

High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition

CHENG JIANG,[†] PATRICK KILCULLEN,[†]  YINGMING LAI, TSUNEYUKI OZAKI, AND JINYANG LIANG* 

Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, 1650 boulevard Lionel-Boulet, Varennes, Québec J3X1S2, Canada

*Corresponding author: jinyang.liang@emt.inrs.ca

Received 4 June 2020; revised 20 September 2020; accepted 20 September 2020; posted 22 September 2020 (Doc. ID 399492); published 30 October 2020

We report dual-view band-limited illumination profilometry (BLIP) with temporally interlaced acquisition (TIA) for high-speed, three-dimensional (3D) imaging. Band-limited illumination based on a digital micromirror device enables sinusoidal fringe projection at up to 4.8 kHz. The fringe patterns are captured alternately by two high-speed cameras. A new algorithm, which robustly matches pixels in acquired images, recovers the object's 3D shape. The resultant TIA-BLIP system enables 3D imaging over 1000 frames per second on a field of view (FOV) of up to 180 mm × 130 mm (corresponding to 1180 × 860 pixels) in captured images. We demonstrated TIA-BLIP's performance by imaging various static and fast-moving 3D objects. TIA-BLIP was applied to imaging glass vibration induced by sound and glass breakage by a hammer. Compared to existing methods in multi-view phase-shifting fringe projection profilometry, TIA-BLIP eliminates information redundancy in data acquisition, which improves the 3D imaging speed and the FOV. We envision TIA-BLIP to be broadly implemented in diverse scientific studies and industrial applications. © 2020 Chinese Laser Press

<https://doi.org/10.1364/PRJ.399492>

1. INTRODUCTION

Three-dimensional (3D) surface imaging has been extensively applied in numerous fields in industry, entertainment, and biomedicine [1,2]. Among existing methods, structured-light profilometry has gained increasing popularity in measuring dynamic 3D objects because of its high measurement accuracy and high imaging speeds [3–6]. As the most widely used method in structured-light profilometry, phase-shifting fringe projection profilometry (PSFPP) uses a set of sinusoidal fringe patterns as the basis for coordinate encoding. In contrast to other approaches of structured light, such as binary pattern projection [7], the pixel-level information carried by the phase of the fringe patterns is insensitive to variations in reflectivity across an object's surface, which enables high accuracy in 3D measurements [8]. The sinusoidal fringes employed in PSFPP are commonly generated using digital micromirror devices (DMDs). Each micromirror on the DMD can be independently tilted to either +12° or −12° from its surface normal to generate binary patterns at up to tens of kilohertz. Despite being a binary amplitude spatial light modulator [9], the DMD can be used to generate grayscale fringe patterns at high speeds [10–14]. The conventional dithering method controls the average reflectance rate of each micromirror to form a grayscale

image. However, the projection rate of fringe patterns is clamped at hundreds of hertz. To improve the projection speed, binary defocusing techniques [13] have been developed to produce a quasi-sinusoidal pattern by slightly defocusing a single binary DMD pattern. Nonetheless, the image is generated at a plane unconjugate to the DMD, which compromises the depth-sensing range and is less convenient to operate with fringe patterns of different frequencies. Recently, these limitations are lifted by the development of band-limited illumination [14], which controls the system bandwidth by placing a pinhole low-pass filter at the Fourier plane of a 4f imaging system. Both the binary defocusing method and the band-limited illumination scheme allow the generation of one grayscale sinusoidal fringe pattern from a single binary DMD pattern. Thus, the fringe projection speed matches the DMD's refresh rate.

High-speed image acquisition is also indispensable to DMD-based PSFPP. In the standard phase-shifting methods, extra calibration patterns must be used to avoid phase ambiguity [15], which reduces the overall 3D imaging speed [16]. A solution to this problem is to place multiple cameras at both sides of the projector to simultaneously capture the full sequence of fringe patterns [17–20]. These multiview approaches bring in enriched observation of 3D objects in data acquisition.

Pixel matching between different views is achieved with various assistance, including epipolar line rectification [17], measurement-volume-dependent geometry [18], and wrapped phase monotonicity [19]. Using these methods, the object's 3D surface information can be directly retrieved from the wrapped phase maps [5]. Consequently, the necessity of calibration patterns is eliminated in data acquisition and phase unwrapping. This advancement, along with the incessantly increasing imaging speeds of cameras [21–24], has endowed multiview PSFPP systems with image acquisition rates that keep up with the DMD's refresh rates.

Despite these advantages, existing multiview PSFPP systems have two main limitations. First, each camera captures the full sequence of fringe patterns. This requirement imposes redundancy in data acquisition, which ultimately clamps the systems' imaging speeds. Given the finite readout rates of camera sensors, a sacrifice of the field of view (FOV) is inevitable for higher imaging speeds. Although advanced signal processing approaches such as image interpolation [25] and compressed sensing [26] have been applied to mitigate this trade-off, they usually are accompanied by high computational complexity and reduced image quality [27]. Second, the cameras are mostly placed on different sides of the projector. This arrangement could induce a large intensity difference from the directional scattering light and the shadow effect from the occlusion by local surface features, both of which reduce the reconstruction accuracy and pose challenges in imaging non-Lambertian surfaces [20].

To overcome these limitations, we have developed dual-view band-limited illumination profilometry (BLIP) with temporally interlaced acquisition (TIA). A new algorithm is developed for coordinate-based 3D point matching from different views. Implemented with two cameras, TIA allows each to capture half of the sequence of the phase-shifted patterns, reducing the data transfer load of each camera by 50%. This freed capacity is used either to transfer data from more pixels on each camera's sensor or to support using higher frame rates of both cameras. In addition, the two cameras are placed as close as possible on the same side of the projector, which largely mitigates the intensity difference and shadow effects. Leveraging these advantages, TIA-BLIP has enabled high-speed 3D imaging of glass vibration induced by sound and glass breakage by a hammer.

2. METHOD

A. Setup

The schematic of the TIA-BLIP system is shown in Fig. 1(a). A 200 mW continuous-wave laser (wavelength $\lambda = 671$ nm, MRL-III-671, CNI Lasers) is used as the light source. After expansion and collimation, the laser beam is directed to a 0.45" DMD (AJD-4500, Ajile Light Industries) at an incident angle of $\sim 24^\circ$ to its surface normal. Four phase-shifting binary patterns, generated by an error diffusion algorithm [28] from their corresponding grayscale sinusoidal patterns, are loaded onto the DMD. A band-limited $4f$ imaging system that consists of two lenses [Lens 1 and Lens 2 in Fig. 1(a)] and one pinhole converts these binary patterns to grayscale fringes at the intermediate image plane. The two lenses have focal lengths

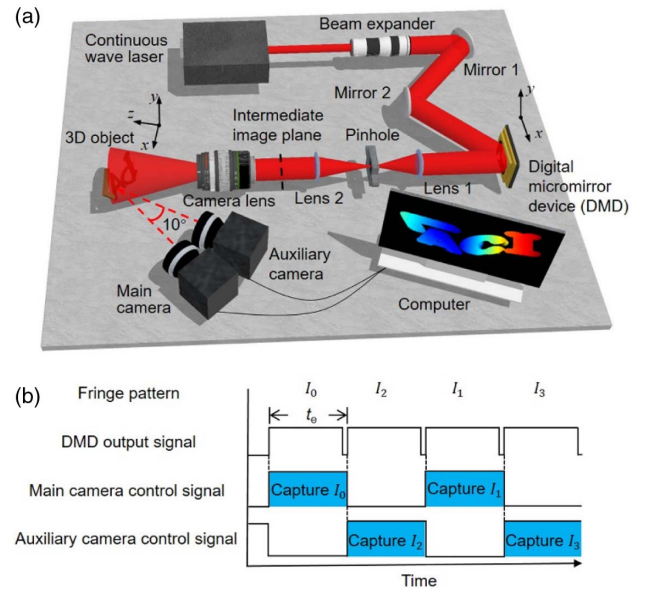


Fig. 1. Operating principle of TIA-BLIP. (a) System schematic. (b) Timing diagram and acquisition sequence. t_e , camera's exposure time.

of $f_1 = 120$ mm and $f_2 = 175$ mm. The pinhole works as a low-pass filter. Its diameter, determined by the system bandwidth, is calculated as

$$D = \frac{\lambda f_1}{p_f}, \quad (1)$$

where $p_f = 324$ μm denotes the fringe period composed of 30 DMD pixels. Thus, the required pinhole diameter is $D = 248.52$ μm . In the experiment, a 300 μm diameter pinhole is used to ensure all spatial frequency content of the sinusoidal fringe pattern passing through the system. Then, a camera lens (AF-P DX NIKKOR, Nikon, 18–55 mm focal length) projects these fringe patterns onto a 3D object. The deformed structure images are captured alternately by two high-speed CMOS cameras (CP70-1HS-M-1900, Optronis) with camera lenses (AZURE-3520MX5M, AZURE Photonics, 35 mm focal length) placed side by side. The distance between these two cameras is ~ 12 cm. The difference in their viewing angles to the 3D object is $\sim 10^\circ$. Depending on their roles in image reconstruction, they are denoted as the main camera and the auxiliary camera. Synchronized by the DMD's trigger signal, each camera captures half of the sequence [Fig. 1(b)]. The acquired images from each camera are transferred to a computer via a CoaXPress cable connected to a frame grabber (Cyton-CXP, Bitflow).

B. System Calibration

To recover 3D information from the mutually incomplete images provided by the interlaced acquisition, TIA-BLIP relies on a coordinate-based understanding of the spatial relationship of the projector and both cameras in image formation. In particular, a "pinhole" model [29],

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & \alpha & u_{pp} \\ 0 & f_v & v_{pp} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad (2)$$

describes the projection of 3D coordinates (x, y, z) onto camera coordinates (u, v) . The extrinsic parameters R and T describe the rotation and translation of coordinates, respectively. The intrinsic parameters, which link to the camera's properties in image formation, consist of f_u and f_v (representing effective focal lengths along each axis of the camera's sensor), u_{pp} and v_{pp} (describing the coordinates of the camera's principal point), and α (accounting for pixel skewness). The column vectors $[u, v, 1]^T$ and $[x, y, z, 1]^T$ represent (u, v) and (x, y, z) in the homogeneous coordinates, which through the scalar factor s allow for the numerical extraction of (u, v) from Eq. (2).

Based on the pinhole model, both cameras and the projector can be calibrated to determine the values of these parameters. Using a checkerboard as the calibration object, we adopted the established calibration procedure and software toolbox [30]. Since the direct image acquisition is not possible for a projector, the phase-based mapping method [29] was used to synthesize projector-centered images of the calibration object. These images were subsequently sent to the toolbox with calibration proceeding in the same manner as for the cameras.

C. Coordinate-Based 3D Point Determination

In the context of the pinhole model of Eq. (2), a coordinate-based method is used to recover 3D information from a calibrated imaging system. Two independent coordinates correspond to a point on a 3D object with the coordinates (x, y, z) : (u, v) for the camera and (u'', v'') for the projector. In a calibrated PSFPP system, any three of these coordinates [i.e., (u, v, u'', v'')] can be determined. Then a linear system of the form $E = M[x, y, z]^T$ is derived. The elements of E and M are found by using each device's calibration parameters as well as by using the scalar factors and the three determined coordinates [31]. In this way, the 3D information of an object point can be extracted via matrix inversion.

This analysis can be adapted to dual-view TIA-BLIP. First, images from a selected calibrated camera are used to provide the coordinates (u, v) of a point on a 3D object. Along with the system's calibration parameters, an epipolar line is determined on the other camera. The horizontal coordinate in the images of this camera is recovered using search-based algorithms along this epipolar line—a procedure commonly referred to as stereo vision. Second, by substituting a calibrated projector for the secondary camera, structured light methods use the intensity values of the pixel (u, v) across a sequence of images to recover information about a coordinate of the projector. By incorporating both aspects, 3D information can be extracted pixel by pixel based on interlaced image acquisition.

1. Data Acquisition

In data acquisition, four fringe patterns, whose phases are equally shifted by $\pi/2$, illuminate a 3D object. The intensity value for the pixel (u, v) in the k th acquired image, $I_k(u, v)$, is expressed as

$$I_k(u, v) = I_b(u, v) + I_{va}(u, v) \cos \left[\varphi(u, v) - \frac{\pi k}{2} \right], \quad (3)$$

where $k \in [0, 3]$. $I_b(u, v)$, $I_{va}(u, v)$, and $\varphi(u, v)$ represent background intensity, the variation of intensity, and depth-dependent phase, respectively.

Equation (3) allows the analysis of two types of intensity-matching conditions for the order of pattern projection shown in Fig. 1(b). The coordinates of a selected pixel in the images of the main camera are denoted by (u_m, v_m) . For a pixel (u'_a, v'_a) in the images of the auxiliary camera that perfectly corresponds with (u_m, v_m) , Eq. (3) allows us to write

$$I_0(u_m, v_m) + I_2(u'_a, v'_a) = I_1(u_m, v_m) + I_3(u'_a, v'_a). \quad (4)$$

Alternatively, rearrangement of Eq. (4) leads to the equivalent expression,

$$I_0(u_m, v_m) - I_1(u_m, v_m) = I_3(u'_a, v'_a) - I_2(u'_a, v'_a). \quad (5)$$

In contrast to Eq. (4), each side of Eq. (5), containing images captured by the same camera, computes images with a residual fringe component. Retaining sinusoidal characteristics, this residual has the effect of improving the efficiency of line-constrained searches by regularizing encountered patterns of local maxima and minima and by including additional phase information. Moreover, by interpreting its right-hand side as a continuously varying function along the epipolar line, Eq. (5) together with bilinear interpolation allows for the selection of discrete candidates with subpixel accuracy. Thus, Eq. (5) is selected as the intensity-matching condition.

2. Image Reconstruction

We developed a four-step algorithm to recover the 3D image of the object pixel by pixel. In brief, for a selected pixel (u_m, v_m) of the main camera, the algorithm locates a matching point (u'_a, v'_a) in the images of the auxiliary camera. From knowledge of the camera calibration, this point then enables the determination of estimated 3D coordinates as well as the ability to recover a wrapped phase. Using knowledge of the projector calibration, this phase value is used to calculate a horizontal coordinate on the projector's plane. A final 3D point is then recovered using the coordinate-based method. A flowchart of this algorithm is provided in Fig. 2(a).

In the first step, $(I_0 + I_1)/2$ and $(I_2 + I_3)/2$ are calculated. Then, a threshold intensity, calculated from a selected background region, is used to eliminate pixels with low intensities. The thresholding results in a binary quality map [see Step I in Fig. 2(a)]. Subsequently, only pixels that fall within the quality map of the main camera are considered for 3D information recovery.

In the second step, the selected pixel (u_m, v_m) determines an epipolar line containing the matching point within the auxiliary camera's images. Then, the algorithm extracts the candidates (u'_{ai}, v'_{ai}) that satisfy the intensity-matching condition [i.e., Eq. (5); illustrative data shown in Fig. 2(b)] in addition to three constraints [see Step II in Fig. 2(a)]. The subscript " i " denotes the i th candidate. As displayed in the illustrative data in Fig. 2(c), the first constraint requires candidates to fall within the quality map of the auxiliary camera. The second constraint requires that candidates occur within a segment of the epipolar line determined by a fixed transformation that approximates

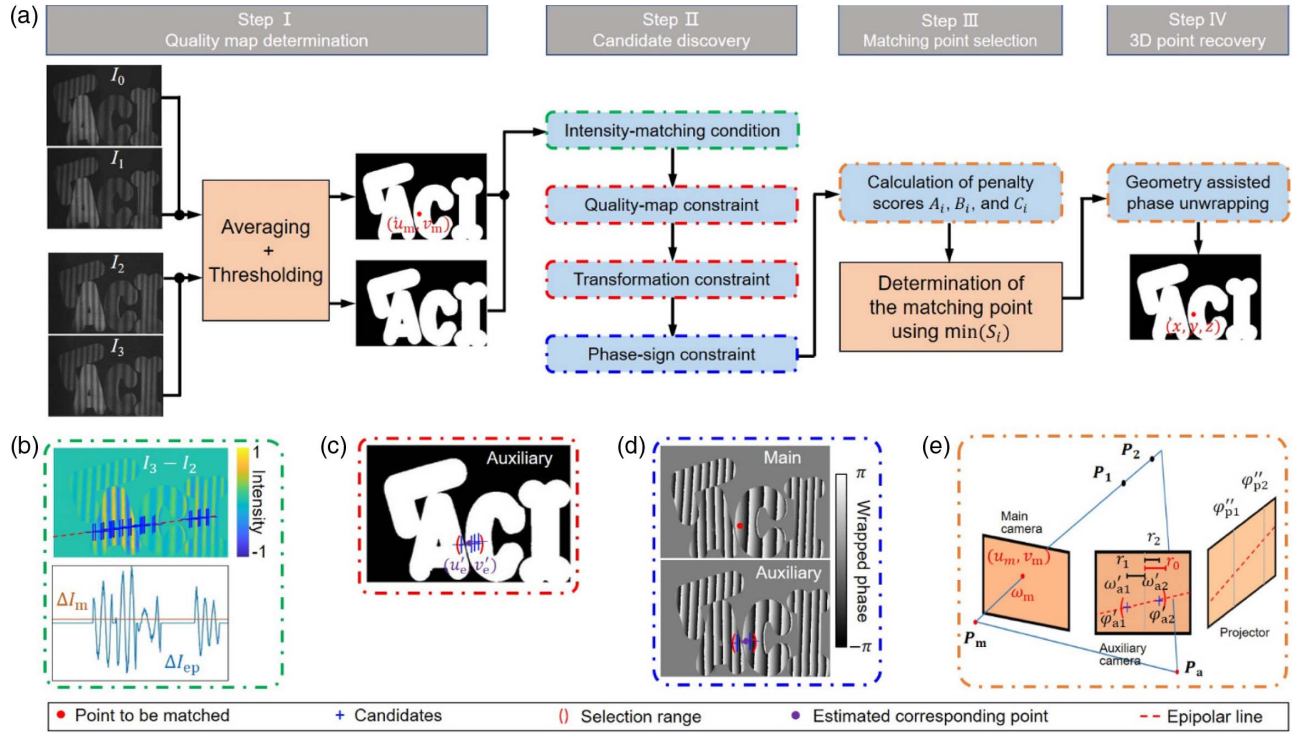


Fig. 2. Coordinate-based 3D point determination algorithm. (a) Flowchart of the algorithm. (u_m, v_m) , coordinates of the point to be matched for the main camera; (x, y, z) , recovered 3D coordinates. (b) Illustrative data of the intensity matching condition. $\Delta I_m = I_0(u_m, v_m) - I_1(u_m, v_m)$; ΔI_{ep} , intensity profile of $I_3 - I_2$ along the epipolar line. (c) Illustrative data of the quality map constraint and the transformation constraint. (u'_c, v'_c) , coordinates of the estimated corresponding point for the auxiliary camera. (d) Illustrative data of the phase sign constraint. (e) Scheme for penalty score calculation and phase unwrapping. r_i , horizontal distance between the candidates and the estimated corresponding point; ω_m , phase value of the selected point in the main camera calculated by the Fourier transform profilometry method; ω'_{ai} , phase value of the candidate points in the auxiliary camera calculated by the Fourier transform profilometry method; φ'_{ai} , phase value calculated by the phase-shifting method; φ'_{pi} , phase value determined on the projector's plane; P_i , 3D points determined by candidates; P_m , principal point of the main camera; and P_a , principal point of the auxiliary camera.

the location of the matching point. This approximation is provided by a 2D projective transformation (or homography) that determines the estimated corresponding point (u'_c, v'_c) by [32]

$$s'[u'_c, v'_c, 1]^T = H[u_m, v_m, 1]^T, \quad (6)$$

where s' is a scalar factor representing extraction of the pair (u'_c, v'_c) from its homogeneous coordinates. H is determined by applying Eq. (6) to four points chosen as the corners of a flat rectangular plane when imaged by both cameras at the approximate center of the measurement volume. Once (u'_c, v'_c) are determined, the search along the epipolar line is confined to the segment occurring over the horizontal interval $[u'_c - r_0, u'_c + r_0]$, where r_0 is an experiment-dependent constant. In general, r_0 should be as small as possible while still covering the targeted depth range. For our experiments, the value of r_0 was set to 40 pixels.

The third constraint requires the selected point and candidates to have the same sign of their wrapped phases [Fig. 2(d)]. Estimates of the wrapped phases are obtained using the technique of Fourier transform profilometry [5]. In particular, by bandpass filtering the left side of Eq. (5), i.e., $I_0 - I_1$, the intensity of pixel (u_m, v_m) in the filtered image is

$$I_f(u_m, v_m) = \frac{\sqrt{2}}{2} I_{va}(u_m, v_m) \exp \left\{ j \left[\varphi(u_m, v_m) + \frac{\pi}{4} \right] \right\}. \quad (7)$$

The wrapped phase estimation ω_m of (u_m, v_m) is recovered by

$$\omega_m = \arctan \left\{ \frac{\Im[I_f(u_m, v_m)]}{\Re[I_f(u_m, v_m)]} \right\}, \quad (8)$$

where $\Im[\cdot]$ and $\Re[\cdot]$, respectively, denote the imaginary part and the real part of a complex variable. The same treatment is applied to the right side of Eq. (5), i.e., $I_3 - I_2$. For the candidate (u'_{ai}, v'_{ai}) , the estimate of its wrapped phase is calculated as

$$\omega'_{ai} = \arctan \left\{ \frac{\Im[I'_f(u'_{ai}, v'_{ai})]}{\Re[I'_f(u'_{ai}, v'_{ai})]} \right\}. \quad (9)$$

The constraint requires ω_m and ω'_{ai} to have the same sign in the interval $(-\pi, \pi]$. The output of the second step is a pool of candidates for further evaluation. If no candidate is found, the algorithm abandons the following process, and this step is reinitiated for the next pixel in the main camera.

In the third step, three criteria are used to calculate penalty scores for each candidate, as shown in Step III in Fig. 2(a). The scheme is shown in Fig. 2(e). The first and primary criterion compares the phase values of the candidates using two methods.

First, the phase inferred from the intensities of candidates and the pixel (u_m, v_m) is calculated by

$$\phi'_{ai} = \arctan \left[\frac{I_1(u_m, v_m) - I_3(u'_{ai}, v'_{ai})}{I_0(u_m, v_m) - I_2(u'_{ai}, v'_{ai})} \right]. \quad (10)$$

Meanwhile, for each candidate (u'_{ai}, v'_{ai}) , the coordinate triple (u_m, v_m, u'_{ai}) and knowledge of camera calibration allow the determination of an estimated 3D point P_i by using the stereo vision method. In addition, with the knowledge of the projector calibration, a point with coordinates (u''_{pi}, v''_{pi}) on the projector's plane is determined for each candidate. Then, an unwrapped phase value ϕ''_{pi} is calculated by

$$\phi''_{pi} = \frac{2\pi}{p} (u''_{pi} - u''_d), \quad (11)$$

where u''_d is a horizontal datum coordinate on the projector's plane associated with the zero phase, and p is the fringe period in units of projector pixels. Since these independently inferred phase values must agree if the candidate correctly matches (u_m, v_m) , a penalty score A_i , as a normalized difference of these two phase values, is calculated by

$$A_i = \frac{|R(\phi'_{ai} - \phi''_{pi})|}{\pi}, \quad (12)$$

where the rewrapping function $R(\cdot)$ computes the subtracted difference between the wrapped and unwrapped phase values.

To improve the robustness of the algorithm, two additional criteria are implemented using data available from the second step. B_i is a normalized distance score favoring candidates located closer to the estimated matching point (u'_e, v'_e) , which is calculated by

$$B_i = \frac{|u'_e - u'_{ai}|}{r_0}. \quad (13)$$

Moreover, C_i is a normalized difference of the wrapped phase values using the wrapped phases ω_m and ω'_{ai} , which is calculated by

$$C_i = \frac{|R(\omega_m - \omega'_{ai})|}{\pi}. \quad (14)$$

A total penalty score S_i for each candidate is then computed as a weighted linear combination of three individual scores,

$$S_i = \eta_1 A_i + \eta_2 B_i + \eta_3 C_i, \quad (15)$$

where the normalized weights $[\eta_1, \eta_2, \eta_3] = [0.73, 0.09, 0.18]$ are empirically chosen to lead to the results that are most consistent with the physical reality. Finally, the candidate with the minimum S_i is chosen as the matching point (u'_a, v'_a) . Its phase values, calculated by using Eqs. (10) and (11), are denoted as ϕ'_a and ϕ''_p , respectively.

In the final step, the algorithm determines the final 3D coordinates [see Step IV in Fig. 2(a) and the scheme in Fig. 2(e)]. First, ϕ'_a is unwrapped as $\phi'_a + 2\pi q$, where q is an integer making $\phi''_p - (\phi'_a + 2\pi q) \in (-\pi, \pi]$. Then, the coordinate on the projector's plane, u''_p , is recovered with subpixel resolution as

$$u''_p = u''_d + P(\phi'_a/2\pi + q), \quad (16)$$

from which the final 3D coordinates (x, y, z) are computed using the calibration information associated with the coordinate triple (u_m, v_m, u''_p) .

3. RESULTS

A. Quantification of Depth Resolution

To quantify the depth resolution of TIA-BLIP with different exposure times, we imaged two stacked planar surfaces offset by $\sim 9^\circ$ (Fig. 3). Reconstructed results at four representative exposure times (denoted as t_e) are shown in Fig. 3(a). One area on each surface [marked as white solid boxes in Fig. 3(a)] was selected in the reconstructed image. The depth information on the x axis was calculated by averaging the depth values along the y axis. The difference in depths between these two surfaces is denoted by z_d . In addition, the noise is defined as the averaged value of the standard deviations in depth from both surfaces. The depth resolution is defined as when z_d equals twice the system's noise level. As shown in the four plots in Fig. 3(a), the reconstruction results deteriorate with shorter exposure time, manifested by increased noise levels and more points incapable of retrieving 3D information. As a result, the depth resolution degrades from 0.06 mm at $t_e = 950 \mu\text{s}$ to 0.45 mm at $t_e = 150 \mu\text{s}$ [Fig. 3(b)]. At $t_e = 100 \mu\text{s}$, TIA-BLIP fails in 3D measurements. The region of unsuccessful reconstruction prevails across most of the planar surfaces. The noise dominates the calculated depth difference, which is attributed to the low signal-to-noise ratio in the captured images.

B. Imaging of Static 3D Objects

To examine the feasibility of TIA-BLIP, we imaged various static 3D objects. First, two sets of 3D distributed letter toys that composed the words "LACI" and "INRS" were imaged. Shown in Fig. 4(a), the two perspective views of the reconstructed results reveal the 3D position of each letter toy. The detailed surface structures are illustrated by the selected depth profiles [see the white dashed lines and the magenta

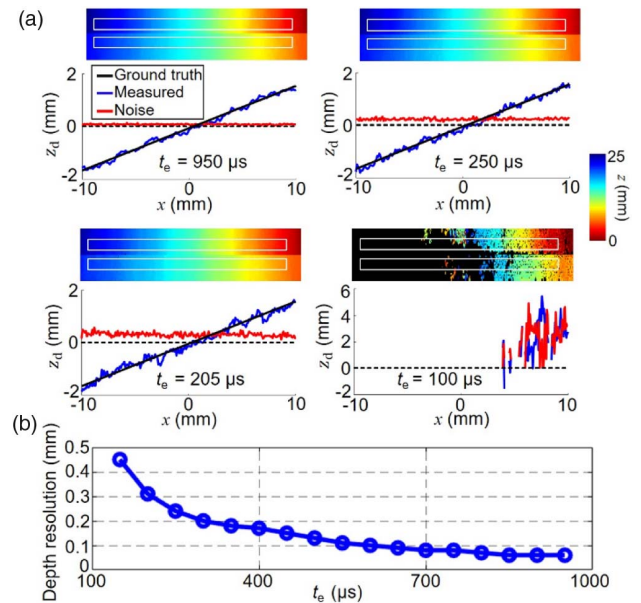


Fig. 3. Quantification of TIA-BLIP's depth resolution. (a) 3D images of the planar surfaces (top image) and measured depth difference (bottom plot) at four exposure times. The white boxes represent the selected regions for analysis. (b) Relation between measured depth differences and their corresponding exposure times.

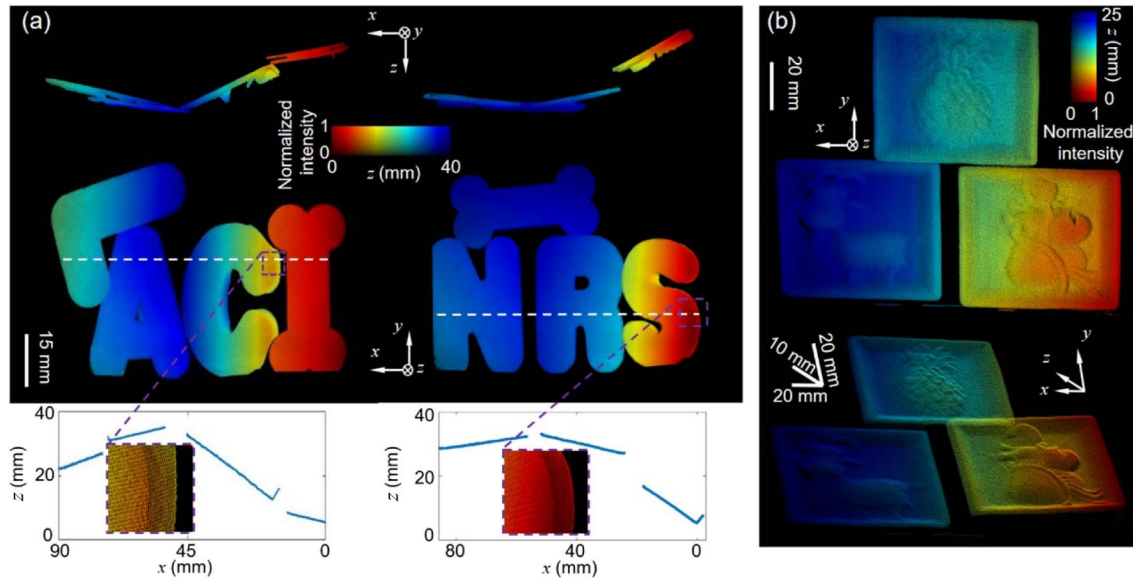


Fig. 4. TIA-BLIP of static 3D objects. (a) Reconstructed results of letter toys. Two perspective views are shown in the top row. Selected depth profiles (marked by the white dashed lines in the top images) and close-up views are shown in the bottom row. (b) Two perspective views of the reconstruction results of three cubes.

dashed boxes in Fig. 4(a)]. We also conducted a proof-of-concept experiment on three cube toys with fine structures (with a depth of ~ 4 mm) on the surfaces. Depicted in Fig. 4(b), the detailed structural information of these cube toys is recovered by TIA-BLIP.

C. Imaging of Dynamic 3D Objects

To verify high-speed 3D surface profilometry, we used TIA-BLIP to image two dynamic scenes: a moving hand and three bouncing balls. The fringe patterns were projected at 4 kHz. The exposure times of both cameras were $t_e = 250$ μ s. Under these experimental conditions, TIA-BLIP had a 3D imaging speed of 1 thousand frames per second (kfps), an FOV of 180 mm \times 130 mm (corresponding to 1180 \times 860 pixels) in captured images, and a depth resolution of 0.24 mm. Figure 5(a) shows the reconstructed 3D images of the moving hand at five time points from 0 ms to 60 ms with a time interval of 15 ms (see the full evolution in Visualization 1). TIA-BLIP's high-speed 3D imaging allowed tracking the movements of four fingertips. Shown in Fig. 5(b), all the four fingers have apparent movement in both the x axis and the z axis, but stay relatively stationary in the y axis, which agrees with the experimental condition.

In the second experiment, three white balls, each of which was marked by a different letter on its surface, bounced in an inclined transparent container. Figure 5(c) shows five representative reconstructed images from 8 ms to 28 ms with a time interval of 5 ms. The changes of the letter "C" on B_1 and the letter "L" on B_2 [marked in the third panel of Fig. 5(c)] clearly show the rotation of the two balls (see the full evolution in Visualization 2). TIA-BLIP enabled tracking the 3D centroids of each ball over time. Shown in Fig. 5(d), B_1 collides with B_2 at 16 ms, resulting in a sudden change in the moving directions. This collision temporarily interrupted the free fall of

B_1 , represented by the two turning points in the curve of evolution along the y axis [see the second panel of Fig. 5(d)]. The collision also changed the moving direction of B_2 , making it touch the base at 27 ms and then bounce up. In this scene, B_3 maintained its movement in a single direction in both the x axis and the z axis. It fell onto the base and bounced back at 16 ms, resulting in a turning point in its y - t curve. Because of the inclined bottom plane, the y value of B_3 at 16 ms was smaller than that of B_2 at 27 ms.

Under the same experimental settings and pattern sequence choice, TIA-BLIP surpasses the existing PSFPP techniques in pixel counts and hence the imaging FOV. At the 1 kfps 3D imaging speed, the systems of standard single-camera PSFPP [14] and multiview PSFPP [19] would restrict their imaging FOV to 512 \times 512 pixels and 768 \times 640 pixels, respectively [33]. In contrast, TIA, with a frame size of 1180 \times 860 pixels, increases the FOV by 3.87 and 2.07 times, respectively.

D. Imaging of Sound-Induced Vibration on Glass

To highlight the broad utility of TIA-BLIP, we imaged sound-induced vibration on glass. In this experiment [Fig. 6(a)], a glass cup was fixed on a table. The glass's surface was painted white. A function generator drove a speaker to produce single-frequency sound signals (from 450 Hz to 550 Hz with a step of 10 Hz) through a sound channel placed close to the cup's wall. To image the vibration dynamics, fringe patterns were projected at 4.8 kHz. The cameras had an exposure time of $t_e = 205$ μ s. This configuration enabled a 3D imaging speed of 1.2 kfps, an FOV of 146 mm \times 130 mm (corresponding to 960 \times 860 pixels) in captured images, and a depth resolution of 0.31 mm. Figure 6(b) shows four representative 3D images of the instantaneous shapes of the glass cup driven by the 500 Hz sound signal (the full sequence is shown in Visualization 3), showing the dynamic of structural deformation

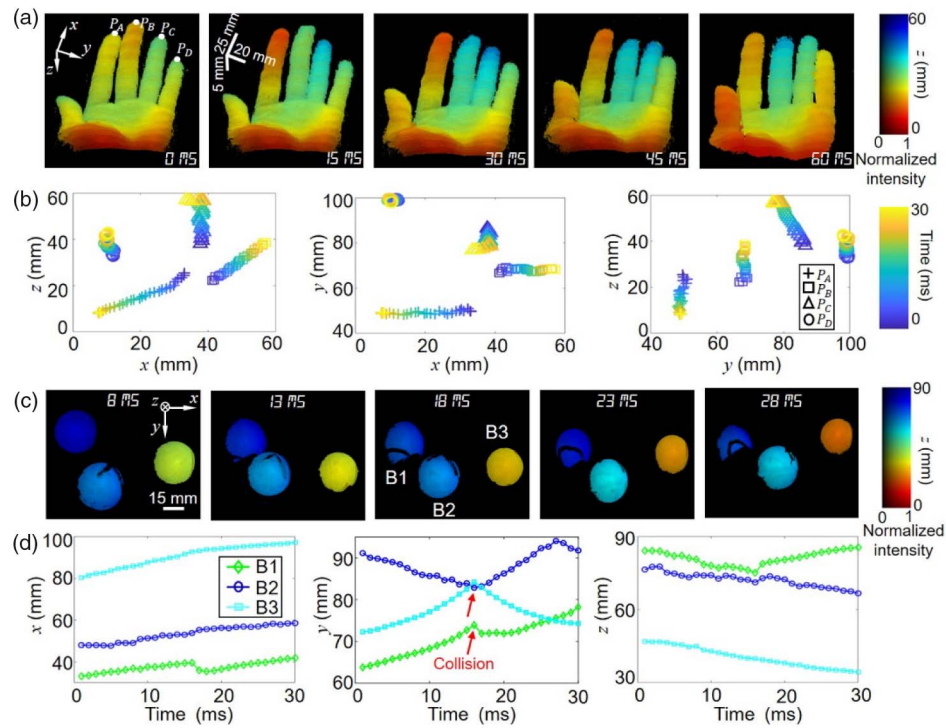


Fig. 5. TIA-BLIP of dynamic objects. (a) Reconstructed 3D images of a moving hand at five time points. (b) Movement traces of four fingertips [marked in the first panel in (a)]. (c) Front view of the reconstructed 3D image of the bouncing balls at five different time points. (d) Evolution of 3D positions of the three balls [marked in the third panel in (c)].

of the glass cup. The evolution of depth changes was analyzed using five selected points [marked by P_A to P_E in the first panel of Fig. 6(b)]. Shown in Fig. 6(c), the depth changes of five points are in accordance, which is attributed to the rigidity of the glass.

We further analyzed time histories of averaged depth displacements under different sound frequencies. Figure 6(d)

shows the results at the driving frequencies of 490 Hz, 500 Hz, and 510 Hz. Each result was fitted by a sinusoidal function with a frequency of 490.0 Hz, 499.4 Hz, and 508.6 Hz, respectively. These results show that the rigid glass cup vibrated in compliance with the driving frequency. Moreover, the amplitudes of fitted results Δz_{fit} were used to

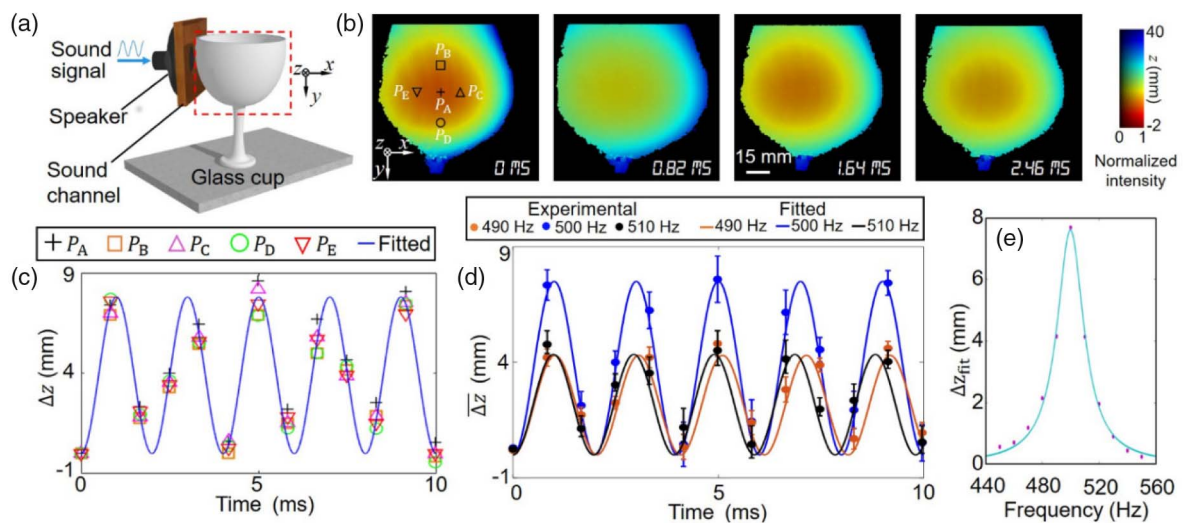


Fig. 6. TIA-BLIP of sound-induced vibration on glass. (a) Schematic of the experimental setup. The field of view is marked by the red dashed box. (b) Four reconstructed 3D images of the cup driven by a 500-Hz sound signal. (c) Evolution of the depth change of five points marked in the first panel of (b) with the fitted result. (d) Evolution of the averaged depth change with the fitted results under the driving frequencies of 490 Hz, 500 Hz, and 510 Hz. Error bar: standard deviation of Δz calculated from the five selected pixels. (e) Response of the depth displacements to sound frequencies. The cyan curve is a fitted result of a Lorentz function.

determine the relationship between the depth displacement and the sound frequency [Fig. 6(e)]. We fitted this result by the Lorentz function, which determined the resonant frequency of this glass cup to 499.0 Hz.

It is worth noting that this phenomenon would be difficult to be captured by using previous methods given the same experimental settings and pattern sequence choice. With frame size of 960×860 pixels, the maximum frame rate for the used cameras is 2.4 kfps, which transfers to a 3D imaging speed of 480 fps for single-camera-based PSFPP and 600 fps for multi-view PSFPP. Neither provides sufficient imaging speed to visualize the glass vibration at the resonance frequency. In contrast, TIA improves the 3D imaging speed to 1.2 kfps, which is fully capable of sampling the glass vibration dynamics in the tested frequency range of 450–550 Hz.

E. Imaging of Glass Breakage

To further apply TIA–BLIP to recording nonrepeatable 3D dynamics, we imaged the process of glass breaking by a hammer (the full sequence is shown in Visualization 4). As displayed in Fig. 7(a), the growth of cracks and the burst of fragments with different shapes and sizes are clearly shown in the reconstructed 3D images. The time courses of velocities of four fragments [marked by F_A to F_D in Fig. 7(a)] are plotted in Fig. 7(b). The velocities in the y axis are considerably small compared to the other two directions, which indicates the impact of the hammer force was exerted on the x – z plane. The v_y of fragments F_A and F_C shows that they moved upward until 13 ms and fell afterward. The v_y of fragments F_B and F_D reveals that they fell onto the remaining base of the cup at 15 ms and kept sliding down on the surface. The data of v_z illustrates that F_A and F_C moved closer to the cameras, which was directly driven by the hammer's force. However, F_B and F_D , which collided with other pieces, maintained their positive

directions in v_z to move away from the cameras. The corresponding accelerations are displayed in Fig. 7(c), which indicates the influence of both the main strike and the ensuing collision among different fragments. At 14 ms, the collision with other fragments, which applied an impact along the $+x$ direction, dominated the acceleration direction for all four tracked fragments. In contrast, at 15 ms, another collision produced an impact in the $-x$ direction, causing a sharp decrease in the acceleration for F_A and F_C . In addition, the direction of acceleration for F_D along the y axis changed several times, which is attributed in several collisions of F_D with the base of the glass cup while sliding down.

4. DISCUSSION AND CONCLUSIONS

We have developed TIA–BLIP with a kfps-level 3D imaging speed over an FOV of up to $180 \text{ mm} \times 130 \text{ mm}$ (corresponding to 1180×860 pixels) in captured images. This technique implements TIA in multiview 3D PSFPP systems, which allows each camera to capture half of the sequence of the phase-shifting fringes. Leveraging the characteristics indicated in the intensity-matching condition [i.e., Eq. (5)], the newly developed algorithm applies constraints in geometry and phase to find the matching pair of points in the main and auxiliary cameras and guides phase unwrapping to extract the depth information. TIA–BLIP has empowered the 3D visualization of glass vibration induced by sound and the glass breakage by a hammer.

TIA–BLIP possesses many advantages. First, TIA eliminates the redundant capture of fringe patterns in data acquisition. The roles of the main camera and the auxiliary camera are interchangeable. Despite being demonstrated only with high-speed cameras, TIA–BLIP is a universal imaging paradigm easily adaptable to other multiview PSFPP systems. Second, TIA reduces the workload for each camera employed in the

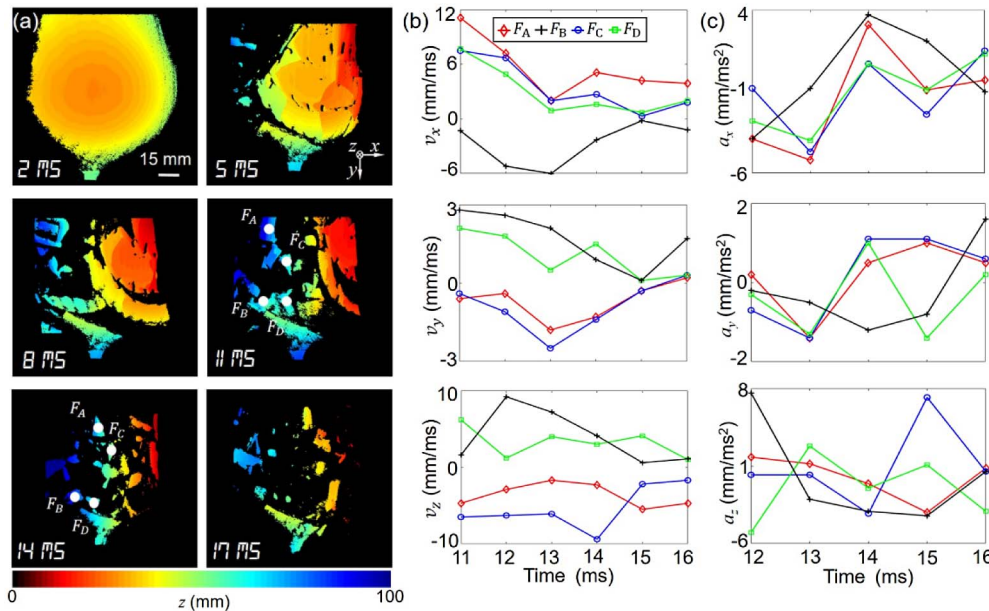


Fig. 7. TIA–BLIP of glass breakage. (a) Six reconstructed 3D images showing a glass cup broken by a hammer. (b) Evolution of 3D velocities of four selected fragments marked in the fourth and fifth panels in (a). (c) Evolution of the corresponding 3D accelerations of the four selected fragments.

multiview systems. The freed capacity is used to enhance the technical specifications in PSFPP. In particular, at a certain frame rate, more pixels on the sensors of the deployed cameras can be used, which increases the imaging FOV. Alternatively, if the FOV is fixed, TIA supports these cameras to have higher frame rates, which thus increases the 3D imaging speed. Both advantages shed light on implementing TIA-BLIP with an array of cameras to simultaneously accomplish high accuracy and high-speed 3D imaging over a larger FOV. Third, the two cameras deployed in the current TIA-BLIP system are placed side by side. Compared with the existing dual-view PSFPP systems that mostly place the cameras at different sides of the projector, the arrangement in TIA-BLIP circumvents the intensity difference induced by the directional scattering light from the 3D object and reduces the shadow effect by occlusion. Both merits support robust pixel matching in the image reconstruction algorithm to recover 3D information on non-Lambertian surfaces.

Future work will be carried out in the following aspects. First, we plan to further improve TIA-BLIP's imaging speed and FOV in three ways: by separating the workload to an array of cameras, by implementing a faster DMD, and by using a more powerful laser. Moreover, we will implement depth-range estimation and online feedback to reduce the time in candidate discovery. Furthermore, parallel computing will be used to increase the speed of image reconstruction toward real-time operation [34]. Finally, to robustly image 3D objects with different sizes and with incoherent light sources, we will generate fringe patterns with adaptive periods by using a slit or a pinhole array as the spatial filter [35]. Automated size calculation [36] also will be integrated into the imaging processing software to facilitate the determination of the proper fringe period.

Besides technical improvements, we will continue to explore new applications of TIA-BLIP. For example, it could be integrated into structure illumination microscopy [37] and frequency-resolved multidimensional imaging [38]. TIA-BLIP could also be implemented in the study of the dynamic characterization of glass in its interaction with the external forces in nonrepeatable safety test analysis [39–41]. As another example, TIA-BLIP could trace and recognize the hand gesture in 3D space to provide information for human–computer interaction [42]. Furthermore, in robotics, TIA-BLIP could provide a dual-view 3D vision for object tracking and reaction guidance [43]. Finally, TIA-BLIP can function as an imaging accelerometer for vibration monitoring in rotating machinery [44] and for behavior quantification in biological science [45].

Funding. Natural Sciences and Engineering Research Council of Canada (ALLRP-549833-20, ALLRP-551076-20, CRDPJ-532304-18, RGPAS-507845-2017, RGPIN-2017-05959); Canada Foundation for Innovation (37146); Fonds de recherche du Québec–Nature et technologies (2019-NC-252960); Fonds de Recherche du Québec–Santé (267406, 280229).

Acknowledgment. The authors thank Xianglei Liu for experimental assistance.

Disclosures. The authors declare no conflicts of interest.

[†]These authors contributed equally to this paper.

REFERENCES

1. X. Su and Q. Zhang, "Dynamic 3-D shape measurement method: a review," *Opt. Lasers Eng.* **48**, 191–204 (2010).
2. P. Kilcullen, C. Jiang, T. Ozaki, and J. Liang, "Camera-free three-dimensional dual photography," *Opt. Express* **28**, 29377–29389 (2020).
3. S. S. Gorthi and P. Rastogi, "Fringe projection techniques: whither we are?" *Opt. Lasers Eng.* **48**, 133–140 (2010).
4. S. Van der Jeught and J. J. Dirckx, "Real-time structured light profilometry: a review," *Opt. Lasers Eng.* **87**, 18–31 (2016).
5. S. Zhang, "Absolute phase retrieval methods for digital fringe projection profilometry: a review," *Opt. Lasers Eng.* **107**, 28–37 (2018).
6. J. Liang, "Punching holes in light: Recent progress in single-shot coded-aperture optical imaging," *Rep. Prog. Phys.* (in press, 2020).
7. I. Ishii, K. Yamamoto, K. Doi, and T. Tsuji, "High-speed 3D image acquisition using coded structured light projection," in *IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2007), 925–930.
8. J. Geng, "Structured-light 3D surface imaging: a tutorial," *Adv. Opt. Photon.* **3**, 128–160 (2011).
9. J. Liang, M. F. Becker, R. N. Kohn, and D. J. Heinzen, "Homogeneous one-dimensional optical lattice generation using a digital micromirror device-based high-precision beam shaper," *J. Micro/Nanolith. MEMS MOEMS* **11**, 023002 (2012).
10. L. J. Hornbeck, "Digital light processing for high-brightness high-resolution applications," *Proc. SPIE* **3013**, 27–41 (1997).
11. J. Liang, S.-Y. Wu, R. N. Kohn, M. F. Becker, and D. J. Heinzen, "Grayscale laser image formation using a programmable binary mask," *Opt. Eng.* **51**, 108201 (2012).
12. S. Lei and S. Zhang, "Flexible 3-D shape measurement using projector defocusing," *Opt. Lett.* **34**, 3080–3082 (2009).
13. B. Li, Y. Wang, J. Dai, W. Lohry, and S. Zhang, "Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques," *Opt. Lasers Eng.* **54**, 236–246 (2014).
14. C. Jiang, P. Kilcullen, X. Liu, J. Gribben, A. Boate, T. Ozaki, and J. Liang, "Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface," *Opt. Lett.* **45**, 964–967 (2020).
15. Y. Wang and S. Zhang, "Superfast multifrequency phase-shifting technique with optimal pulse width modulation," *Opt. Express* **19**, 5149–5155 (2011).
16. C. Zuo, Q. Chen, G. Gu, S. Feng, and F. Feng, "High-speed three-dimensional profilometry for multiple objects with complex shapes," *Opt. Express* **20**, 19493–19510 (2012).
17. D. Li, H. Zhao, and H. Jiang, "Fast phase-based stereo matching method for 3D shape measurement," in *International Symposium on Optomechatronic Technologies* (IEEE, 2010), pp. 1–5.
18. C. Bräuer-Burchardt, C. Munkelt, M. Heinze, P. Kühmstedt, and G. Notni, "Using geometric constraints to solve the point correspondence problem in fringe projection based 3D measuring systems," in *International Conference on Image Analysis and Processing* (Springer, 2011), pp. 265–274.
19. Z. Li, K. Zhong, Y. F. Li, X. Zhou, and Y. Shi, "Multiview phase shifting: a full-resolution and high-speed 3D measurement framework for arbitrary shape dynamic objects," *Opt. Lett.* **38**, 1389–1391 (2013).
20. W. Yin, S. Feng, T. Tao, L. Huang, M. Trusiak, Q. Chen, and C. Zuo, "High-speed 3D shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system," *Opt. Express* **27**, 2411–2431 (2019).
21. B. Li, P. Ou, and S. Zhang, "High-speed 3D shape measurement with fiber interference," *Proc. SPIE* **9203**, 920310 (2014).
22. N. L. Karpinsky, M. Hoke, V. Chen, and S. Zhang, "High-resolution, real-time three-dimensional shape measurement on graphics processing unit," *Opt. Eng.* **53**, 024105 (2014).
23. C. Jiang, T. Bell, and S. Zhang, "High dynamic range real-time 3D shape measurement," *Opt. Express* **24**, 7337–7346 (2016).

24. J.-S. Hyun and S. Zhang, "Superfast 3D absolute shape measurement using five binary patterns," *Opt. Lasers Eng.* **90**, 217–224 (2017).
25. M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 277–285 (1991).
26. X. Liu, J. Liu, C. Jiang, F. Vetrone, and J. Liang, "Single-shot compressed optical-streaking ultra-high-speed photography," *Opt. Lett.* **44**, 1387–1390 (2019).
27. C. Lei, Y. Wu, A. C. Sankaranarayanan, S.-M. Chang, B. Guo, N. Sasaki, H. Kobayashi, C.-W. Sun, Y. Ozeki, and K. Goda, "GHz optical time-stretch microscopy by compressive sensing," *IEEE Photon. J.* **9**, 7500207 (2017).
28. J. Liang, R. N. Kohn, Jr., M. F. Becker, and D. J. Heinzen, "1.5% root-mean-square flat-intensity laser beam formed using a binary-amplitude spatial light modulator," *Appl. Opt.* **48**, 1955–1962 (2009).
29. S. Zhang and P. S. Huang, "Novel method for structured light system calibration," *Opt. Eng.* **45**, 083601 (2006).
30. J.-Y. Bouguet, "Camera calibration toolbox for MATLAB," http://www.vision.caltech.edu/bouguetj/calib_doc/.
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).
32. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision* (Cambridge University, 2003).
33. OPTRONIS GMBH, *User Manual CP70-1HS-M/C-1900* (2018), Vol. 4.
34. W. Gao and Q. Kemao, "Parallel computing in experimental mechanics and optical measurement: a review," *Opt. Lasers Eng.* **50**, 608–617 (2012).
35. C. Chang, J. Liang, D. Hei, M. F. Becker, K. Tang, Y. Feng, V. Yakimenko, C. Pellegrini, and J. Wu, "High-brightness X-ray free-electron laser with an optical undulator by pulse shaping," *Opt. Express* **21**, 32013–32018 (2013).
36. C. Hoppe, M. Klopschitz, M. Rumpler, A. Wendel, S. Kluckner, H. Bischof, and G. Reitmayr, "Online feedback for structure-from-motion image acquisition," in *British Machine Vision Conference (BMVC)*, 2012, pp. 1–12.
37. J. Qian, M. Lei, D. Dan, B. Yao, X. Zhou, Y. Yang, S. Yan, J. Min, and X. Yu, "Full-color structured illumination optical sectioning microscopy," *Sci. Rep.* **5**, 14513 (2015).
38. K. Dorozynska, V. Kornienko, M. Aldén, and E. Kristensson, "A versatile, low-cost, snapshot multidimensional imaging approach based on structured light," *Opt. Express* **28**, 9572–9586 (2020).
39. A. Ramos, F. Pelayo, M. Lamela, A. F. Canteli, C. Huerta, and A. Acios, "Evaluation of damping properties of structural glass panes under impact loading," in *COST Action TU0905 Mid-Term Conference on Structural Glass* (2013).
40. C. Bedon, M. Fasan, and C. Amadio, "Vibration analysis and dynamic characterization of structural glass elements with different restraints based on operational modal analysis," *Buildings* **9**, 13 (2019).
41. M. Haldimann, A. Luible, and M. Overend, *Structural Use of Glass* (IABSE, 2008), Vol. 10.
42. Y. Li, "Hand gesture recognition using Kinect," in *IEEE International Conference on Computer Science and Automation Engineering* (IEEE, 2012), pp. 196–199.
43. S. Huang, K. Shinya, N. Bergström, Y. Yamakawa, T. Yamazaki, and M. Ishikawa, "Dynamic compensation robot with a new high-speed vision system for flexible manufacturing," *Int. J. Adv. Manuf. Technol.* **95**, 4523–4533 (2018).
44. R. B. Randall, "State of the art in monitoring rotating machinery-part 1," *Sound Vib.* **38**, 14–21 (2004).
45. T. A. Van Walsum, A. Perna, C. M. Bishop, C. P. Murn, P. M. Collins, R. P. Wilson, and L. G. Halsey, "Exploring the relationship between flapping behaviour and accelerometer signal during ascending flight, and a new approach to calibration," *Int. J. Avian Sci.* **162**, 13–26 (2020).