Multi-scale band-limited illumination profilometry for robust three-dimensional surface imaging at video rate

CHENG JIANG, PATRICK KILCULLEN, YINGMING LAI, SIQI WANG, TSUNEYUKI OZAKI, AND JINYANG LIANG*

Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, Université du Québec, 1650 Boulevard Lionel-Boulet, Varennes, Québec J3X1P7, Canada
*jinyang.liang@inrs.ca

Abstract: Dynamic three-dimensional (3D) surface imaging by phase-shifting fringe projection profilometry has been widely implemented in diverse applications. However, existing techniques fall short in simultaneously providing the robustness in solving spatially isolated 3D objects, the tolerance of large variation in surface reflectance, and the flexibility of tunable working distances with meter-square-level fields of view (FOVs) at video rate. In this work, we overcome these limitations by developing multi-scale band-limited illumination profilometry (MS-BLIP). Supported by the synergy of dual-level intensity projection, multi-frequency fringe projection, and an iterative method for distortion compensation, MS-BLIP can accurately discern spatially separated 3D objects with highly varying reflectance. MS-BLIP is demonstrated by dynamic 3D imaging of a translating engineered box and a rotating vase. With an FOV of up to 1.7 m × 1.1 m and a working distance of up to 2.8 m, MS-BLIP is applied to capturing full human-body movements at video rate.

1. Introduction

Three-dimensional (3D) surface imaging techniques [1–4] have found diverse applications, including industrial manufacturing, archaeological inspection, entertainment, and biomedicine [5–9]. The incessant demands of higher accuracy in extracting surface spatial information in complex scenes pose strict requirements to existing 3D imaging methods. For example, a millimeter-level spatial resolution is desired in reliable 3D facial recognition for attendance checks [10]. Meanwhile, video-rate 3D imaging is necessary for fluid-flag interaction analysis in applied bionics [11]. Moreover, spatially isolated 3D objects with large depth discontinuities must be identified for *in-situ* monitoring and robotic classification [12]. Furthermore, a high dynamic range is indispensable to measure highly reflective objects in defect inspection of steel castings for the automobile industry [13]. Finally, a large field of view (FOV), ideally covering the full human body, is needed to detect and collect 3D movements for virtual reality gaming [14].

Among existing techniques, phase-shifting fringe projection profilometry (PSFPP) has proven to be a potent technique for 3D surface imaging [15]. In the most widely used configuration, PSFPP works by first projecting sets of phase-shifting sinusoidal fringe patterns onto 3D objects and then analyzing deformed structure images reflected from the objects to retrieve 3D surface information [16]. In contrast to other structured-light 3D imaging techniques [17–20], PSFPP uses sinusoidal fringe patterns to encode grayscale phase values for each pixel, followed by phase unwrapping to extract 3D information with a high depth resolution [21].

The speed of sinusoidal fringe projection is one of the most important specifications in PSFPP. These patterns are commonly generated by digital micromirror devices (DMDs) [22]. Though being a binary amplitude spatial light modulator [23], the DMD can generate grayscale sinusoids

in various ways. The conventional dithering method forms a grayscale image by controlling the average reflectance of each micromirror over time, which clamps the projection rate at hundreds of hertz [24]. To improve the fringe projection speed, binary defocusing techniques [25,26] and band-limited illumination profilometry (BLIP) [27,28] have been developed, both of which can produce a grayscale sinusoidal pattern from a single binary DMD mask. Their fringe pattern projection speeds can keep up with the DMD's refreshing rate [up to tens of kilohertz (kHz)], showing promise for high-speed 3D visualizations.

Although overcoming the limitation in projection speed, binary defocusing could bring instability in pattern quality. In this method, quasi-sinusoidal fringe patterns are generated behind the system's image plane. Consequently, DMD's uneven surface could induce image distortion to the defocused binary DMD masks at any unconjugate plane, which may decrease the 3D measurement accuracy, especially under coherent illumination [29]. In addition, because the proper defocusing degree to eliminate the high-order harmonics from the binary mask is frequency-dependent [30], this technique is less compatible with sinusoidal patterns of different periods. Furthermore, the defocusing operation reduces the contrast of quasi-sinusoidal fringes [31], especially at the far end of the measurement volume. Difficult to maintain the defocusing degree for fringe patterns with different frequencies at a fixed image plane, binary defocusing techniques pose challenges in imaging spatially isolated 3D objects [32]. Finally, the binary defocusing technique compromises the depth-sensing range [25]. Hence, it has not reached a cubic meter (m^3)-level measurement volume, falling short in multi-scale 3D imaging with tunable working distances and FOVs.

BLIP can overcome the limitations in binary defocusing techniques. It generates sinusoids by controlling the optical bandwidth of a $4f$ imaging system. As a result, regardless of their frequencies, sinusoidal fringe patterns are always generated on the image plane of the DMD. This configuration eliminates the distortion brought by uneven DMD surface and maintains the contrast of projected fringe patterns. It is also inherently compatible with multi-frequency fringe projection, varying working distances, and different FOVs. Despite these attractive advantages, to date, BLIP has not yet been demonstrated for imaging 3D objects with highly reflective surfaces. Meanwhile, BLIP has only been applied with the single-frequency fringe projection method, which is less robust for reconstructing the 3D information of spatially separated objects. Finally, limited by the power of the light sources and the focal lengths of the lenses on the projector and the camera in the existing apparatus [33], BLIP's FOV has not yet reached the meter-square (m^2) level.
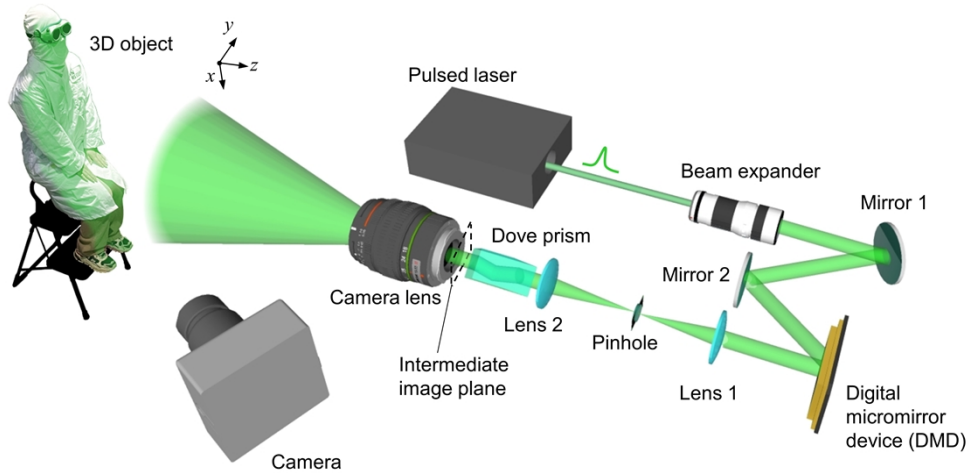
To surmount these limitations, in this paper, we report multi-scale (MS) BLIP for robust 3D surface imaging at video rate. MS-BLIP implements multi-frequency fringe projection with the associated phase unwrapping, which enables imaging spatially isolated 3D objects. Meanwhile, MS-BLIP adopts dual-level intensity projection to enhance its dynamic range, which allows recovering 3D information from surfaces with high reflectance. Moreover, MS-BLIP applies an iterative method for distortion compensation, which improves the 3D reconstruction quality over a m^2-level FOV. Finally, MS-BLIP is demonstrated for video-rate 3D surface measurements at varying working distances of up to 2.8 m with tunable FOVs of up to 1.7 m × 1.1 m.

2. Method

2.1. Setup

The schematic of the MS-BLIP system is shown in Fig. 1. A pulsed laser with an average power of 400 mW (AONano 559-0.2-10, Advanced Optowave) 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 ~24° to its surface normal. Binary fringe masks, generated by an error diffusion algorithm from their corresponding grayscale patterns [34], are loaded onto the DMD and displayed at up to 1 kHz. A band-limited $4f$ imaging system that consists of two lenses

(Lens 1 and Lens 2 in Fig. 1; AC254-075-A, Thorlabs) and one pinhole converts these binary patterns to grayscale fringes at the intermediate image plane. The smallest period in the used sinusoidal fringe patterns is 388.8 µm (i.e., 36 DMD pixels), which demands a 150-µm-diameter pinhole (P150K, Thorlabs) to pass the spatial frequency components of these patterns while filtering all noise induced by the digital halftoning [35]. A dove prism (PS992M, Thorlabs), placed between Lens 2 and the intermediate image plane, rotates the generated fringe patterns to match the aspect ratio of the targeted scene. Then, a camera lens (AF-P DX NIKKOR 10-20mm f/4.5-5.6G VR, Nikon) projects these fringe patterns onto 3D objects. The deformed structure images are captured by a high-speed CMOS camera (CP70-1HS-M-1900, Optronis) with a camera lens (AZURE-3520MX5M, AZURE Photonics). Synchronized by the DMD's trigger signal, the acquired images are transferred to a computer via a CoaXPress cable connected to a frame grabber (Cyton-CXP, Bitflow).

Fig. 1. Schematic of the MS-BLIP system.

2.2. Operating principle

2.2.1. System modeling

The projective behaviors of the camera and the projector in the MS-BLIP system are modeled based on the "pinhole model" with the consideration of image distortion. For a 3D point (x, y, z) in the world coordinate system, in a perfect imaging condition, the pinhole model describes its corresponding pixel on the camera sensor, $(\hat{u}, \hat{v})$, to be

$$s \begin{bmatrix} \hat{u} \\ \hat{v} \\ 1 \end{bmatrix} = A[R\ T] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_{pp} \\ 0 & f_v & v_{pp} \\ 0 & 0 & 1 \end{bmatrix} [R\ T] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \tag{1}$$

Here, the matrix A contains the camera's intrinsic parameters: f_u and f_v describe the effective focal lengths along the axes of the camera, and (u_{pp}, v_{pp}) is the coordinate of the camera's principal point. Defined as the camera's extrinsic parameters, R is a 3×3 matrix accounting for rotation, and T is a 3×1 matrix for translation. Finally, s is a scalar factor of numerical extraction of $(\hat{u}, \hat{v})$. Overall, this operation is expressed by $(\hat{u}, \hat{v}) = \text{Proj}_c(x, y, z)$. Similarly, the

projection of the 3D point to the projector is modeled by $(\hat{u}'', \hat{v}'') = \mathrm{Proj}_\mathrm{p}(x, y, z)$. The superscript of "double prime" is used hereafter to refer to the coordinates for the projector.

To take into consideration of distortion in the acquisition of deformed structure images, we first define the normalized camera coordinate without distortion as

$$\begin{bmatrix} \hat{u}_\mathrm{n} \\ \hat{v}_\mathrm{n} \end{bmatrix} = \begin{bmatrix} (\hat{u} - u_\mathrm{pp})/f_u \\ (\hat{v} - v_\mathrm{pp})/f_v \end{bmatrix}. \tag{2}$$

Then, the distorted normalized camera coordinate $(u_\mathrm{n}, v_\mathrm{n})$ is determined by [36]

$$\begin{bmatrix} u_\mathrm{n} \\ v_\mathrm{n} \end{bmatrix} = (1 + d_1\hat{r}_\mathrm{n}^2 + d_2\hat{r}_\mathrm{n}^4 + d_3\hat{r}_\mathrm{n}^6) \begin{bmatrix} \hat{u}_\mathrm{n} \\ \hat{v}_\mathrm{n} \end{bmatrix} + \begin{bmatrix} 2d_4\hat{u}_\mathrm{n}\hat{v}_\mathrm{n} + d_5(\hat{r}_\mathrm{n}^2 + 2\hat{u}_\mathrm{n}^2) \\ 2d_5\hat{u}_\mathrm{n}\hat{v}_\mathrm{n} + d_4(\hat{r}_\mathrm{n}^2 + 2\hat{v}_\mathrm{n}^2) \end{bmatrix}, \tag{3}$$

where $d_1, \ldots, d_5$ are the camera system distortion coefficients, and $\hat{r}_\mathrm{n}^2 = \hat{u}_\mathrm{n}^2 + \hat{v}_\mathrm{n}^2$. Once $(u_\mathrm{n}, v_\mathrm{n})$ is computed, the distorted camera pixel coordinate (u, v) is found by using Eq. (2). Overall, this operation is expressed as $(u, v) = \mathrm{Dist}_\mathrm{c}(\hat{u}, \hat{v})$. Similarly, the distortion for the projector is modeled by $(u'', v'') = \mathrm{Dist}_\mathrm{p}(\hat{u}'', \hat{v}'')$.

2.2.2. System calibration

MS-BLIP's system calibration aims to compute the intrinsic parameters, extrinsic parameters, and image distortion coefficients for both the camera and the projector. In the calibration of the camera, a checkerboard pattern was imaged with 20–40 poses. A MATLAB toolbox [37] was used to extract the grid corners, which allowed calculating all the calibration parameters. In the calibration of the projector, the projector was treated as another camera to adopt the similar process used in camera calibration [38]. The method involved capturing additional images of the checkerboard pattern under the illumination of both horizontally and vertically shifted fringe patterns to determine the pixel-wise mapping between the camera view and the projector view. Then, the camera-captured images of the checkerboard were transformed to generate the corresponding images from the perspective of the projector. Finally, the same MATLAB toolbox used in camera calibration was implemented to compute the corresponding calibration parameters of the projector.

2.2.3. Dual-level intensity projection

To image 3D objects with a large variety of reflectance, projection intensity modulation [39] is implemented to improve MS-BLIP's dynamic range. In particular, two normalized intensity levels, empirically chosen to be 1.0 and 0.3, are set for each fringe pattern. For any camera pixel in the acquired deformed structure images, if any values in the sequence of high-intensity projection are saturated, then it is replaced with the corresponding sequence with low intensity.

2.2.4. Coordinate recovery using multi-frequency fringe projection

MS-BLIP implements the multi-frequency fringe projection method [40]. The intensity of the camera pixel (u, v) in the captured deformed structure images can be expressed as

$$I_{km}(u, v) = I_\mathrm{b}(u, v) + I_\mathrm{va}(u, v)\cos[\varphi_m(u, v) - 2\pi k/K]. \tag{4}$$

Here, $I_{km}(u, v)$ represents the intensity in the kth deformed structure image in the mth sequence. K is the number of phase-shifting steps. In this work, we use four-step phase shifting in six sets of fringes, so that $K = 4$, $k \in [0, 3]$, and $m \in [0, 5]$. $I_\mathrm{b}(u, v)$, $I_\mathrm{va}(u, v)$, and $\varphi_m(u, v)$ represent the background, intensity variation, and depth-dependent phase for the mth sequence, respectively. From Eq. (4), the calculated $\varphi_m(u, v)$ is wrapped in the interval $(-\pi, \pi]$. Then, phase unwrapping

is performed to determine the horizontal coordinate of the projector, which is then used to solve the 3D surface information by triangulation.

In single-frequency fringe projection, a single unwrapped phase map is generated by a two-dimensional weighted phase unwrapping algorithm [41], which must make assumptions about the surface smoothness. The unwrapped phase value of each pixel is derived according to the phase values within a local neighborhood of this pixel. Consequently, this method falls short in solving the phase ambiguities induced by the depth-discontinued regions or isolated objects [42].

This limitation is lifted by multi-frequency fringe projection that unwraps the phase value of each pixel individually. The values of P_p^m in our work were determined by synthetically considering the system's signal-to-noise ratios and the measurement accuracy. The coarsest fringe pattern is chosen to have no more than one period so that its absolute phase, defined as $\tilde{\varphi}_0(u, v)$, equals the wrapped phase map $\varphi_0(u, v)$. The ensuing sets of fringe patterns, i.e., $m>0$, have the periods of $P_p^m = Round(N_u/2^m)$. Here, N_u is the width of the DMD expressed in the unit of DMD pixel. Because unwrapped phase maps of two adjacent sets of fringe projection have the relationship of $\tilde{\varphi}_m(u, v) = (P_p^{m-1}/P_p^m)\tilde{\varphi}_{m-1}(u, v)$, the absolute phase map for $m>0$ can be calculated by

$$\tilde{\varphi}_m(u, v) = \varphi_m(u, v) + 2\pi \cdot Round\left[\frac{(P_p^{m-1}/P_p^m)\tilde{\varphi}_{m-1}(u, v) - \varphi_m(u, v)}{2\pi}\right]. \qquad (5)$$

Equation (5) shows that, unlike single-frequency phase unwrapping, the phase of each pixel is unwrapped independently, which avoids incorrect depth quantification in analyzing spatially isolated objects [42]. Using this method, the unwrapped phase map of each set of fringe patterns is computed successively, and $\tilde{\varphi}_5(u, v)$ is used to compute the horizontal coordinate of the projector, denoted by u'', by

$$u'' = \tilde{\varphi}_5(u, v)\frac{P_p^5}{2\pi} + \frac{1}{2}(N_u - 1). \qquad (6)$$

2.2.5. Distortion compensation and 3D point recovery

Distortion compensation contains two major steps. First, the undistorted camera pixel coordinate $(\hat{u}, \hat{v})$ is extracted by performing the inverse operation of the distortion model (i.e., $Dist_c$). However, the direct inversion of Eq. (3) takes the form of a 7th degree polynomial in both $\hat{u}_n$ and $\hat{v}_n$, rendering direct recovery difficult. Thus, we apply a fixed point iteration method [43] for computing the undistorted normalized camera coordinate $(\hat{u}_n, \hat{v}_n)$. With the initial condition of

$\begin{bmatrix} \hat{u}_{n,0} \\ \hat{v}_{n,0} \end{bmatrix} = \begin{bmatrix} (u - u_{pp})/f_u \\ (v - v_{pp})/f_v \end{bmatrix}$, the ith iteration is described by

$$\begin{bmatrix} \hat{u}_{n, i+1} \\ \hat{v}_{n, i+1} \end{bmatrix} = \frac{\begin{bmatrix} \hat{u}_{n, i} \\ \hat{v}_{n, i} \end{bmatrix} - \begin{bmatrix} 2d_4\hat{u}_{n, i}\hat{v}_{n, i} + d_5(\hat{r}_{n, i}^2 + 2\hat{u}_{n, i}^2) \\ 2d_5\hat{u}_{n, i}\hat{v}_{n, i} + d_4(\hat{r}_{n, i}^2 + 2\hat{v}_{n, i}^2) \end{bmatrix}}{1 + d_1\hat{r}_{n, i}^2 + d_2\hat{r}_{n, i}^4 + d_3\hat{r}_{n, i}^6}, \qquad (7)$$

where $\hat{r}_{n,i}^2 = \hat{u}_{n,i}^2 + \hat{v}_{n,i}^2$. The corresponding undistorted pixel coordinate $(\hat{u}, \hat{v})$ on the camera plane is then calculated by using Eq. (2). Overall, this operation is expressed as $(\hat{u}, \hat{v}) = DistComp_c(u, v)$.

The next step is to obtain the corresponding undistorted projector horizontal pixel coordinate $\hat{u}''$. In Eq. (1), we define $A[R\,T] = \begin{bmatrix} a_{11} & \cdots & a_{14} \\ \vdots & \ddots & \vdots \\ a_{31} & \cdots & a_{34} \end{bmatrix}$ and $A''[R''\,T''] = \begin{bmatrix} b_{11} & \cdots & b_{14} \\ \vdots & \ddots & \vdots \\ b_{31} & \cdots & b_{34} \end{bmatrix}$. By

using Proj_c and Proj_p, the coordinate of the distortion-compensated 3D point is calculated by triangulation as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \text{Tri}(\hat{u}, \ \hat{v}, \ \hat{u}'') = \begin{bmatrix} a_{11} - \hat{u}a_{31} & a_{12} - \hat{u}a_{32} & a_{13} - \hat{u}a_{33} \\ a_{21} - \hat{v}a_{31} & a_{22} - \hat{v}a_{32} & a_{23} - \hat{v}a_{33} \\ b_{11} - \hat{u}''b_{31} & b_{12} - \hat{u}''b_{32} & b_{13} - \hat{u}''b_{33} \end{bmatrix}^{-1} \begin{bmatrix} \hat{u}a_{34} - a_{14} \\ \hat{v}a_{34} - a_{24} \\ \hat{u}''b_{34} - b_{14} \end{bmatrix}. \quad (8)$$

Ideally, $\hat{u}''$ could be computed by $(\hat{u}'', \ \hat{v}'') = \text{DistComp}_p(u'', v'')$. However, v'' cannot be calculated from the projection of vertical fringe patterns. To overcome this limitation, we have developed an iterative method that recovers distortion-compensated 3D information without prior knowledge of v'' from fringe measurements. We first estimate the coordinate of the 3D point by the Tri function [i.e., Eq. (8)] by using the coordinate $(\hat{u}, \ \hat{v}, \ u'')$. This 3D point is then projected to the projector plane to extract the initial estimate of v'' with the function of Dist_p, which is input to an iterative algorithm. Each iteration starts with performing the function DistComp_p by using u'' calculated from (Eq. (6)) and the variable v'' to compute the distortion compensated projector coordinate $\hat{u}''$, which is then used to extract the 3D coordinate by the function Tri. The method is presented as the following pseudo-code:

Algorithm to recover 3D information with distortion compensation

---

    Input: $(\hat{u}, \hat{v}, u'')$, *maximum iterations N, vertical projector pixel error tolerance TOL,*
    calibration parameters of camera: A, R, T, s, d_1, d_2, d_3, d_4, d_5;
    calibration parameters of projector: A'', R'', T'', s'', d_1'', d_2'', d_3'', d_4'', d_5''
    Output: (x,y,z)
    Variables: q, h_1, h_2, h_3, h_4, v_q'', $\hat{u}''$

1   *Set $q = 1$*

2   *compute $(x,y,z) = \text{Tri}(\hat{u}, \hat{v}, u'')$*

3   *compute $(h_1, h_2) = \text{Proj}_p(x,y,z)$*

4   *compute $(\sim, v_q'') = \text{Dist}_p(h_1, h_2)$*

5   **while** *($q \leq N$) do steps 6 – 11*

6   *compute $(\hat{u}'', \sim) = \text{DistComp}_p\left(u'', v_q''\right)$*

7   *compute $(x,y,z) = \text{Tri}(\hat{u}, \hat{v}, \hat{u}'')$*

8   *compute $(h_3, h_4) = \text{Proj}_p(x,y,z)$*

9   *compute $\left(\sim, v_{q+1}''\right) = \text{Dist}_p(h_3, h_4)$*

10   *if $\left|v_{q+1}'' - v_q''\right| \leq TOL$, go to step 13*

11   *Set $q = q + 1$*

12   **End**

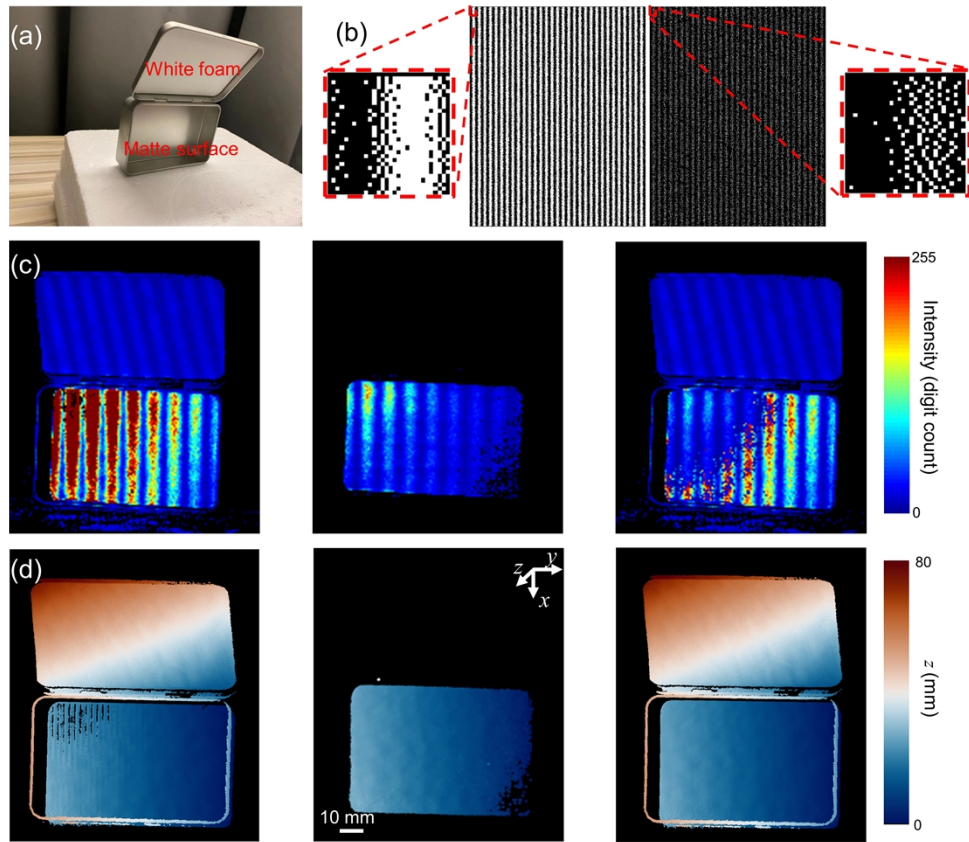**13**   **Output** (x,y,z)

14   **End**

---

3. Results

3.1. Demonstration of MS-BLIP

To verify the feasibility of MS-BLIP, we imaged static 3D objects at a working distance of 0.5 m, with an FOV of 0.3 m × 0.2 m. The depth resolution, quantified by the standard deviation of a

reconstructed planar surface [28], was measured to be 0.3 mm. The lateral resolution, quantified by imaging the sharp edge of the planar surface [44], was 0.4 mm.
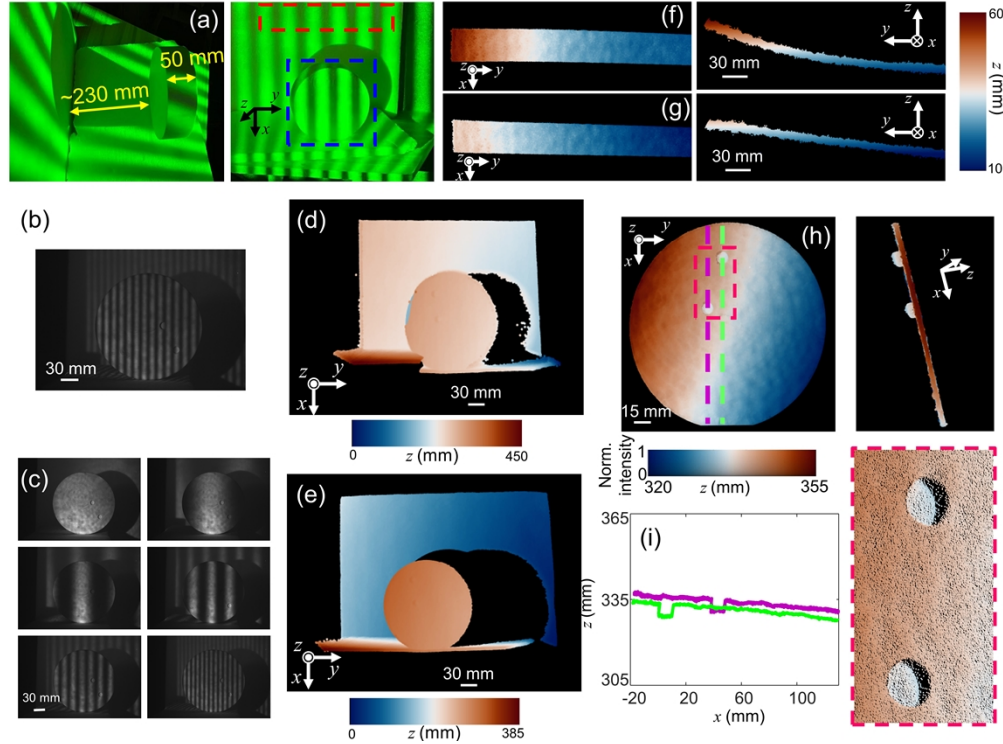
To demonstrate MS-BLIP's ability to tolerate a large variety in reflectance, we imaged an open box that had a highly reflective matte surface with a weakly reflective foam surface [Fig. 2(a)]. The representative fringe masks with two intensity levels are displayed in Fig. 2(b). The effect of different intensity levels in deformed structure images is compared in Fig. 2(c). The high-intensity projection (i.e., $I_{km}^{1.0}$) results in a large number of saturated pixels in the left region of the matte surface. The low-intensity projection (i.e., $I_{km}^{0.3}$), despite eliminating saturation, fails to resolve the white foam surface and the right bottom part of the matte surface. The dynamic ranges of $I_{km}^{1.0}$ and $I_{km}^{0.3}$ were calculated to be 48.1 dB and 41.8 dB, respectively. In comparison, dual-level intensity projection preserves low-intensity pixels while avoiding any saturation [Fig. 2(d)]. With an enhanced dynamic range of 58.6 dB, this method fully recovers this 3D object.

Fig. 2. MS-BLIP of a 3D object with varying reflectance. (a) Photo of the 3D object. (b) Representative binary fringe masks with the normalized intensity levels of 1.0 (left) and 0.3 (right). Red dashed boxes: close-up views. (c) Representatives fringe images by using high-intensity projection (left), low-intensity projection (middle), and their combination (right). (d) As (c), but showing the 3D reconstruction results in the three cases.

MS-BLIP's high accuracy in imaging 3D objects with depth discontinuities is proven by comparing the multi-frequency phase unwrapping method applied in MS-BLIP with the standard single-frequency counterpart. As shown in Fig. 3(a), a 50-mm-long cylinder with a 15-cm-diameter flat front surface stands on a base. A square foam plate is placed ~230 mm behind the cylinder. The single-frequency method used four phase-shifting fringe patterns with a period of

36 pixels (i.e., P_p^5) and an additional pattern of the DMD's vertical central line. Figures 3(b)–(c) show the captured deformed structure images by using the two methods. Because the distance between the cylinder and the foam plate resulted in a shift of multiple periods of fringes in deformed structure images, single-frequency fringe projection fails to correctly resolve the depth in the left half of the cylinder [Fig. 3(d)]. In contrast, displayed in Fig. 3(e), the multi-frequency method, adopted in MS-BLIP, accurately recovers the detailed 3D information. Specifically, the calculated depth between the cylinder's front surface to the foam plate is 281.7 mm, which well matches the ground truth (with a deviation of 0.61%).

Fig. 3. MS-BLIP of 3D objects with depth discontinuity. (a) Side view (left) and front view (right) of the 3D objects. (b) Representative deformed structure image by using single-frequency fringe projection. (c) As (b), but using multi-frequency fringe projection. (d)–(e) 3D reconstruction from single-frequency (d) and multi-frequency (e) fringe projection. (f)–(g) 3D reconstruction without (f) and with (g) distortion compensation of a selected region on the foam plate, indicated by the red dashed box in (a). Front view (left) and side view (right) are displayed. (h) Front view (left) and side view (right) of the reconstructed flat round surface, indicated as the blue dashed box in (a). Red dashed box: Close-up view of the two dents. (i) Selected depth profiles [marked by the purple and green dashed lines in (h)].

To demonstrate the necessity of distortion compensation, we analyzed the 3D reconstruction results of a selected region on the foam plate [red dashed box in Fig. 3(a)]. As shown in Fig. 3(f), the system distortion mistakenly guides the 3D reconstruction, which displays a curved structure of the reconstructed surface. With the iterative method for distortion compensation applied in MS-BLIP, the surface with correct depth information is presented in Fig. 3(g).

MS-BLIP can accurately resolve 3D information of fine structures. Figure 3(h) shows the 3D reconstruction of the cylinder's flat round surface with two dents [blue dashed box in Fig. 3(a)]. The quantitative analysis of selected line profiles shows that the two dents on the flat surface

have a depth of 6.0 mm and a diameter of 10.0 mm [Fig. 3(i)]. The deviations from the ground truth are 0.3 mm and 0.5 mm, respectively, which show an excellent agreement with the object's configuration.

3.2. MS-BLIP of translational movements

To highlight the potential application of MS-BLIP in automated detection and classification, we imaged an engineered box moving translationally at a speed of ~15 cm/s [Fig. 4(a)]. The box used a piece of curved cardboard as the base, on which pipes made of different materials were fixed. On the top row, the left pipe was made of black foam with low reflectance, while the other three metal pipes, as well as the white swelling glue, had high reflectance. MS-BLIP had a 3D imaging speed of 10.4 fps, an FOV of 0.3 m × 0.2 m at a working distance of 0.5 m, a depth resolution of 1.3 mm, and a lateral resolution of 1.1 mm.

Figures 4(b)–(d) show the advantage of dual-level intensity projection. For the $I_{km}^{1.0}$ projection, because saturated intensity distorted the pixel-wised sinusoidal intensity profile, fringe residues show up as artifacts on the regions with high reflectance [Fig. 4(b)]. Meanwhile, due to the lack of scattered light intensity from the $I_{km}^{0.3}$ projection, part of the black foam pipe was failed to be reconstructed [Fig. 4(c)]. In contrast, the dual-level intensity projection enhanced the dynamic range by over 12 dB. Presented in Fig. 4(d), the dual-level intensity projection successfully reconstructs the full shape of this 3D object.

Figure 4(e) shows the reconstructed 3D images of the moving engineered box at four time-lapse frames with an interval of 960 ms (see the full evolution in Visualization 1). MS-BLIP allowed tracking the 3D movements of all spatial points. As an example, the time-resolved positions of two points, selected from the pipes with both low and high reflectance [marked by P_A and P_B in Fig. 4(e)], are shown in Fig. 4(f). The results reveal that the two points have linear movement in the y axis at 15.4 cm/s but stay relatively stationary in the z axis, well matching the pre-set experimental condition.
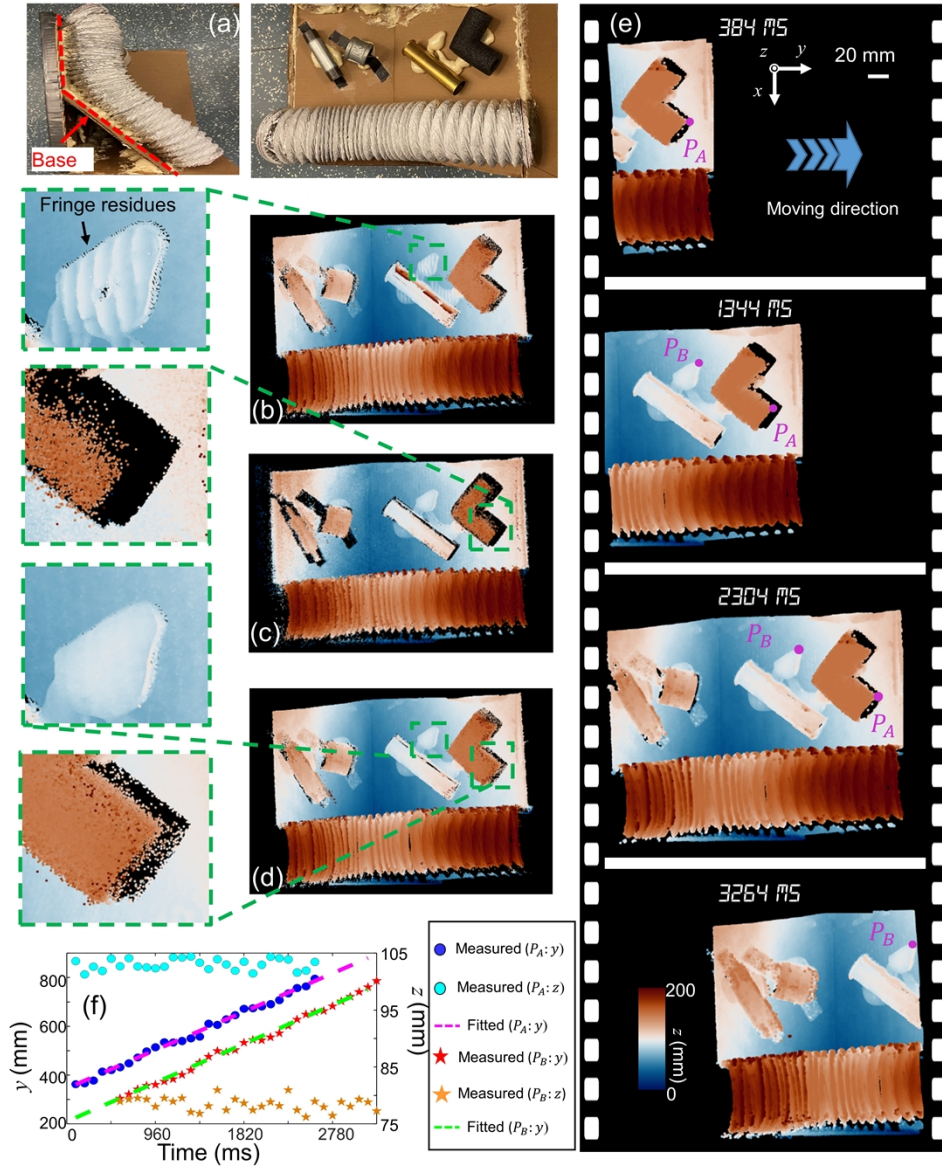
3.3. MS-BLIP of rotational movement

To demonstrate MS-BLIP's potential in industrial inspection, we imaged the rotational movement of a bamboo vase with extending branches [Fig. 5(a)]. With a height of 1.3 m, this 3D object was glued on a stand rotating at 0.6 rad/s. To fit the projection region with the desired imaging area, the dove prism was placed at 45°, which results in a 90° rotation of the projected patterns with respect to the fringe masks loaded onto the DMD. MS-BLIP was operated at a working distance of 2 m, with an FOV of 1.5 m × 1.0 m, and at a 3D imaging speed of 20.8 fps. Under these working conditions, the depth resolution was quantified to be 3.7 mm, and the lateral resolution was measured to be 1.7 mm.
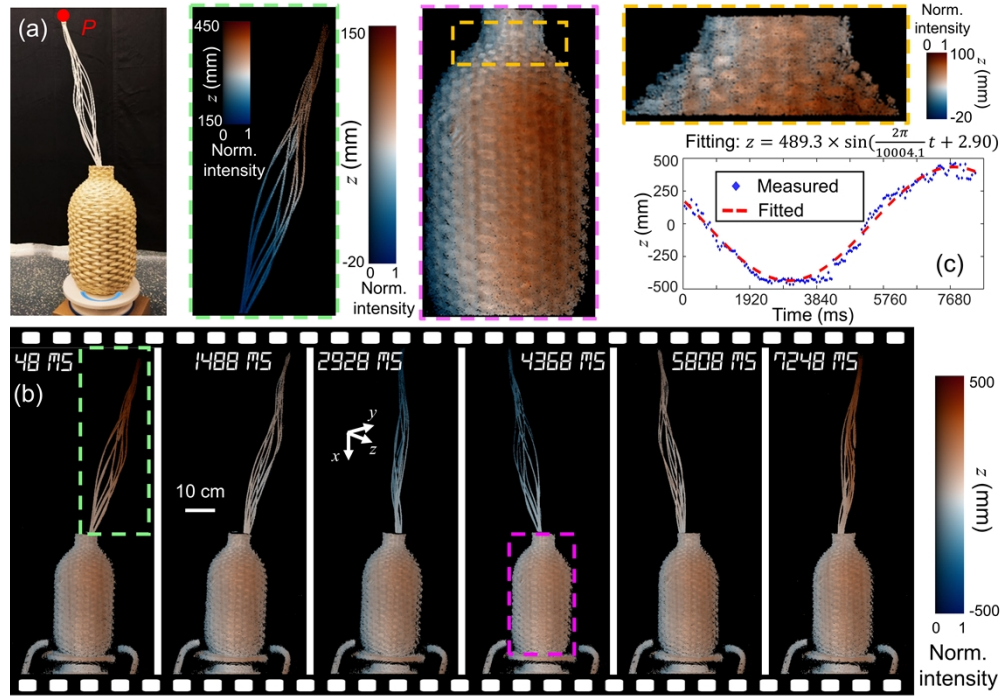
Figure 5(b) shows six time-lapse 3D images of the object (the full sequence is shown in Visualization 2). The close-up view of the vase mouth presents detailed structural information on its surface. The depth-encoded color change of the branches reflects the rotation movement of the object. By tracking the 3D position of the tip of the branches [marked by P in Fig. 5(a)], MS-BLIP reveals its rotational radius of 489.3 mm and the period of 10004.1 ms [Fig. 5(c)], showing a uniform rotation with the pre-set speed.

3.4. MS-BLIP of human-body movement

To explore the potential of MS-BLIP in human-computer interaction, we imaged the full-body movements of a volunteer (the full sequence is shown in Visualization 3). All the experiments were conducted in accordance with the protocol (Project CÉR-22-649) approved by the human ethics research committee of Institut National de la Recherche Scientifique, Université du Québec. A photo of the scene is shown in Fig. 6(a). The volunteer had a height of ~1.6 m, wearing proper protective clothing and a laser goggle. Both hands were exposed to the laser illumination with

Fig. 4. MS-BLIP of translational movement of an engineered box. (a) Photo of the 3D object. (b)–(d) Reconstructed 3D images with high-intensity projection (b), low-intensity projection (c), and their combination (d). Close-up views are shown on the left column. (e) Time-lapse 3D images of the translationally moving engineered box. (f) Traces of two points [marked in (e)] with fitting.
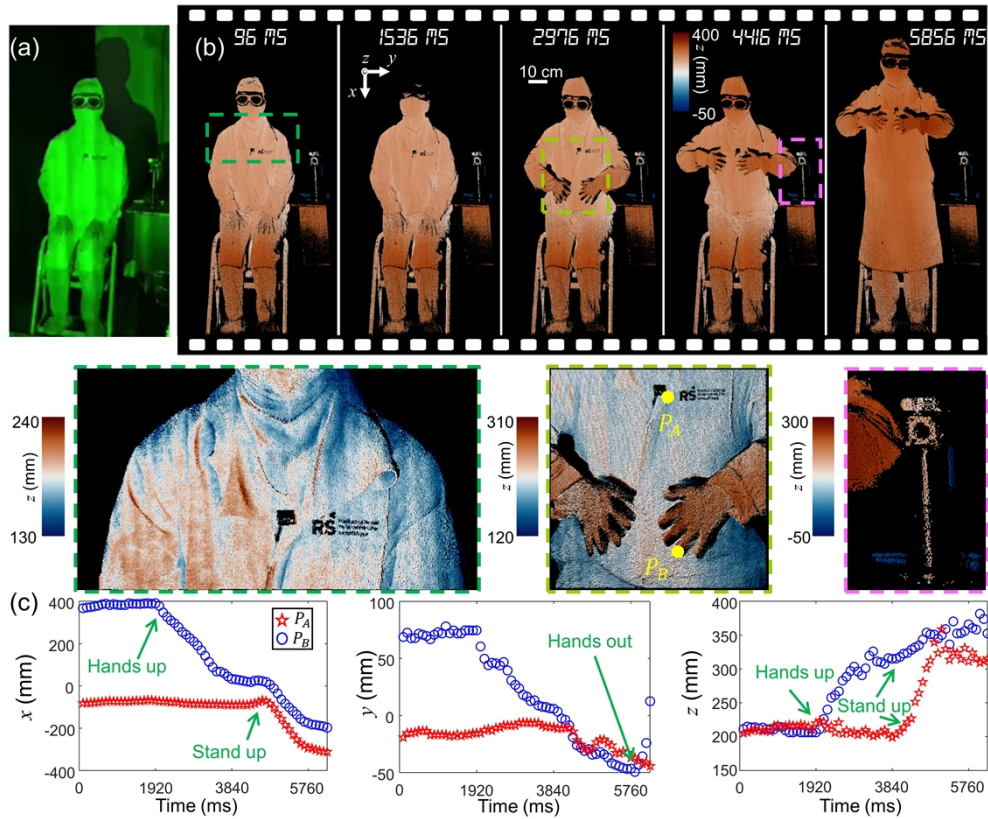
Fig. 5. MS-BLIP of rotational movement of a vase with branches. (a) Photo of the 3D object. (b) Time-lapse 3D images of the object. Close-up views are shown in the dashed boxes. (c) Evolution of the depth change of a selected point marked in (a) with sinusoidal fitting.

intensity (and fluence) of 7.8×10^{-7} W/cm^2 (and 3.9×10^{-11} J/cm^2), which is much lower than the maximum permissible exposure (i.e., 0.2 W/cm^2 and 2×10^{-2} J/cm^2) regulated by the ANSI safety limit [45]. A part of an optical table on the left of this volunteer was also included in the FOV. Akin to the 3D imaging of rotational movement, the dove prism was set to 45°. The MS-BLIP system was operated with a working distance of 2.8 m and an FOV of 1.7 m × 1.1 m. It had a 3D imaging speed of 10.4 fps. The depth and the lateral resolutions were measured to be 4.7 mm and 2.4 mm, respectively.

Figure 6(b) shows five time-lapse 3D images of the instantaneous poses of the volunteer. The detailed surface structures of the volunteer's lab coat and hands are illustrated by the close-up views. The edge of an optical table and optomechanical components can also be seen as a static background. As displayed in Fig. 6(c), we tracked the 3D positions of two selected points over time: the ring finger's tip of the volunteer's left hand and the edge point of a marker on the volunteer's coat. In the experiment, the 3D image acquisition began when the volunteer sat on a stepladder with both hands on the knees. Afterward, the volunteer moved up both hands at 2016 ms, resulting in a change of the fingertip position along all three axes [marked by the arrows for "Hands up" in Fig. 6(c)]. The volunteer's body stayed at the same position until 4128 ms when the volunteer stood up, which results in movements along the -*x* axis for both points [marked by the arrow for "Stand up" in Fig. 6(c)]. In addition, the forward incline movement of the volunteer's body while standing up was also reviewed according to the position changes along the +*z* axes. At 6048 ms, the volunteer stretched out both hands, which is responsible for the sudden increase of the *y* value [marked by the arrow for "Hands out" in Fig. 6(c)].

Fig. 6. MS-BLIP of full human-body movements. (a) Photo of the 3D object. (b) Five time-lapse reconstructed 3D images. Close-up views are shown in the dashed boxes. (c) Evolution of the 3D positions of the selected points [marked in the middle close-up view of (b)].

4. Discussion and conclusions

We have developed MS-BLIP for robust multi-scale 3D surface imaging at video rate. Compared to existing BLIP systems, MS-BLIP employs dual-level intensity projection to tolerate the variation in 3D objects' reflectance. It implements the multi-frequency fringe projection method to robustly resolve spatially isolated 3D objects with depth discontinuity. Finally, it uses an iterative method for distortion compensation in 3D point recovery to improve the quality of reconstruction. These supports have empowered MS-BLIP for dynamic 3D visualization of translational movements of an engineered box, rotational movements of a craft vase, and full human body movements at an imaging speed of up to 20.8 fps and a measurement volume of up to 1.5 m^3—a three order of magnitude increase compared to existing BLIP systems. Specifications used for main demonstrations are summarized in Table 1.

MS-BLIP contributes several conceptual advantages to PSFPP. First, its illumination scheme could be readily adapted to other multi-frequency PSFPP systems. Holding a strict imaging relationship between the DMD and the 3D objects, MS-BLIP preserves high accuracy and high contrast of grayscale fringe patterns at different frequencies. Despite being demonstrated only with six selected frequencies, MS-BLIP can be easily adapted to accommodate varying frequencies, arbitrary sets of patterns, and different working distances and FOVs for specific studies. Second, MS-BLIP employs a nanosecond pulsed laser as the illumination source. Since

Table 1. Major specifications of MS-BLIP in dynamic 3D imaging experiments

Application	Working distance (m)	(x, y) FOV (m^2)	Projection/acquisition speed (Hz)	Camera exposure time (ms)	3D imaging speed(fps)	Depth resolution (mm)
Translating engineered box	0.5	0.3×0.2	500	2	10.4	1.3
Rotating vase	2.0	1.0×1.5	1000	1	20.8	3.7
Full human body	2.8	1.1×1.7	500	2	10.4	4.7

the laser pulse width is much shorter than the DMD's display time, no photons would be wasted during pattern switching, which is otherwise unavoidable for a continuous-wave laser. This advantage also sheds light on implementing MS-BLIP using high-speed cameras [46,47] to increase 3D imaging speeds while maintaining signal-to-noise ratios. Third, a dove prism is used to adjust the orientation of the FOV. In existing PSFPP techniques, the aspect ratio of the projected area is fixed by the DMD's shape. Changing the aspect ratio limits the DMD's active region, which, however, reduces the overall light throughput. The dove prism increases the flexibility in the orientation of the projection region, which thus helps MS-BLIP fit different experimental scenarios.

The future work will be carried out in the following aspects. First, we could further improve MS-BLIP's imaging speed by adopting multiple cameras, a faster DMD, and a more powerful light source. Moreover, we plan to implement online feedback to adaptively adjust the intensity of projected patterns [48], which would optimize MS-BLIP's dynamic range for different working conditions. Furthermore, to take into account possible measurement inaccuracy induced by laser speckles, we could implement a superluminescent diode and a rotating diffuser in the MS-BLIP system as well as apply a filtering algorithm in image reconstruction [49]. Finally, a graphic processing unit could be used to accelerate the computation of phase maps for real-time display of the 3D reconstruction results. Besides technical improvement, we will continue exploring new applications of MS-BLIP, including automated industrial inspection [50], archaeology [51], and human-computer interaction [52].

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. C. Zuo, T. Tao, S. Feng, L. Huang, A. Asundi, and Q. Chen, "Micro Fourier Transform Profilometry (μFTP): 3D shape measurement at 10,000 frames per second," Optics and Lasers in Engineering **102**, 70–91 (2018).
2. J.-S. Hyun and S. Zhang, "Superfast 3D absolute shape measurement using five binary patterns," Optics and Lasers in Engineering **90**, 217–224 (2017).
3. W. Yu and B. Xu, "A portable stereo vision system for whole body surface imaging," Image and Vision Computing **28**(4), 605–613 (2010).
4. M.n. Gilles, H. Fayad, P. Miglierini, J.-F. Clément, S. Scheib, L. Cozzi, J. Bert, N. Boussion, U. Schick, and O. Pradier, "Patient positioning in radiotherapy based on surface imaging using time of flight cameras," Med. Phys. **43**(8Part1), 4833–4841 (2016).

5. K. Zhong, Z. Li, X. Zhou, Y. Li, Y. Shi, and C. Wang, "Enhanced phase measurement profilometry for industrial 3D inspection automation," Int J Adv Manuf Technol **76**(9-12), 1563–1574 (2015).

6. C. Jiang, S. Jia, Y. Xu, Q. Bao, J. Dong, and Q. Lian, "The application of multi-frequency fringe projection profilometry on the measurement of biological tissues," Bio-Med. Mater. Eng. **26**(s1), S395–S403 (2015).

7. B. Li and S. Zhang, "Superfast high-resolution absolute 3D recovery of a stabilized flapping flight process," Opt. Express **25**(22), 27270–27282 (2017).

8. Z. Wu, W. Guo, B. Pan, Q. Kemao, and Q. Zhang, "A DIC-assisted fringe projection profilometry for high-speed 3D shape, displacement and deformation measurement of textured surfaces," Optics and Lasers in Engineering **142**, 106614 (2021).

9. J. Qian, S. Feng, T. Tao, Y. Hu, K. Liu, S. Wu, Q. Chen, and C. Zuo, "High-resolution real-time 360° 3D model reconstruction of a handheld object with fringe projection profilometry," Opt. Lett. **44**(23), 5751–5754 (2019).

10. A. Jha, "Class room attendance system using facial recognition system," The International Journal of Mathematics, Science, Technology and Management **2**, 4–7 (2007).

11. S. Alben and M. J. Shelley, "Flapping states of a flag in an inviscid fluid: bistability and the transition to chaos," Phys. Rev. Lett. **100**(7), 074301 (2008).

12. M. Mazzetto, L. F. Southier, M. Teixeira, and D. Casanova, "Automatic classification of multiple objects in automotive assembly line," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, (IEEE, 2019), 363–369.

13. D. Kosmopoulos and T. Varvarigou, "Automated inspection of gaps on the automobile production line through stereo vision and specular reflection," Computers in Industry **46**(1), 49–63 (2001).

14. Y. Cui, W. Chang, T. Nöll, and D. Stricker, "Kinectavatar: fully automatic body capture using a single kinect," in *Asian Conference on Computer Vision*, (Springer, 2012), 133–147.

15. S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," Optics and Lasers in Engineering **48**(2), 149–158 (2010).

16. J. Geng, "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photonics **3**(2), 128–160 (2011).

17. I. Ishii, K. Yamamoto, K. Doi, and T. Tsuji, "High-speed 3D image acquisition using coded structured light projection," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (IEEE, 2007), 925–930.

18. O. Hall-Holt and S. Rusinkiewicz, "Stripe boundary codes for real-time structured-light range scanning of moving objects," in *Proceedings Eighth IEEE International Conference on Computer Vision*. ICCV 2001, (IEEE, 2001), 359–366.

19. P. S. Huang, S. Zhang, and F.-P. Chiang, "Trapezoidal phase-shifting method for three-dimensional shape measurement," Opt. Eng. **44**(12), 123601 (2005).

20. P. Jia, J. Kofman, and C. E. English, "Two-step triangular-pattern phase-shifting method for three-dimensional object-shape measurement," Opt. Eng. **46**(8), 083201 (2007).

21. S. Zhang, "High-speed 3D shape measurement with structured light methods: A review," Optics and Lasers in Engineering **106**, 119–131 (2018).

22. 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**(10), 108201 (2012).

23. J. Liang, "Punching holes in light: recent progress in single-shot coded-aperture optical imaging," Rep. Prog. Phys. **83**(11), 116101 (2020).

24. Z. Song, *Handbook of 3D machine vision : optical metrology and imaging* (CRC press, 2013), Chap. 9.

25. J.-S. Hyun, B. Li, and S. Zhang, "High-speed high-accuracy three-dimensional shape measurement using digital binary defocusing method versus sinusoidal method," Opt. Eng. **56**(7), 074102 (2017).

26. B. Li, Y. Wang, J. Dai, W. Lohry, and S. Zhang, "Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques," Optics and Lasers in Engineering **54**, 236–246 (2014).

27. 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**(4), 964–967 (2020).

28. C. Jiang, P. Kilcullen, Y. Lai, T. Ozaki, and J. Liang, "High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition," Photonics Res. **8**(11), 1808–1817 (2020).

29. V. Parthiban, R. N. Kohn Jr, J. Liang, and M. F. Becker, "Experimental demonstration of precise holograms using complex light modulation," in *Emerging Digital Micromirror Device Based Systems and Applications VIII*, (International Society for Optics and Photonics 2016) 97610M.

30. Y. Wang, H. Zhao, H. Jiang, and X. Li, "Defocusing parameter selection strategies based on PSF measurement for square-binary defocusing fringe projection profilometry," Opt. Express **26**(16), 20351–20367 (2018).

31. X. Li, M. Gong, Z. Zhang, Z. Zhang, and C. Yang, "Calibration method for fringe projection profilometry with a binary defocusing technique," Opt. Rev. **24**(4), 495–504 (2017).

32. C. Zuo, L. Huang, M. Zhang, Q. Chen, and A. Asundi, "Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review," Optics and Lasers in Engineering **85**, 84–103 (2016).

33. C. Jiang, P. Kilcullen, X. Liu, Y. Lai, T. Ozaki, and J. Liang, "High-speed three-dimensional surface measurement using band-limited illumination profilometry (BLIP)," in *Emerging Digital Micromirror Device Based Systems and Applications XIII*, (International Society for Optics and Photonics 116980 (2021). V.

34. 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**(26), 32013–32018 (2013).

35. D. L. Lau and G. R. Arce, *Modern digital halftoning* (CRC Press, 2018).

36. J. Heikkila and O. Silvén, "A four-step camera calibration procedure with implicit image correction," in *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, (IEEE, 1106–1112 (1997).

37. J.-Y. Bouguet, "Camera Calibration Toolbox for Matlab," retrieved http://www.vision.caltech.edu/bouguetj/calib_doc/

38. S. Zhang and P. S. Huang, "Novel method for structured light system calibration," Opt. Eng. **45**(8), 083601 (2006).

39. S. Zhang and S.-T. Yau, "High dynamic range scanning technique," Opt. Eng. **48**(3), 033604 (2009).

40. J. Tian, X. Peng, and X. Zhao, "A generalized temporal phase unwrapping algorithm for three-dimensional profilometry," Optics and Lasers in Engineering **46**(4), 336–342 (2008).

41. D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," J. Opt. Soc. Am. A **11**(1), 107–117 (1994).

42. H. O. Saldner and J. M. Huntley, "Temporal phase unwrapping: application to surface profiling of discontinuous objects," Appl. Opt. **36**(13), 2770–2775 (1997).

43. R. Burden, J. D. Faires, and A. Reynolds, "Numerical analysis, brooks/cole," Boston, Mass, USA (2010).

44. J. Liang, Y. Zhou, A. W. Winkler, L. Wang, K. I. Maslov, C. Li, and L. V. Wang, "Random-access optical-resolution photoacoustic microscopy using a digital micromirror device," Opt. Lett. **38**(15), 2683–2686 (2013).

45. "ANSI Z136.1-2014 Laser Safety Standard for the Safe Use of Lasers," Laser Institute of America, Orlando, Florida (2014).

46. X. Liu, A. Skripka, Y. Lai, C. Jiang, J. Liu, F. Vetrone, and J. Liang, "Fast wide-field upconversion luminescence lifetime thermometry enabled by single-shot compressed ultrahigh-speed imaging," Nat. Commun. **12**(1), 6401 (2021).

47. J. Liang and L. V. Wang, "Single-shot ultrafast optical imaging," Optica **5**(9), 1113–1127 (2018).

48. C. Chen, N. Gao, X. Wang, and Z. Zhang, "Adaptive pixel-to-pixel projection intensity adjustment for measuring a shiny surface using orthogonal color fringe pattern projection," Meas. Sci. Technol. **29**(5), 055203 (2018).

49. H. Liu, G. Lu, S. Wu, S. Yin, and T. Francis, "Speckle-induced phase error in laser-based phase-shifting projected fringe profilometry," J. Opt. Soc. Am. A **16**(6), 1484–1495 (1999).

50. E. W. Y. So, M. Munaro, S. Michieletto, M. Antonello, and E. Menegatti, "Real-time 3d model reconstruction with a dual-laser triangulation system for assembly line completeness inspection," in *Intelligent Autonomous Systems* 12 (Springer, 2013), pp. 707–716.

51. F. Menna, E. Nocerino, and A. Scamardella, "Reverse engineering and 3D modelling for digital documentation of maritime heritage," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38**, W16 (2011).

52. L. Zhu, K. Rematas, B. Curless, S. M. Seitz, and I. Kemelmacher-Shlizerman, "Reconstructing NBA players," in *European Conference on Computer Vision*, (Springer, 177–194 (2020).