camera.

the object in the homogeneous world coordinate system, and s^c is the scale factor. The matrix $[\mathbf{R}^c, \mathbf{t}^c]$ is composed of extrinsic parameters, where $\mathbf{R}^c$ is a 3×3 matrix representing the rotation between the world coordinate system and the camera coordinate system, whereas $\mathbf{t}^c$ is a 3×1 vector representing the translation between those two coordinate systems. $\mathbf{A}^c$ is the matrix of the intrinsic parameters described by

$$\mathbf{A}^c = \begin{bmatrix} \alpha & \gamma & u_0^c \\ 0 & \beta & v_0^c \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

where (u_0^c, v_0^c) is the coordinate of the principle point, α and β are elements implying the focal lengths along the u^c and v^c axes, respectively, on the image plane, and γ is the skew factor of the two image axes.

In reality, the camera lens can have distortion; the nonlinear lens distortion is mainly composed of radial and tangential distortion coefficients, which can be modeled as a vector of five elements:

$$\mathbf{Dist}^c = \begin{bmatrix} k_1 & k_2 & p_1 & p_2 & k_3 \end{bmatrix}^T, \tag{5}$$

where k_1, k_2, and k_3 are the radial distortion coefficients, which can be corrected using the following formula:

$$u^{c'} = u^c(1 + k_1 r^2 + k_2 r^4 + k_3 r^6), \tag{6}$$

$$v^{c'} = v^c(1 + k_1 r^2 + k_2 r^4 + k_3 r^6). \tag{7}$$

Here, (u, v) and $(u^{c'}, v^{c'})$ refer to the camera point coordinates before and after correction, respectively, and $r = \sqrt{u^2 + v^2}$ represents the absolute distance between the camera point and the origin. Similarly, tangential distortion can be corrected using the following formula:

$$u^{c'} = u^c + [2p_1 uv + p_2(r^2 + 2u^2)], \tag{8}$$

$$v^{c'} = v^c + [p_1(r^2 + 2v^2) + 2p_2 uv]. \tag{9}$$

C. Camera Calibration

Essentially, the camera calibration procedure is to estimate the intrinsic and the extrinsic matrices so that the relationship between the world coordinate system and the image coordinate system is determined. The determination of the world coordinate system and the extrinsic matrix will be introduced in Section 2.E. The estimation of intrinsic parameters follows the model described by Zhang [5]; these parameters were estimated using the standard OpenCV camera calibration toolbox. Here, instead of using traditional checkerboard for calibration, we used 7×21 arrays of white circles printed on a flat black board as shown in Fig. 2, and the centers of the circles were extracted as feature points. The calibration board was positioned with different poses, and a total of 18 poses were captured to estimate the intrinsic parameters of the camera. It is worthwhile to note that the nonlinear calibration model was considered and the distortion was corrected for the camera calibration.

D. Out-of-Focus Projector Calibration

Basically, the projector has the inverse optics with respect to the camera since it projects images rather than capturing them. In order to enable the projector to have a similar calibration procedure to the camera, we need to create "captured images" for the projector by establishing the correspondence between the camera coordinate and the projector coordinate. However, in our system, we took advantage of binary defocusing technology. In this case, a defocused projector will make the calibration procedure quite challenging. This is because the model for calibrating the projector briefly follows the model for calibrating the camera, since there does not exist a technology that could calibrate a defocused camera, not to mention the calibration for an out-of-focus projector. In this section, we will introduce the model of a defocused imaging system and the solution to the calibration of an out-of-focus projector.

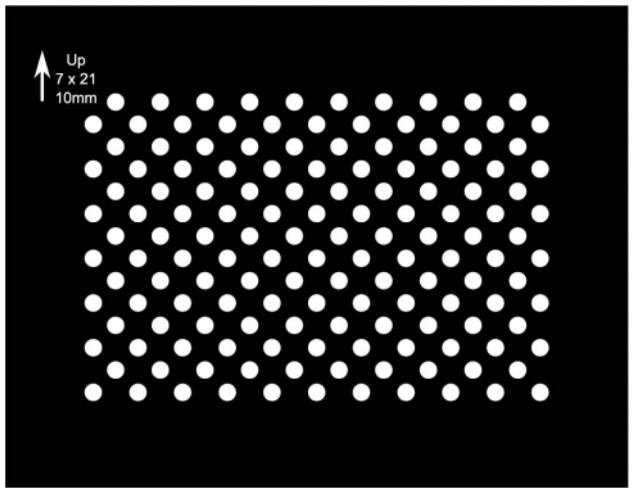


Fig. 2. Design of calibration board.

1. Model of Defocused Imaging System

The model of an imaging system in general can be described as follows. According to Ellenberger [28], suppose that $o(x,y)$ is the intensity distribution of a known object; its image $i(x,y)$ after passing through an imaging system can be described by a convolution of the object intensity $o(x,y)$ and the point spread function (PSF) $\mathrm{psf}(x,y)$ of the imaging system:

$$i(x,y) = o(x,y) \otimes \mathrm{psf}(x,y). \tag{10}$$

Here, $\mathrm{psf}(x,y)$ is determined by a pupil function of the optical system $f(u,v)$

$$\begin{aligned}\mathrm{psf}(x,y) &= \left|\frac{1}{2\pi}\iint_{-\infty}^{+\infty} f(u,v)e^{i(xu+yv)}\mathrm{d}u\mathrm{d}v\right|^2 \\ &= |F(x,y)|^2,\end{aligned} \tag{11}$$

where $F(x,y)$ denotes the Fourier transform of the pupil function $f(u,v)$. In general, the pupil function $f(u,v)$ is described as

$$f(u,v) = \begin{cases} t(u,v)e^{j\frac{2\pi}{\lambda}\omega(u,v)}, & \text{for } u^2+v^2 \le 1 \\ 0, & \text{for } u^2+v^2 > 1 \end{cases}, \tag{12}$$

where $t(u,v)$ represents the transmittance of the pupil, and $\omega(u,v)$ describes all source of aberrations. When describing the system in the Fourier domain by applying the convolution theorem, we can obtain

$$I(s_0,t_0) = O(s_0,t_0) \times \mathrm{OTF}'(s_0,t_0). \tag{13}$$

Here, $I(s_0,t_0)$ and $O(s_0,t_0)$ represent the Fourier transform of the function denoted by their corresponding lowercase letters, and $\mathrm{OTF}'(s_0,t_0)$ is the Fourier transform of the PSF. The optical transfer function (OTF) $OTF(s_0,t_0)$ is defined by its normalized form

$$\mathrm{OTF}(s_0,t_0) = \frac{\mathrm{OTF}'(s_0,t_0)}{\mathrm{OTF}'(0,0)}. \tag{14}$$

Specifically, if the system is circular symmetric and aberration free with the only defect of defocusing, the pupil transfer function can be simplified as

$$f(u,v) = \begin{cases} e^{j\frac{2\pi}{\lambda}\omega(u^2+v^2)}, & \text{for } u^2+v^2 \le 1 \\ 0, & \text{for } u^2+v^2 > 1 \end{cases}, \tag{15}$$

where ω is a circular symmetric function that describes the amount of defocusing, which can also be represented by the maximal optical distance between the emergent wavefront S and the reference sphere S_r, as shown in Fig. 3. Meanwhile, the OTF degenerates to [29]

$$\mathrm{OTF}(s) = \frac{\iint_{-\infty}^{+\infty} f\left(u+\frac{s}{2},v\right)f^*\left(u-\frac{s}{2},v\right)\mathrm{d}u\mathrm{d}v}{\int_{-\infty}^{+\infty}|f(u,v)|^2\mathrm{d}u\mathrm{d}v}, \tag{16}$$

with s being $\sqrt{(s_0^2+t_0^2)}$. The expression of the exact OTF can be very complicated and almost impossible to compute efficiently. However, according to Hopkins [30], if we neglect the diffraction properties of the light and approximate the OTF based on geometric optics, the OTF can be simplified as

$$\mathrm{OTF}(s) = 2J_1(a)/a, \qquad a = \left(\frac{4\pi}{\lambda}\omega s\right), \tag{17}$$

where J_1 is the Bessel function of the first kind with order $n = 1$. A visualization of the simplified OTF is shown in Fig. 4. Figure 4(a) shows an example of the OTF with defocusing parameter $\omega = 2\lambda$, and Fig. 4(b) shows a cross section of OTFs with different defocusing parameter ω, from which we can see that when there is no defect of defocusing ($\omega = 0$), the OTF has a uniform amplitude. However, when there exists a defect of defocusing, the OTF follows an "Airy rings"-like profile with the cut-off frequency being decreased when the defocusing degree ω increases.

A more intuitive understanding on how the defect of defocusing influences the resultant image can be obtained by looking at the PSF, as is illustrated in Fig. 5, which is the inverse Fourier transform of its corresponding OTF. Figure 5(a) shows an example of the normalized PSF with the defocusing parameter $\omega = 2\lambda$, while Fig. 5(b) illustrates a cross section of the normalized PSFs with different defocusing parameters. The normalized PSF indicates when the optical system is in focus ($\omega = 0$), the PSF becomes a unit impulse function centered at the origin, which means that a point on the object will still map to a point on the image after passing through the optical system, since the resultant image is simply a convolution between the object intensity distribution and the PSF. However, with the optical system becoming more and more defocused, the PSF expands to be a blurred circular disk, which means that a point on the object will no longer map to a single pixel on the image plane, but rather spread to the nearby region.

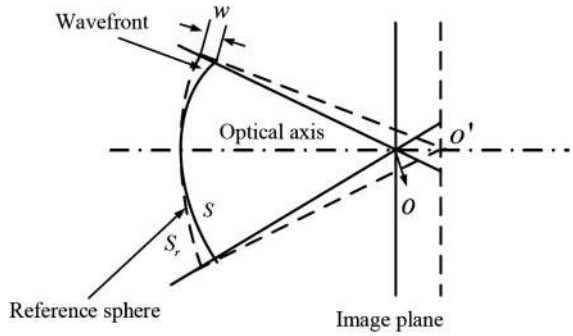


Fig. 3. Defocused optical system with the parameter ω.

2. Phase-Domain Invariant Mapping

Section 2.D.1 showed that if the optical system is defocused, a point on the object will no longer converge to a point on the image plane, but rather a blurred circular disk. For a structured-light system with an out-of-focus projector, as is illustrated in Fig. 6,

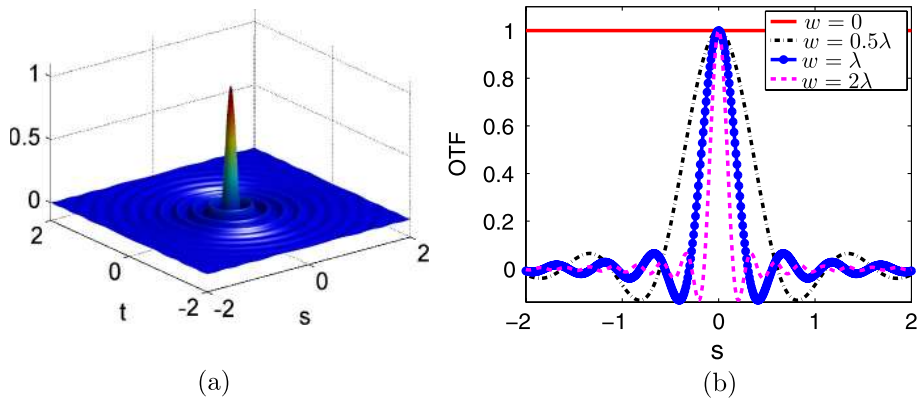


Fig. 4. Illustration of the optical transfer function (OTF) for a defocusing system. (a) Example of OTF with defocusing parameter $\omega = 2\lambda$. (b) Cross section of OTFs with different defocusing parameter ω.

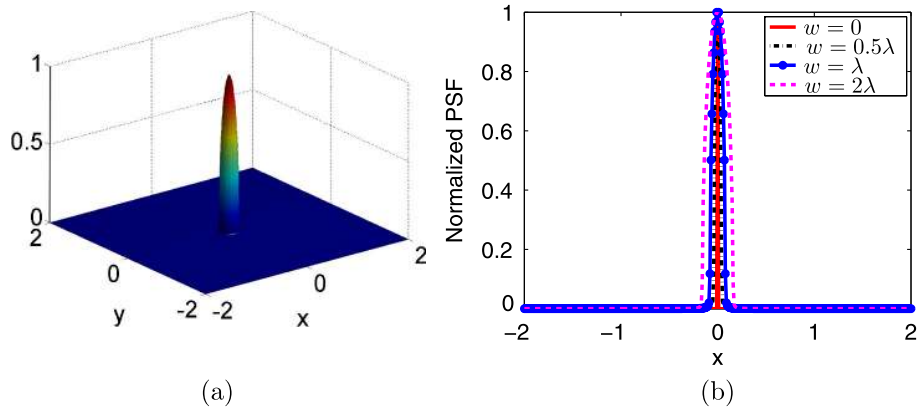


Fig. 5. Illustration of the point spread function (PSF) for a defocusing system. (a) Example of normalized PSF with defocusing parameter $\omega = 2\lambda$. (b) Cross section of normalized PSFs with different defocusing parameter ω.

a projector's pixel does not correspond to the one single pixel on the camera, but rather "pollutes" to its nearby region, as shown in the dashed area A. However, considering the infinite light ray of the optical system, the center of the projector pixel still corresponds to the center of a camera pixel regardless the amount of defocusing if they indeed are corresponding to each other. Therefore, if the center of the pixel can be found, one-to-one mapping between the projector and the camera can still be virtually established. From our previous discussion, the center point corresponds to the peak value of the circular disk whose phase value maintains regardless of the defocusing. Therefore, one-to-one mapping between the projector pixel center, (u^p, v^p), which is actually the pixel itself, and the camera pixel center, (u^c, v^c), can be established in the phase domain using the phase-shifting algorithm, albeit being impractical to generate the mapped projector images as proposed in [20].

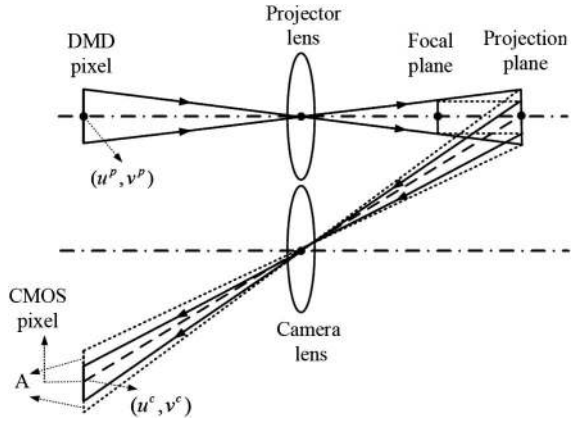


Fig. 6. Model of a structured-light system with an out-of-focus projector.

Theoretically, the mapping in the phase domain is invariant between the central points of a projector pixel and a camera pixel, which can be seen from the system model in the frequency domain. Based on the aforementioned model of the imaging system, as shown in Eq. (13), the Fourier transform $I_p(u^p, v^p)$ of the projector image $i_p(u, v)$ at the pixel center (u^p, v^p) can be represented by

$$I_p(u^p, v^p) = I_c(u^c, v^c) \times \frac{\mathrm{OTF}'_p(0,0)}{\mathrm{OTF}'_c(0,0)}, \tag{18}$$

where $I_c(u^c, v^c)$ is the Fourier transform of the corresponding camera image $i_c(u, v)$ at the pixel center

(u^c, v^c), $\mathrm{OTF}'_p(0,0)$ is the unnormalized OTF of the projector optical system at the center pixel, and $\mathrm{OTF}'_c(0,0)$ is the unnormalized OTF of the camera optical system at the center pixel. From Eq. (17), it is indicated that the OTF is a circular symmetric and real-valued function that does not have contribution to phase information. In other words, the phase of a point (u^p, v^p) on the camera image is not altered after passing through the two optical systems, and has the same value as the phase of the point (u^c, v^c) on the camera censor. Therefore, we can indeed establish one-to-one correspondence between the central points of a camera pixel and a projector pixel using the phase information.

The basic principle of the mapping can be described as follows. Without loss of generality, if the horizontal patterns are projected onto the calibration board and the absolute phase ϕ_{va} in the vertical gradient direction is retrieved, the camera coordinate can be mapped to the projector horizontal line using the constraint of equal phase values

$$\phi^c_{va}(u^c, v^c) = \phi^p_{va}(v^p) = \phi_{va}. \tag{19}$$

Similarly, if vertical patterns are projected and the absolute phase ϕ_{ha} in the horizontal gradient direction is extracted, another constraint can be established as

$$\phi^c_{ha}(u^c, v^c) = \phi^p_{ha}(u^p) = \phi_{ha}, \tag{20}$$

to correspond the one camera pixel to the vertical line of the projector image plane. The intersecting point of these two lines on the projector image plane (u^p, v^p) is the unique mapping point of the camera pixel (u^c, v^c) in (and only in) the phase domain. The intrinsic matrix of the projector can be estimated by using the same 18 different poses of the camera calibration board using the mapped points of the circle centers for each pose.

E. System Calibration

After estimating the intrinsic parameters for the camera and the projector, the system extrinsic parameters can be calibrated using the standard stereo camera calibration method.

Our previous method that used the world coordinate system was established based on one calibration image [20]. However, this method is far from optimal since rotation and translation parameters estimated by one calibration pose can only provide limited accuracy. Therefore, in this research, the world coordinate coincides with the camera lens coordinate system. Since the camera is unique, we can estimate the transformation matrix $[\mathbf{R}, \mathbf{t}]$ from the camera to the projector with a number of different poses, which can essentially improve the accuracy of calibration.

Figure 7 shows the diagram of our structured-light system. Here, $(o^c; x^c, y^c, z^c)$ and $(o^c_0; u^c, v^c)$ represent the camera coordinate system and its image coordinate system. $(o^p; x^p, y^p, z^p)$ and $(o^p_0; u^p, v^p)$ denote the

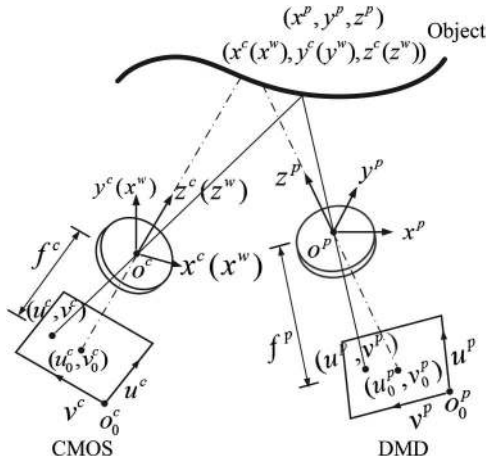


Fig. 7. Pinhole model of structured-light system.

projector coordinate system and its image coordinate system. f^c and f^p represent the focal length of the camera and the projector, respectively. And $(o^w; x^w, y^w, z^w)$ is the world coordinate system, which is the same as the camera coordinate system $(o^c; x^c, y^c, z^c)$. The model of the whole system can be described as follows:

$$s^c\mathbf{I}^c = \mathbf{A}^c[\mathbf{R}^c, \mathbf{t}^c]\mathbf{X}^w, \tag{21}$$

$$s^p\mathbf{I}^p = \mathbf{A}^p[\mathbf{R}^p, \mathbf{t}^p]\mathbf{X}^w. \tag{22}$$

The estimation of the intrinsic matrices $\mathbf{A}^c$ and $\mathbf{A}^p$ follows the procedure described in Sections 2.C and 2.D. For the extrinsic matrices, first of all, we can estimate the transformation matrix $[\mathbf{R}, \mathbf{t}]$ from the camera to the projector with a number of 18 poses with high accuracy. Then, since the world coordinate system $\mathbf{X}^w$ is set to be identical to the camera coordinate system $(o^c; x^c, y^c, z^c)$, the extrinsic matrix $[\mathbf{R}^c, \mathbf{t}^c]$ of the camera can be simplified as $[\mathbf{E}_3, \mathbf{0}]$, where $\mathbf{E}_3$ is a 3×3 identity matrix, and $\mathbf{0}$ is a 3×1 translation vector. Then, the transformation matrix $[\mathbf{R}, \mathbf{t}]$ will be the extrinsic matrix $[\mathbf{R}^p, \mathbf{t}^p]$ of the projector. The simplified model of the system can be represented as follows:

$$s^c\mathbf{I}^c = \mathbf{A}^c[\mathbf{E}_3, \mathbf{0}]\mathbf{X}^w, \tag{23}$$

$$s^p\mathbf{I}^p = \mathbf{A}^p[\mathbf{R}, \mathbf{t}]\mathbf{X}^w. \tag{24}$$

Once we obtain all the intrinsic and extrinsic parameters, then by simultaneously solving Eqs. (23) and (24) and applying the absolute phase constraint [20], the unknowns $(s^c, s^p, x^w, y^w, z^w)$ can be uniquely determined. Then we obtain the piece of 3D geometry from the projector–camera pair.

3. Structured-Light System Calibration Procedures

Figure 8 shows a photograph of our system setup. The projector projects phase-shifted fringe images generated by the computer onto the object; then

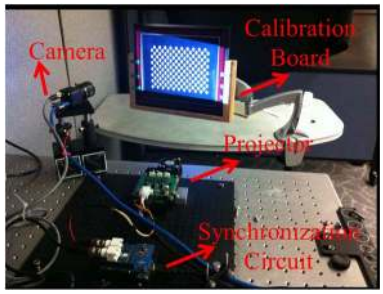


Fig. 8. Photograph of dual-camera structured-light system.

the distorted images will be captured by the camera from another view angle. A synchronization circuit is used to ensure that the camera is triggered with the projector while capturing fringe images. In this system, we used a digital light processing (DLP) projector (Model: LightCrafter 3000) with a resolution of 608×684. It has a micromirror pitch of 7.6 μm, and its pixel geometry is shown in Fig. 9(a). The camera that we used is a CMOS camera with an image resolution of 1280×1024 and a sensor size of 4.8 μm × 4.8 μm (Model: PointGrey FL3-U3-13Y3M-C). Its pixel geometry is shown in Fig. 9(b). The lens used for the camera is a Computar M0814-MP2 lens with a focal length of 8 mm at $f/1.4$ to $f/16$.

The system was calibrated using the aforementioned approach. Specifically, the system calibration requires the following major steps:

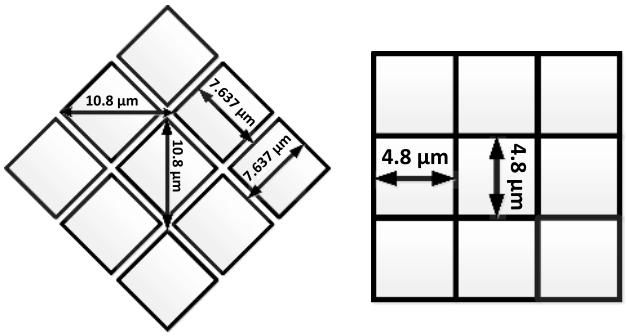


Fig. 9. Pixel geometry of the structured-light system devices. (a) DMD projector pixel geometry. (b) CMOS camera pixel geometry.

• *Step 1: Image capture*. The required images to calibrate our system include both fringe images and the actual circle pattern images for each pose of the calibration target. The fringe images were captured by projecting a sequence of horizontal and vertical phase-shifted fringe patterns for absolute phase recovery using the phase-shifting algorithm discussed in Section 2.A. The circle board image was captured by projecting uniform white images onto the board. In total, for each pose, 31 images were recorded for further analysis. Figure 10 shows an example of the captured fringe images with horizontal pattern projection, vertical pattern projection, and pure white image projection.

• *Step 2: Camera calibration*. The 18 circle board images were then used to find the circle center and then used to estimate the intrinsic parameters and lens distortion parameters of the camera. Both circle center finding and the intrinsic calibration were performed by the OpenCV camera calibration toolbox. Figure 11(a) shows one of the circle board images, and Fig. 11(b) shows the circle center we detected with the OpenCV circle center finding software algorithm. The circle detected circle centers were stored for further analysis.

• *Step 3: Projector circle center determination*. For each calibration pose, we obtained the absolute horizontal and vertical gradient phase maps (i.e., ϕ^c_{ha} and ϕ^c_{va}) using the phase-shifting algorithm. For each circle center, (u^c, v^c), found from Step 2 for this pose, the corresponding mapping point on the projector (u^p, v^p) was determined by

$$v^p = \phi^c_{va}(u^c, v^c) \times P/2\pi \tag{25}$$

$$u^p = \phi^c_{ha}(u^c, v^c) \times P/2\pi, \tag{26}$$

where P is the fringe period for the narrowest fringe pattern (18 pixels in our example). These equations simply convert phase into projector pixel. The circle center phase values were obtained by bilinear interpolation because of the subpixel circle center detection algorithm for the camera image. Figure 11(c) shows mapped circle centers for the projector. From

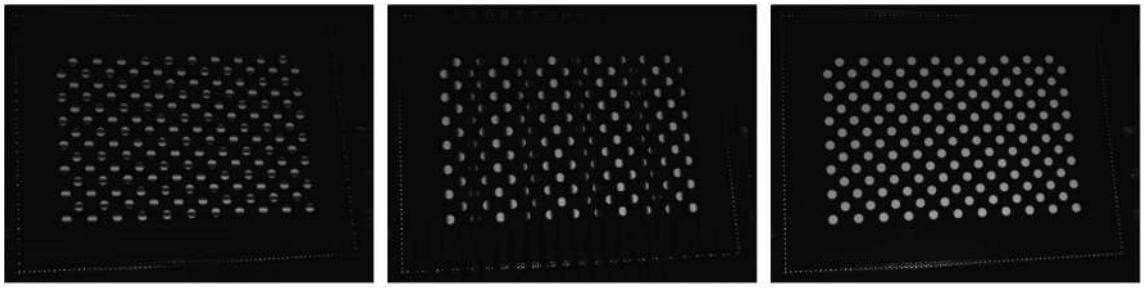
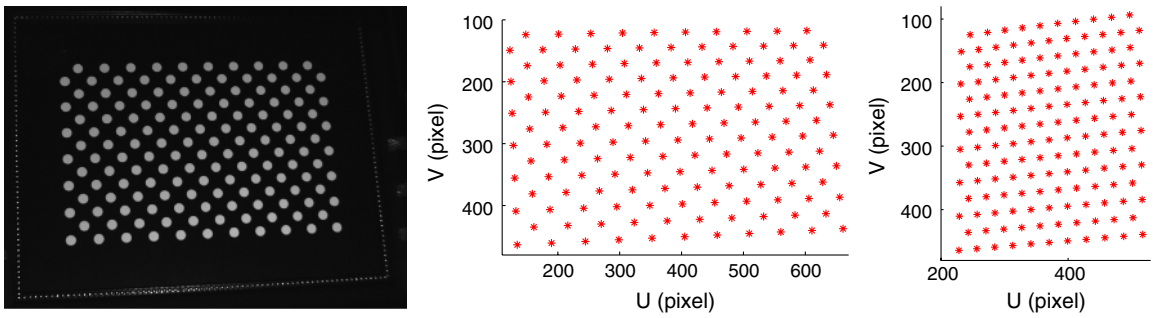


Fig. 10. Example of captured images. (a) Example of one captured fringe image with horizontal pattern projection. (b) Example of one captured fringe image with vertical pattern projection. (c) Example of one captured fringe image with pure white image projection.



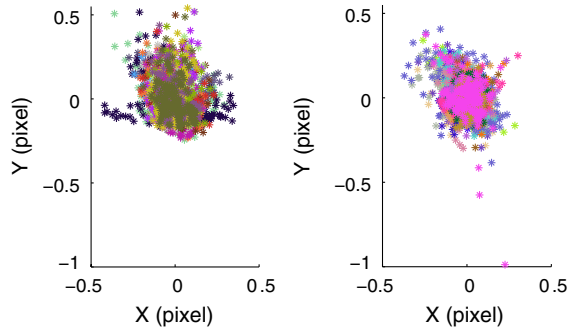
(a) (b) (c)

Fig. 11. Example of finding circle centers for the camera and the projector. (a) Example of one calibration pose. (b) Circle centers extracted from (a). (c) Mapped image for the projector from (b).

Eqs. (25) and (26), we can deduce that the mapping accuracy is not affected by the accuracy of camera parameters. However, the mapping accuracy could be influenced by the accuracy of circle center extraction and the phase quality. Since the camera circle centers were extracted by the standard OpenCV toolbox, we could obtain the coordinates of the circle centers with high accuracy. For high-quality phase generation, in general, the narrower the fringe patterns used, the better the phase accuracy that will be obtained; the more fringe patterns used, the lower the noise effect. In our research, we reduced the phase error by using a nine-step phase-shifting algorithm and the narrow fringe patterns (fringe period of $T = 18$ pixels). Our experiments to be shown in Section 4 found that this mapping was fairly accurate, which can result in a highly accurate projector calibration, similar to camera calibration, as shown in Fig. 12

It is important to note that for the calibration board, we used white circles on a black background. The main reason for this particular setup is that if we use black circles instead, the contrast of the fringe image within the area of the circles will be significantly reduced, which could lead to inaccurate phase near the circle centers, and thus inaccurate mapping point on the projector.

- *Step 4: Projector intrinsic calibration*. Once the circle centers for the projector were found from Step 3, the same software algorithms for camera calibration were used to estimate the projector's intrinsic parameters. Again, the OpenCV camera calibration toolbox is used in this research. Our experiments found that it was not necessary to consider the lens distortion for the projector, and thus we used a linear model for the projector calibration.
- *Step 5: Extrinsic calibration*. Using the OpenCV stereo calibration toolbox and the intrinsic parameters estimated previously, the extrinsic parameters can be estimated. The extrinsic parameter calibrates the transformation from the camera lens coordinate system to the projector lens coordinate system. In other words, the world coordinate system is perfectly aligned with the camera lens coordinates, making the rotation matrix $\mathbf{R}^c$ an identity matrix, and the translation vector $\mathbf{t}$ be $\mathbf{0}$.



(a) (b)

Fig. 12. Reprojection error caused by nonlinear distortion: (a) error for the camera and (b) error for the projector.

For our structured-light system, we used a total of 18 different poses for the whole system calibration. An example of calibration results (calibrated under defocusing degree 2 in Section 4) will be shown as follows. The intrinsic parameter matrices for the camera and the projector are, respectively,

$$\mathbf{A}^c = \begin{bmatrix} 1698.02 & 0 & 383.062 \\ 0 & 1691.49 & 294.487 \\ 0 & 0 & 1 \end{bmatrix}, \quad (27)$$

$$\mathbf{A}^p = \begin{bmatrix} 1019.05 & 0 & 316.763 \\ 0 & 2014.01 & 841.891 \\ 0 & 0 & 1 \end{bmatrix}, \quad (28)$$

all in pixels. As aforementioned, though the projector can be accurately calibrated, camera lens distortion is required; we found that we only need to consider the radial distortion k_1 and k_2 in Eq. (5). For our particular camera, the lens distortion is

$$\mathbf{Dist}^c = [-0.0905249 \quad 0.320865 \quad 0 \quad 0 \quad 0]^T. \quad (29)$$

Figure 12 shows the reprojection error for the camera and projector intrinsic parameter calibration. It clearly shows that the reprojection error is very small

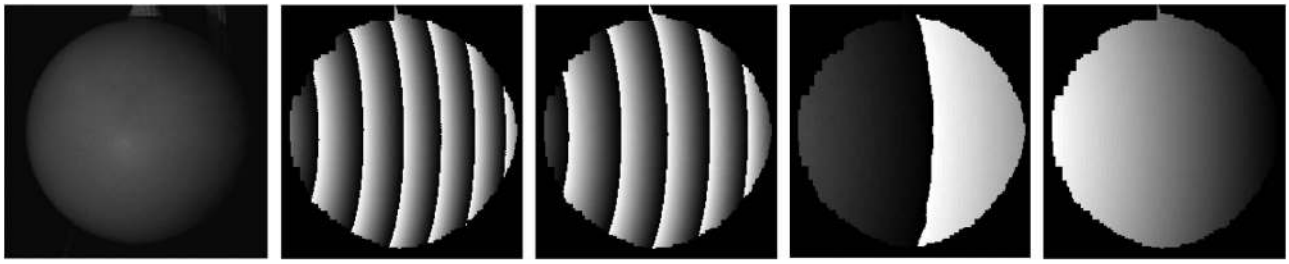


Fig. 13. Absolute phase retrieval using three-frequency phase unwrapping algorithm. (a) Picture of the spherical object. (b) Wrapped phase map obtained from patterns with fringe period $T = 18$ pixels. (c) Wrapped phase map obtained from patterns with fringe period $T = 21$ pixels. (d) Wrapped phase map obtained from patterns with fringe period $T = 154$ pixels. (e) Unwrapped phase map by applying the temporal phase unwrapping algorithm with three frequencies.

[root mean square (rms) of 0.15 pixels for the camera and 0.13 pixels for the projector], confirming that the out-of-focus projector can be accurately calibrated. One may notice that there were a few points that have relatively large reprojection errors (around 0.5 pixels). We believe the large error was caused by the circle center finding uncertainty. As described above, the calibration processes involve reorienting and repositioning the calibration target to a number of conditions. When the calibration target is parallel to the camera sensor plane, the camera imaging pixels are square and small, and thus circle centers can be accurately determined. However, when the angle between the calibration target plane and the camera sensor plane is larger, the camera imaging pixels are no longer square or small, resulting in difficulty in locating circle centers accurately from the camera image. Nevertheless, the reprojection error is overall very small, all smaller than a pixel size.

The extrinsic parameters for the camera and the projectors are, respectively, in pixels,

$$\mathbf{M}^c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{30}$$

$$\mathbf{M}^p = \begin{bmatrix} 0.952329 & -0.00367422 & 0.305051 & -162.986 \\ 0.0281659 & 0.996716 & -0.0759252 & -146.152 \\ -0.30377 & 0.0808978 & 0.949305 & 95.6518 \end{bmatrix}. \tag{31}$$

4. Experiments

To verify the performance of the proposed system calibration approach, we measured a spherical object,

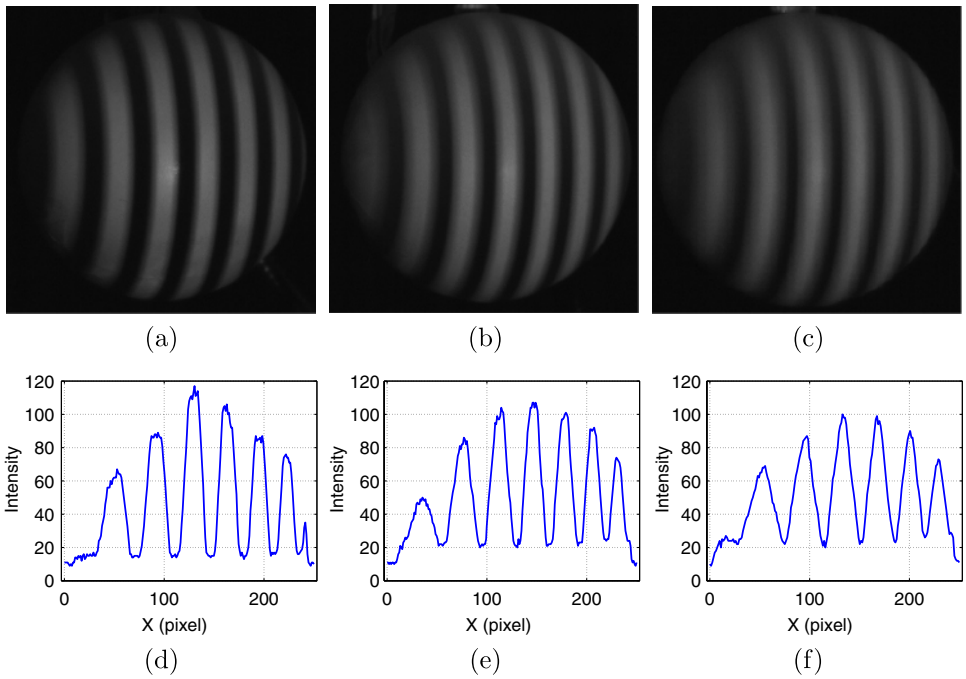


Fig. 14. Illustration of three different defocusing degrees. (a) One captured fringe image under defocusing degree 1 (projector in focus). (b) One captured fringe image under defocusing degree 2 (projector slightly defocused). (c) One captured fringe image under defocusing degree 3 (projector greatly defocused). (d)–(f) Corresponding cross sections of intensity of (a)–(c).

as is shown in Fig. 13(a); Figs. 13(b)–13(e) illustrate the three-frequency phase unwrapping algorithm that we adopted for absolute phase retrieval, which include the phase maps obtained from high-frequency ($T = 18$), medium-frequency ($T = 21$), and low-frequency ($T = 154$) fringe patterns, together with the unwrapped phase map after applying the phase unwrapping algorithm. Then, by applying the absolute phase to coordinate conversion algorithm introduced in [31], we can reconstruct the 3D geometry of the measured object. In this experiment, we measured the sphere under three different defocusing degrees: (1) the projector is in focus, (2) the projector is slightly defocused, and (3) the projector is greatly defocused. Figure 14 shows the captured fringe images under the three defocusing degrees and their corresponding cross sections of intensity. It demonstrates that when the projector is in focus, the pattern in the distorted fringe image has clear binary structure, as is shown in Fig. 14(d). However, as the projector becomes more and more defocused, the pattern will be more and more smoothed to approximate a sinusoidal structure, as is shown in Figs. 15(d)–15(h). The measurement results under three defocusing degrees are shown in Fig. 15. Figures 15(a)–15(c) show the measurement results under defocusing degree 1 (i.e., the projector is in focus), where Fig. 15(a) shows the reconstructed 3D surface. The smooth spherical surface indicates good accuracy. To further evaluate its accuracy, we took a cross section of the sphere and fitted it with an ideal circle. Figure 15(b) shows the overlay of the ideal circle and the measured data points. The difference between these two curves is shown in Fig. 15(c). The error is quite small with a rms error of 0.071 mm or 71 μm. Figures 15(d)–15(f) and Figs. 15(g)–15(i), respectively, show the measurement results under defocusing degree 2 (i.e., the projector is slightly defocused) and defocusing degree 3 (i.e., the projector is greatly defocused). In both defocusing degrees, good measurement accuracies can also be achieved, with rms errors of 77 and 73 μm, respectively. It is important to note that the whole volume of the calibration board poses was around

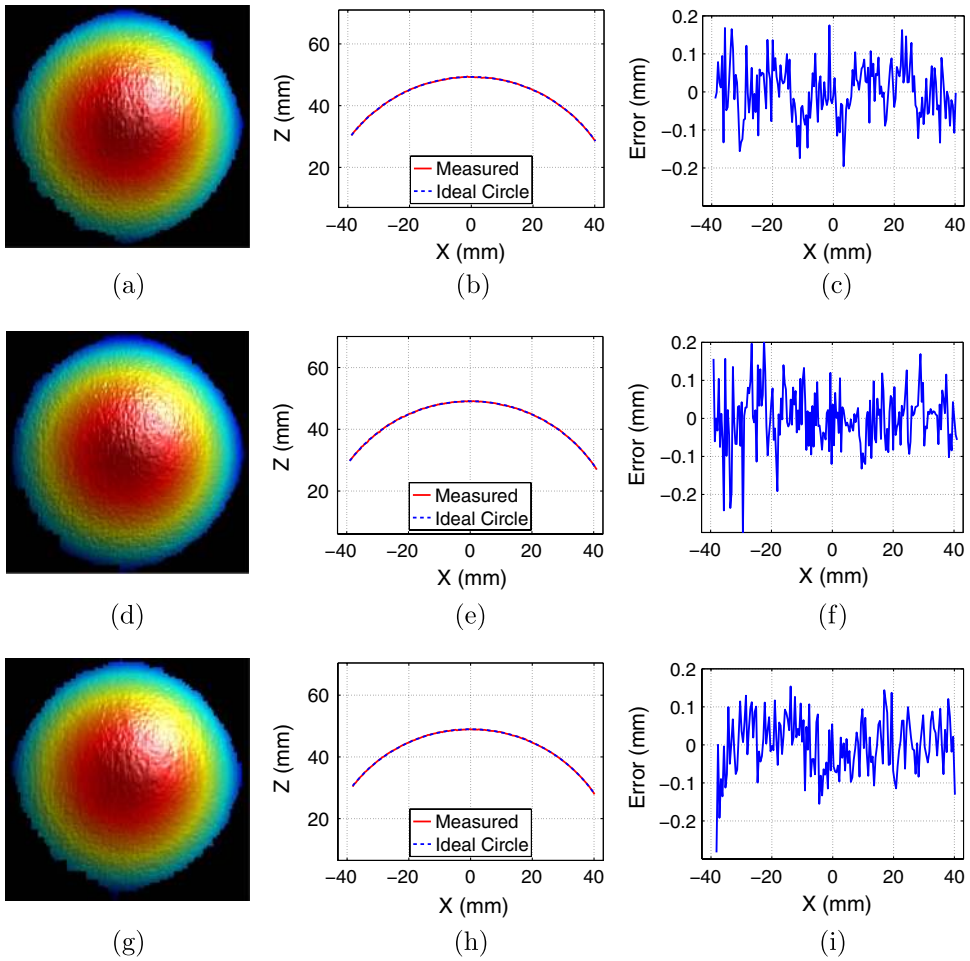


Fig. 15. Measurement results of a spherical surface under three different defocusing degrees; the rms errors estimated on (d), (h), and (l) are ±71, ±77, and ±73 μm, respectively. (a) One captured fringe image under defocusing degree 1 (projector in focus). (b) Reconstructed 3D result under defocusing degree 1 (projector in focus). (c) Cross section of the 3D result and the ideal circle under defocusing degree 1 (projector in focus). (d) Error estimated based on (b). (e)–(h) Corresponding figures of (a)–(d) under defocusing degree 2 (projector slightly defocused). (i)–(l) Corresponding figures of (a)–(d) under defocusing degree 3 (projector greatly defocused).

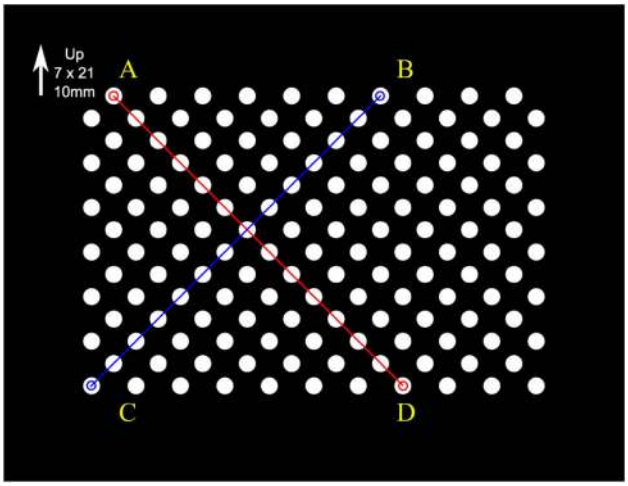


Fig. 16. Illustration of the measured diagonals on calibration board.

150(H) mm × 250(W) mm × 200(D) mm. These experimental results clearly illustrate that for such a large calibration volume, the proposed method can consistently achieve fairly high accuracy from an in-focus condition to a greatly defocused condition.

To further evaluate the calibration accuracy, we also measured the lengths of two diagonals on the calibration board under the aforementioned three different defocusing degrees, and compared the results with their actual lengths obtained using a highly accurate digital caliper. The two measured diagonals $\overline{AD}$ and $\overline{BC}$ are shown in Fig. 16, where $\overline{AD}$ is formed by the top-left and bottom-right circle center pixels, and $\overline{BC}$ is formed by the remaining two circle center pixels. It is worthwhile to note that circle centers were detected automatically with subpixel accuracy through Hough transform, and the 3D coordinates of the subpixel were obtained through bilinear interpolation. The measurement results are shown in Table 1. It again illustrates that good measurement accuracy can be achieved in all three defocusing degrees. On average, the measurement error is around 0.20 mm. Considering the lengths of the diagonals (around 183.10 mm), the relative error is quite small (around 0.12%). The major sources of error could be the error introduced by circle center detection and bilinear interpolation of 3D coordinates. Moreover, the accuracy is also subject to the precision of caliper measurement.

Table 1. Measurement Result of Two Diagonals on Calibration Board

System Setup	$\overline{AD}$ (mm)	Error (mm)	$\overline{BC}$ (mm)	Error (mm)
Defocusing degree 1	182.90	0.20	183.50	0.36
Defocusing degree 2	183.20	0.10	183.29	0.15
Defocusing degree 3	183.36	0.26	182.92	0.22
Actual	183.10	NA	183.14	NA

Furthermore, we measured a dynamically changing human face under defocusing degree 2 to demonstrate that our system can perform high-speed 3D shape measurement. In this experiment, the projection and capturing speeds were both set at 500 Hz. Moreover, in order to reduce the motion artifacts, we adopted the three-step phase-shifting algorithm for the smallest fringe period ($T = 18$ pixels) instead of the nine-step phase-shifting used previously. Figure 17 and its associated video (Media 1) demonstrate the real-time measurement results. This experiment demonstrated that high-quality 3D shape measurement can also be achieved even for the real-time 3D shape measurement.

5. Conclusion

This paper presented a calibration approach for the structured-light system with an out-of-focus projector. Our theoretical analysis provided the foundation that the out-of-focused projector can be calibrated accurately by creating one-to-one mapping between the camera pixel and the projector pixel center in the phase domain. For a calibration volume of 150(H) mm × 250(W) mm × 200(D) mm, our calibration approach has consistent performance over different amounts of defocusing, and the accuracy can reach about 73 μm. Our experimental results confirmed that high calibration accuracy can indeed be achieved by this calibration approach. One may realize that this research ignored the projector lens nonlinear distortion, which could be further considered for higher accuracy measurement. The reason for ignoring the projector nonlinearity is that our research aims at high-speed 3D shape measurement,

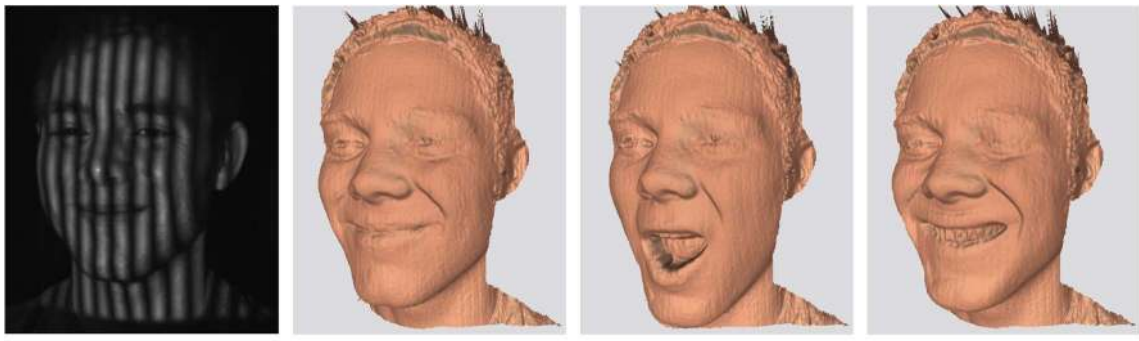


Fig. 17. Real-time 3D shape measurement result. (a) One captured fringe image. (b)–(d) Three frames of the video (Media 1) we recorded.

and only uses one-directional fringe patterns, making it difficult to directly rectify the nonlinear distortion effect caused by the projector since Eqs. (6)–(9) indicate that both u and v coordinates are needed to consider nonlinear distortion. Despite this deficiency, the achieved measurement accuracy is still very high, proving the success of the proposed calibration method.

The authors thank Willima Lohry for his valuable suggestions on OpenCV based calibration. We also thank Tyler Bell for serving as the model for evaluating the system. This study was sponsored by the National Science Foundation (NSF) under grant nos. CMMI-1150711 and CMMI-1300376. The views expressed in this paper are those of the authors and not necessarily those of the NSF.

References

1. S. Zhang, "Recent progresses on real-time 3-D shape measurement using digital fringe projection techniques," Opt. Lasers Eng. **48**, 149–158 (2010).
2. C. B. Duane, "Close-range camera calibration," Photogramm. Eng. **37**, 855–866 (1971).
3. I. Sobel, "On calibrating computer controlled cameras for perceiving 3-D scenes," Artif. Intell. **5**, 185–198 (1974).
4. R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," IEEE J. Robot. Autom. **3**, 323–344 (1987).
5. Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000).
6. J. Lavest, M. Viala, and M. Dhome, "Do we really need an accurate calibration pattern to achieve a reliable camera calibration?" in *Proceedings of the European Conference on Computer Vision* (Springer, 1998), pp. 158–174.
7. A. Albarelli, E. Rodolà, and A. Torsello, "Robust camera calibration using inaccurate targets," IEEE Trans. Pattern Anal. Mach. Intell. **31**, 376–383 (2009).
8. K. H. Strobl and G. Hirzinger, "More accurate pinhole camera calibration with imperfect planar target," in *IEEE International Conference on Computer Vision* (IEEE, 2011), pp. 1068–1075.
9. L. Huang, Q. Zhang, and A. Asundi, "Flexible camera calibration using not-measured imperfect target," Appl. Opt. **52**, 6278–6286 (2013).
10. C. Schmalz, F. Forster, and E. Angelopoulou, "Camera calibration: active versus passive targets," Opt. Eng. **50**, 113601 (2011).
11. L. Huang, Q. Zhang, and A. Asundi, "Camera calibration with active phase target: improvement on feature detection and optimization," Opt. Lett. **38**, 1446–1448 (2013).
12. Q. Hu, P. S. Huang, Q. Fu, and F.-P. Chiang, "Calibration of a three-dimensional shape measurement system," Opt. Eng. **42**, 487–493 (2003).
13. X. Mao, W. Chen, and X. Su, "Improved Fourier-transform profilometry," Appl. Opt. **46**, 664–668 (2007).
14. E. Zappa and G. Busca, "Fourier-transform profilometry calibration based on an exhaustive geometric model of the system," Opt. Lasers Eng. **47**, 754–767 (2009).
15. H. Guo, M. Chen, and P. Zheng, "Least-squares fitting of carrier phase distribution by using a rational function in fringe projection profilometry," Opt. Lett. **31**, 3588–3590 (2006).
16. H. Du and Z. Wang, "Three-dimensional shape measurement with an arbitrarily arranged fringe projection profilometry system," Opt. Lett. **32**, 2438–2440 (2007).
17. L. Huang, P. S. Chua, and A. Asundi, "Least-squares calibration method for fringe projection profilometry considering camera lens distortion," Appl. Opt. **49**, 1539–1548 (2010).
18. M. Vo, Z. Wang, B. Pan, and T. Pan, "Hyper-accurate flexible calibration technique for fringe-projection-based three-dimensional imaging," Opt. Express **20**, 16926–16941 (2012).
19. R. Legarda-Sáenz, T. Bothe, and W. P. Ju, "Accurate procedure for the calibration of a structured light system," Opt. Eng. **43**, 464–471 (2004).
20. S. Zhang and P. S. Huang, "Novel method for structured light system calibration," Opt. Eng. **45**, 083601 (2006).
21. Z. Li, Y. Shi, C. Wang, and Y. Wang, "Accurate calibration method for a structured light system," Opt. Eng. **47**, 053604 (2008).
22. Y. Yin, X. Peng, A. Li, X. Liu, and B. Z. Gao, "Calibration of fringe projection profilometry with bundle adjustment strategy," Opt. Lett. **37**, 542–544 (2012).
23. D. Han, A. Chimienti, and G. Menga, "Improving calibration accuracy of structured light systems using plane-based residual error compensation," Opt. Eng. **52**, 104106 (2013).
24. S. Lei and S. Zhang, "Flexible 3-D shape measurement using projector defocusing," Opt. Lett. **34**, 3080–3082 (2009).
25. Y. Gong and S. Zhang, "Ultrafast 3-D shape measurement with an off-the-shelf DLP projector," Opt. Express **18**, 19743–19754 (2010).
26. L. Merner, Y. Wang, and S. Zhang, "Accurate calibration for 3D shape measurement system using a binary defocusing technique," Opt. Lasers Eng. **51**, 514–519 (2013).
27. Y. Wang and S. Zhang, "Superfast multifrequency phase-shifting technique with optimal pulse width modulation," Opt. Express **19**, 5143–5148 (2011).
28. S. L. Ellenberger, *Influence of Defocus on Measurements in Microscope Images*, ASCI Dissertation Series (ASCI, 2000).
29. P. A. Stokseth, "Properties of a defocused optical system," J. Opt. Soc. Am. **59**, 1314–1321 (1969).
30. H. Hopkins, "The frequency response of a defocused optical system," Proc. R. Soc. London A **231**, 91–103 (1955).
31. S. Zhang, D. Royer, and S.-T. Yau, "GPU-assisted high-resolution, real-time 3-D shape measurement," Opt. Express **14**, 9120–9129 (2006).