Optics and Lasers in Engineering 102 (2018) 70–91

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Micro Fourier Transform Profilometry (μFTP): 3D shape measurement at 10,000 frames per second

Chao Zuo[a,b,*], Tianyang Tao[a,b], Shijie Feng[a,b], Lei Huang[c], Anand Asundi[d], Qian Chen[b]

[a] Smart Computational Imaging (SCI) Laboratory, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China
[b] Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China
[c] Brookhaven National Laboratory, NSLS II 50 Rutherford Drive, Upton, NY 11973-5000, United States
[d] Centre for Optical and Laser Engineering (COLE), School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore

ARTICLE INFO

Keywords:
Fringe projection profilometry
Micro Fourier Transform Profilometry (μFTP)
3D shape measurement
Phase unwrapping

ABSTRACT

Fringe projection profilometry is a well-established technique for optical 3D shape measurement. However, in many applications, it is desirable to make 3D measurements at very high speed, especially with fast moving or shape changing objects. In this work, we demonstrate a new 3D dynamic imaging technique, Micro Fourier Transform Profilometry (μFTP), which can realize an acquisition rate up to 10,000 3D frame per second (fps). The high measurement speed is achieved by the number of patterns reduction as well as high-speed fringe projection hardware. In order to capture 3D information in such a short period of time, we focus on the improvement of the phase recovery, phase unwrapping, and error compensation algorithms, allowing to reconstruct an accurate, unambiguous, and distortion-free 3D point cloud with every two projected patterns. We also develop a high-frame-rate fringe projection hardware by pairing a high-speed camera and a DLP projector, enabling binary pattern switching and precisely synchronized image capture at a frame rate up to 20,000 fps. Based on this system, we demonstrate high-quality textured 3D imaging of 4 transient scenes: vibrating cantilevers, rotating fan blades, flying bullet, and bursting balloon, which were previously difficult or even unable to be captured with conventional approaches.

1. Introduction

The desire to capture and record fast phenomena can be traced back to invention of film photography in the nineteenth century, which used controllable mechanical light shutters to limit the film exposure time [1–3]. However, it is only with the advances made during the last few decades in the field of solid-state electronic imaging sensors based on the charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) technology that high-speed imaging has gained substantial interests and applications [4,5]. Investigation into the scientific origins of fast phenomena has benefited enormously from the development of such high-speed cameras [6,7], and other applications exist in almost all areas of industry, television, defense, sports, and health care. More recently, ultra-high-speed imaging at picosecond (10^{-12} s) timescales has been demonstrated by either using ultra-short pulse illumination to provide temporal resolution [8,9], or combining streak imaging technology with scanning [10] or temporal pixel coding strategy [11] to achieve two-dimensional (2D) information, enabling a frame rate fast enough to visualize photons in motion.

Despite these tremendous advances, most of the current high-speed imaging sensors can only record instantaneous phenomena as 2D image sequences that lack the depth information. Nevertheless, high-speed three-dimensional (3D) reconstruction of transient scenes has been long sought in a host of fields that include biomechanics [12], industrial inspection [13], solid mechanics [14], and vehicle impact testing [15]. In these areas, it is always desirable that the 3D information can be acquired at a high frame rate so that transient geometric changes of an object or an environment can be captured. These can then be reviewed in slow motion to provide in-depth insights into fast changing events in a broad range of application scenarios.

Over the past decades, 3D image acquisition technologies have also rapidly evolved, benefiting from the advances in electronic imaging sensors, optical engineering, and computer vision. For a small scale depth or shape, micrometer or even nanometer measurements can be reached by using interferometry [16], confocal [17] or other depth-resolved 3D microscopic techniques [18,19]. Time-of-flight methods [20,21], in which the distance is resolved by measuring the time of flight of a light pulse with the known speed of light, are well suited for measuring large-scale scenes or depth ($>$ 0.5 m). For the 3D measurement of medium-size

* Corresponding author.
E-mail addresses: zuochao@njust.edu.cn, surpasszuo@163.com (C. Zuo), chenqian@njust.edu.cn (Q. Chen).

https://doi.org/10.1016/j.optlaseng.2017.10.013
Received 14 September 2017; Received in revised form 15 October 2017; Accepted 16 October 2017
0143-8166/

objects, full field triangulation based active structured light (SL), has proven to be one of the most promising techniques [22–24]. Just as human vision system, the basic principle of SL is optical triangulation, in which correspondences between a projector and a camera are established with some coded patterns projected onto the scene. Recent years, numerous SL approaches have been proposed and investigated, and there is a clear trend towards improving the measurement speed to video rates ($\sim$ 25 Hz) and far beyond [25–27]. This trend is being driven by the increasing demand for high-speed depth data coupled with rapid advances in high-frame-rate image sensors and digital projection technology. Today's high-speed cameras can capture video at speeds up to tens of thousands full frames per second (fps) and even faster at a reduced resolution. On the other hand, the digital micro-mirror device (DMD), as a key component of digital light processing (DLP) projection system, is able to create, store, and display high-speed binary pattern sequences through optical switching ("on" and "off") at rates in excess of 10,000 Hz as well. By operating the DMD in binary (1-bit) mode, quasi-sinusoidal fringe patterns can be created at the maximum frame rates with lens defocusing and binary dithering techniques [28–30].

Once the speed of hardware is no longer a limiting factor, the main hurdle to overcome in high-speed 3D sensing is to reduce the number of patterns required for reliable 3D reconstruction. Single-shot approaches [31,32], e.g., de Bruijin sequences [32,33], M-arrays [34], and symbol coded patterns [35], are well-suited for dynamic 3D sensing. Generally, the pattern needs to be wisely designed so that each point can be uniquely identified from its neighboring pixels. Nevertheless, the spatial resolution and depth accuracy of single-shot approaches are limited due to the local depth smoothness assumption, which does not hold around the regions with depth discontinuities and fine details. Fourier transform profilometry (FTP) [25,36–38] is another representative single-shot approach in which the phase is extracted from a single high-frequency fringe image by applying a properly designed band-pass filter in the frequency domain. Compared with other single-shot methods, FTP has the advantages of pixel-wise measurement and effective noise removal, yet with the precondition that the fundamental frequency component, which carries the phase information of the object, is separable from zero-order background. In practice, this precondition can be easily violated when the measured surface contains sharp edges, discontinuities, and/or large surface reflectivity variations [37,39–42]. Several methods have been proposed to alleviate this problem by carefully designing the band-pass filter [40,43], or capturing an additional flat [41] or π-phase shifted fringe pattern [39]. However, for objects with complex surface properties, the measurement accuracy of these FTP approaches is still quite limited due to the inherent spectrum leakage.

To achieve high-quality dense 3D reconstructions, the multi-frame methods [44–46] are usually preferred for their advantages of high accuracy and low complexity. In general, multi-frame methods require several predetermined patterns to be projected onto the measured object. Hence, their measurement speed is limited by the number of patterns per sequence and both camera and projector frame rate. The most widely used multi-frame approach is phase shifting profilometry (PSP) [47], which requires a minimum of three fringe images to provide high-accuracy pixel-wise phase measurement. Furthermore, the measurement is quite robust to ambient illumination and varying surface properties. However, when measuring dynamic scenes, motion will lead to phase distortion artifacts, especially when the object motion during the inter-frame time gap is non-negligible [42,46,48–50]. This is an intrinsic problem of PSP, since the phase information is spread over multiple fringe images. Another challenging problem in both PSP and FTP is the phase ambiguity resulting from the periodical nature of the sinusoidal signal. Though high-frequency patterns with dense fringes are usually preferred for high-accuracy 3D reconstruction (especially for FTP), they also introduce severe ambiguities. To recover the absolute phase, a common practice is to use temporal phase unwrapping (TPU) algorithms with the help of Gray-code patterns [51] or multi-wavelength fringes [52]. However, this requires a large number of additional images, which are only used for phase disambiguation purpose but not contribute to the final 3D results. The prolonged pattern sequence (e.g., a minimum of 9 patterns are required per 3D reconstruction for three-wavelength PSP [53]) greatly limits the performance of PSP and FTP in high-speed, time-critical scenarios.

To address this limitation, several composite phase shifting schemes [54–59] (e.g., dual-frequency PSP [54], bi-frequency PSP [59], 2 + 2 PSP [55], and period coded PSP [56,57]) have been proposed. They can solve the phase ambiguity problem without significantly increasing the number of projected patterns. However, in order to guarantee a reasonable reliability of phase unwrapping, the fringe frequency often cannot be too high, which provides only a comparatively low accuracy [60]. Alternatively, stereo unwrapping methods based on geometric constraint can be used to determine the fringe order without capturing additional images [46,58,61–63]. But the measured objects must be within a restricted depth of the measurement volume. Moreover, some of these techniques require multiple (≥ 2) high-speed cameras, which could considerably increase the overall cost of the system. On a different note, high-speed 3D sensing can also be realized via active stereo-photogrammetry without explicit evaluation of phase information [64–68]. In such approaches, the projected patterns are merely used to establish precise camera-to-camera correspondences based on correlation algorithms, so various types of structured patterns (e.g., statistical speckle [64] and aperiodic sinusoid [69]) can be applied. This idea allows to develop alternative projection units without the need of a DMD, yielding simpler optical designs and/or much higher projection rates. For example, the LED-based array projection system and the GOBO (GOes Before Optics) projection system enable high-speed pattern switching with frame rates up to tens of kHz [66,67], which is much higher than the maximum speed of commercial DLP projectors (typically 180–360 Hz after disassembling the color wheel [45,55,70]). However, these techniques still require at least two high-speed cameras. Moreover, as stated by Grosse et.al. [71], more than 9 images are typically required to establish dense, accurate, and outlier-free correspondences. Thus, the final 3D frame rate achievable is still much lower (almost an order of magnitude) than the native frame rate of the camera and projector, which is typically in the range of only several hundreds Hz (e.g., 330 Hz at a projection frame rate of 3 kHz by using array projection [66] and 1,333 Hz at a camera rate of 12,000 Hz by using GOBO projection [67]).

In this work, we present *Micro Fourier Transform Profilometry* (μFTP), which enables highly-accurate dense 3D shape reconstruction at 10,000 Hz, without posing restrictions on surface texture, scene complexity, and object motion. In contrast to conventional FTP which uses a single high-frequency sinusoidal pattern with a fixed spatial frequency, μFTP introduces very small temporal variations in the frequency of multiple spatial sinusoids to eliminate phase ambiguity. So in μFTP, the word *Micro* just refers to the small values for both the frequency variations and periods of fringe patterns. This is similar to that of *Micro* Phase Shifting [72], in which band-limited high-frequency phase-shifted fringes are used to reduce measurement errors due to global illumination. The key idea of μFTP is to freeze the high-speed motion by encoding the phase information within a single high-frequency sinusoidal pattern. And the phase ambiguities are resolved spatio-temporally with the extra information from the small frequency variations along the temporal dimension. Besides, high-quality 2D texture can be acquired by inserting additional white patterns (all mirrors of the DMD are in the "on" state) between each high-frequency sinusoidal patterns, which also remove spectrum overlapping and enable high-accuracy phase measurements even in the presence of large surface reflectivity variations. Unlike previous approaches in which the phase retrieval and disambiguation were separately addressed in a pixel-wise and time-dependent fashion, the μFTP extends phase unwrapping into the space-time domain. The main contributions of this paper are two-fold:

(1) We propose and analyze a complete computational framework (based on a combination of improved phase recovery, phase unwrapping, error compensation, and system calibration algorithms) that al-

lows to effectively recover an accurate, unambiguous, and distortion-free 3D point cloud with every two projected patterns (Section **Principle of μFTP**). Each projected pattern serves the dual purpose of phase disambiguation and 3D reconstruction, which allows to minimize the number of fringe patterns for absolute depth recovery. The comprehensive theory, implementation, and demonstration of each algorithm involved in the μFTP framework are provided in details in **Appendices A-D**.

(2) We develop a high-frame-rate DLP fringe projection system by pairing a high-speed CMOS camera and a high-speed projection system based on a DLP development kit (Section **Experimental setup**). By applying the μFTP framework to our hardware system, we can achieve a 3D image acquisition rate of up to 10,000 fps, with a depth accuracy better than 80 μm and a temporal uncertainties below 75 μm within a measurement volume of 400 mm × 275 mm × 400 mm (Section **Quantitative analysis of 3D reconstruction accuracy**). In Section **Results**, we demonstrate for the first time high-quality textured 3D imaging of vibrating cantilevers, rotating fan blades, flying bullet, and bursting balloon, which were previously difficult or even unable to be captured with conventional approaches.

2. Materials and methods

2.1. Experimental setup

The high-frame-rate fringe projection system consists of a high-speed CMOS camera (Vision Research Phantom V611) and a high-speed projection system. The high-speed projection system includes a DLP development kit (Texas Instruments DLP Discovery 4100) with an XGA resolution (1024 × 768) DMD, and a custom-designed optics module. By omitting any grayscale capabilities, we drive the DMD at a refresh rate up to 20,000 binary fps. The light source is a green LED module with an output of 600 lm. The high-speed camera used in this system has a frame rate of 6246 fps for maximum image resolution (1280 × 800). A 24 mm (focal length) lens (Nikon AF–S) with a variable aperture from f/3.5 to f/5.6 is attached to the camera. The camera lens aperture is fully opened to allow the maximum amount of light to enter. In this work, the camera is operated at a reduced image resolution (640 × 440) to match the frame rate of the DMD (20,000 fps) with an exposure time of 46 μs. It is also precisely synchronized with the projection system with the help of the DLP development hardware, which will be described in details in the next subsection.

2.2. Phase unwrapping using geometry constraint

To realize high-frame-rate 3D shape reconstruction, the key technical issue regarding the hardware system is the precisely synchronized high-speed pattern projection and capture. Over the years, a number of fringe projection systems have been developed by re-engineering off-the-shelf DLP projectors for high-speed applications [44,45,55,70]. The common idea is to remove the color wheel of the DLP projector to make it work in the monochrome mode so that the projection rate can be tripled theoretically. For conventional 3-step PSP, one can simply make the projector display one static color image with three phase-shifted patterns encoded in its RGB channels respectively [44,45]. Once the projector receives the video signal, it will automatically decode the input color image and project the three phase-shifted patterns in each color channel sequentially. The situation becomes more complicated when the pattern sequence contains more than 3 images because the input color image needs to be switched sequentially and periodically at high speed. In order to handle the pattern switching, addition hardware, like Field-Programmable Gate Array (FPGA) circuits can be used to generate the desired video sequence with high precision and stability. Another possible solution is to directly access the memory of the graphical device based on CUDA programming [70]. Nevertheless, the maximum projection rate can be achieved is still around 180–360 Hz, which is ultimately limited by the projection mechanism of the DLP projector (the gray-scale

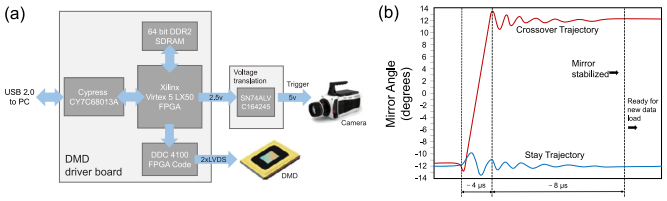

Fig. 1. (a) Schematic diagram of the high-frame-rate fringe projection system developed. (b) Mirror trajectory for the DMD after "reset" operation [73].

pattern is generated with the binary pulse width modulation along the time-axis, and the intensity level is reproduced by integrating smaller binary intensities over time by an image sensor, like eye or camera) and the use of standard (typically 60 Hz–120 Hz) graphic adapters for DMD control. Such bottlenecks in speed block a regular consumer DLP projector from many important applications where very fast motion needs to be acquired.

To overcome these current limitations, in this work, we use a DLP Discovery 4100 development kit from Texas Instruments as the basis of our projection system. As shown in the hardware diagram [Fig. 1(a)], the DLP kit includes a DMD chip and a DMD driver board. The DMD is based on the 0.7 XGA (1024 × 768) chipset with a mirror pitch of 13.6 μm (DLP7000, Texas Instruments). The driver board has a programmable FPGA (Xilinx Virtex 5 LX50) and a USB 2.0 interface (Cypress CY7C68013A) for receiving input patterns from the computer. The on-board DDR2 SDRAM memory is used to store the pattern sequences that are pre-loaded for subsequent high-speed projection. The FPGA connects with the DMD controller chip (DDC4100, Texas Instruments) by parallel interface and transfers pattern data to it directly. At the same time, DDC4100 provides a high-speed 2xLVDS data interface and controls the DMD mirrors to turn for generating a pattern according to the pattern data. Based on these specific hardware, DLP Discovery 4100 offers advanced micro-mirror control as well as flexible formatting and sequencing light patterns. It enables the DMD to operate in binary mode without any temporal dithering, allowing binary light patterns to be projected with speed far surpassing that of a regular DLP projector. However, there remain a few constraints underlying the basic operation of a DMD.

(1) There hardware requires a certain amount of time to transition from one micro-mirror configuration to another - a limitation that is imposed due to data transmission to the DMD. The data transmission bus between FPGA and DMD operates at 64 bits and at 400 MHz. Hence, for our DMD with 1024 × 768 micro-mirrors, it takes at least 30.72 μs (τ_{LD}) to load a full frame binary image

$$\tau_{LD}=\frac{1024\times 768}{64\times 400\times 10^{6}}=30.72\ \mu s \tag{1}$$

(2) When the data is loaded in memory, "reset" operation can be performed, which tilts the mirrors into their new "on" or "off" states. It requires some time to physically tilt the mirrors and some time for mirrors to settle down. As illustrated in Fig. 1(b), there are about 4 μs transition time (τ_{TT}) and 8 μs settling time (τ_{ST}) during the mirror state conversion process [73]. Besides, during the reset and mirror settling time, no data can be loaded in the DMD, thus slowing down the pattern output rate.

When both two factors are considered, the total time required to output one pattern on the mirrors is given in Eq. (2):

$$\tau=\tau_{LD}+\tau_{TT}+\tau_{ST}=30.72+4+8=42.72\ \mu s \tag{2}$$

so the highest possible rate for binary pattern output is 1/42.72 μs = 23408 fps. For the sake of simplicity and security, we update all micro-mirrors with a switching period of 50 μs, based on the timing shown in Fig. 2. With this timing configuration, we are able to project the binary patterns at 20,000 Hz. However, the synchronization between the DMD operation and the high-speed camera needs to be carefully designed in

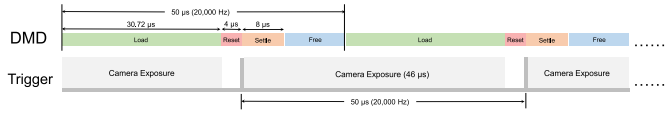

Fig. 2. Timing diagram of the designed hardware system. The exposure time of the camera is set at 46 μs to skip the mirror transition stage.

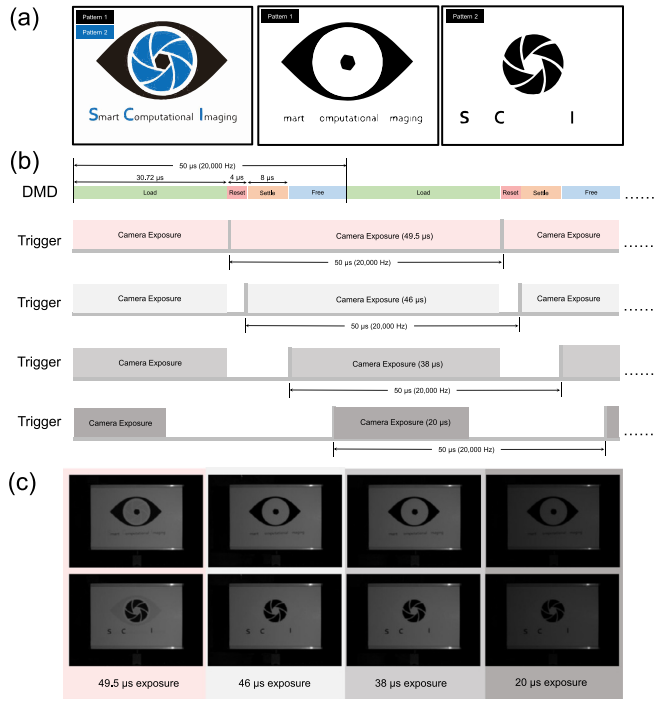

Fig. 3. Captured images with different exposure and trigger settings. (a) Two complementary patterns which are repeatedly projected by the DLP system at 20,000 Hz. (b) Timing diagrams of different exposure and trigger settings used under the test. (c) Corresponding images captured by the high-speed camera. Crosstalk can be obviously seen at 49.5 μs exposure but disappears when the exposure reduces to 46 μs. Further reducing the exposure offers no visible improvement but leads to light intensity attenuation.

order to maximize the exposure time while avoiding potential cross-talk due to the mirror transition. Thus, we program the trigger output from the DMD to the camera according to the waveform shown in the second row of Fig. 2 and set the exposure time of the camera at 46 μs to skip the mirror transition stage. Further reducing the exposure time does not provide visible improvement but leads to light intensity attenuation, as clearly illustrated in Fig. 3. This is due to the fact that the intensity fluctuation during the 8 μs mirror settling time can be effectively averaged out during camera exposure. It should also be noted that the trigger signal output from the DLP board is 2.5 V logic level only while the trigger level required for the camera is 5 V. To fix this problem, we build a voltage translation circuit (SN74ALVC164245, Texas Instruments) to take the output trigger of the DMD at 2.5 V, condition it, and level convert to the required 5 V for the camera. In **Supplementary Video 1**, it is further demonstrated that the designed pattern sequences (see Section **Principle of μFTP** for details) can be repeatedly projected onto a dynamic scene (a rotating desk fan) and precisely captured by a synchronized high-speed camera at 20,000 Hz, indicating the synchronization between the DMD and the camera works well.

2.3. Projection optics

Since the DLP development kit does not include optics, a custom-designed module is attached to the DMD to provide both illumination and projection optics. The projection optics module has two ports, one connecting the DMD and the other connecting the light source. The light source we use is a green LED with an output of 600 lm (Osram). When the light emitting from the LED enters the projection optics module, it is firstly spatially smoothed by an integration rod to create homogeneous illumination, and then directed onto the DMD by relay optics. The projection lens has a working distance of 800–2,000 mm and an aperture of f/3.8. All lens components are coated for optimal transmission between 381 and 650 nm. Besides, a copper heat sink, piping, and a fan are added to improve heat dissipation of the whole projection system.

2.4. Principle of μFTP

The whole framework of μFTP is illustrated in Fig. 4, which operates in two stages: acquisition and reconstruction. In the image acquisition stage, μFTP uses few ($n \geq 2$, e.g., $n = 3$ as illustrated in Fig. 1) high-frequency sinusoidal fringe patterns with slightly different wavelengths (fringe pitches)$\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Between two sinusoids, a "white" pattern with all mirrors of the DMD in the "on" state is inserted in the pattern sequence. Thus, there are totally $2n$ patterns that will be rapidly projected onto the measured object surface sequentially at 20,000 fps. To create quasi-sinusoidal gray-scale fringe patterns with the DMD operating in the binary mode, the ideal sinusoids are binarized with Floyd–Steinberg's error diffusion dithering algorithm [74], and the gray-scale intensity is then reproduced by properly defocusing the projector lens [28,29]. For μFTP, the fringe wavelength set $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ must meet the following two conditions (see **Appendix B** for details):

(1) λ_i should be sufficiently small (frequency is high) as required for successful phase retrieval in conventional FTP based techniques.

(2) The least common multiple (LCM) of the fringe wavelength set should be greater than the total pixel number in the axis wherein the sinusoidal intensity value varies ($LCM(\lambda_1, \lambda_2, \cdots, \lambda_n) \geq W$) so that the phase ambiguities can be theoretically excluded.

For our DMD projector with a resolution of 1024 × 768, we find that three wavelengths $\{\lambda_1, \lambda_2, \lambda_3\} = \{14, 16, 18\}$ pixels are sufficient to make a good tradeoff between the fringe contrast and unambiguous phase range (discussion about the wavelength selection is provided in **Appendix B4**), and they are used throughout the present work.

In the reconstruction stage, the captured pattern images are processed sequentially, with a $2n$-frame sliding window moving across the whole video sequence. Considering $2n$ frames within a window centered on the current frame (I_2 in Fig. 4) at a given time point (t_0), we have n sinusoid images $\{I_1, I_2, \ldots, I_n\}$ and corresponding n white images $\{I_{1B}, I_{2B}, \ldots, I_{nB}\}$, as illustrated in Fig. 4. The 3D shape of the measured object at t_0 is reconstructed from these images based on the 4 steps as follows:

(1) **Background-normalized Fourier transform profilometry (BNFTP)**: BNFTP is an improved version of existing FTP-based approaches [36,38,41] which is specially designed for high-speed 3D measurement with binary patterns. It uses a sinusoidal fringe image and an additional "white" image with all "1"s in the projection pattern to recover a high-quality wrapped phase map. The basic theory and implementation of BNFTP is given in **Appendix A** in details. Based on the idea of the modified FTP approach [41], the "white" image is firstly used to eliminate the background of the fringe image. Then, an additional normalization step is applied to the background-subtracted fringe image, further alienating the influence of surface reflectivity variations. This simple modification allows BNFTP to provide improved performance when measuring textured surfaces. Considering n sinusoidial images $\{I_1, I_2, \ldots, I_n\}$ and corresponding n white images $\{I_{1B}, I_{2B}, \ldots, I_{nB}\}$ within a $2n$-frame sliding window (Fig. 4), we can recover n wrapped phase maps $\{\phi_1, \phi_2, \cdots, \phi_n\}$ based on BNFTP.

(2) **Temporal phase unwrapping based on projection distance minimization (PDM)**: The phase maps obtained by Step (1) are wrapped to principle values of the arctangent function, and consequently, phase unwrapping is required to remove the phase ambiguities and correctly extract the object depth. For μFTP, the phase of the current time point (t_0) is firstly unwrapped temporally by exploiting

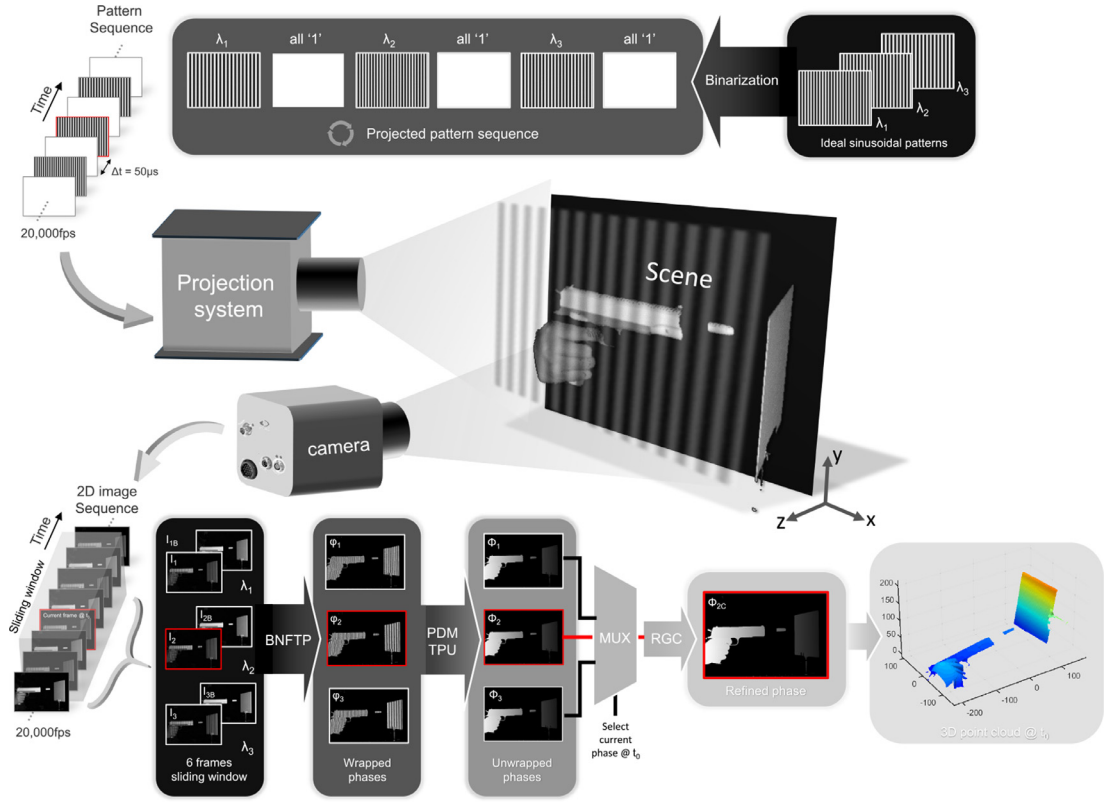

Fig. 4. Overview of μFTP framework (frequency number $n = 3$). 3 high-frequency binarized spatial sinusoids with slightly different wavelengths $\{\lambda_1, \lambda_2, \lambda_3\}$ and 3 white patterns are sequentially projected onto a scene. The camera captures the corresponding synchronized 2D video at 20,000 fps. To reconstruct 3D depth of the scene at a given time t_0, a 6-frame sliding window centered on the current frame at t_0 (indicated by the red box) is applied to extracting 3 sinusoid fringe images $\{I_1, I_2, I_3\}$ and corresponding 3 flat images $\{I_{1B}, I_{2B}, I_{3B}\}$ with a frame interval of 50 μs. Three wrapped phase maps with different wavelengths can be recovered by background-normalized FTP (BNFTP) algorithm, and further unwrapped by projection distance minimization temporal phase unwrapping (PDM TPU) algorithm. The phase corresponding to current time point t_0 (Φ_2) is selected through multiplexer (MUX) and further refined with a reliability-guided compensation (RGC) algorithm in spatial domain. Finally, the refined phase is used to establish the projector-camera correspondences and reconstruct 3D point cloud through triangulation.

information from neighboring frames (as shown in Fig. 4, the current phase ϕ_2 is unwrapped with previous and next wrapped phases ϕ_1 and ϕ_3), based on a newly developed TPU algorithm called PDM (detailed in **Appendix B**). The basic idea of PDM is to determine the optimum fringe order combination $\{k_1, k_2, \ldots, k_n\}$ (for each wrapped phase maps $\{\phi_1, \phi_2, \cdots, \phi_n\}$) so that the corresponding unwrapped phase value combination $\{\Phi_1, \Phi_2, \cdots, \Phi_n\}$ is "closest" (in the Euclidean sense) to the following straight line in dimension n

$$\Phi_1\lambda_1 = \Phi_2\lambda_2 = \ldots = \Phi_n\lambda_n \tag{3}$$

In **Appendix B2**, we also prove that the phase unwrapping results obtained by PDM algorithm is *optimal in a maximum likelihood sense*. It provides a larger unwrapping range as well as better noise robustness than classic multi-wavelength TPU approaches (e.g. heterodyne approach, as demonstrated in **Appendices B3 and B4**). Furthermore, the minimum projection distance in PDM also provides an inherent metric to quantitatively evaluate the unwrapping reliability for each pixel, which is used in the following reliability guided compensation algorithm [Step (3)]. It should be also noted that, if the approximate depth range of the measured scene can be estimated, geometric constraint [58,61–63] can be applied to restricting the search range for possible fringe orders and ruling out several false candidates beforehand (detailed in **Appendix D2**). After this step, a group of unwrapped phase maps $\{\Phi_1, \Phi_2, \cdots, \Phi_n\}$ can be obtained, and only the phase map corresponding the current time point t_0 (Φ_2 shown in Fig. 4) will be further processed by the following steps.

(3) **Reliability guided compensation (RGC) of fringe order error**: Although the initial unwrapped phase map obtained by Step (2) (Φ_2 shown in Fig. 4) encodes depth information of the measured scene at t_0 with a temporal resolution of 100 μs (remember that in Step (1) the principal value of Φ_2 is recovered only from 2 patterns), fringe order errors are inevitable especially around dark regions (lower fringe quality) and object edges (higher motion sensitivity). We propose an approach called RGC for identifying and compensating those fringe order errors by exploiting additional information in the spatial domain (detailed in **Appendix C**). Our key observation is that the fringe order errors are usually isolated (at least less concentrated than the correct phase values) delta-spike artifacts with a phase error of integral multiples of 2π. Inspired by quality guided (spatial) phase unwrapping approaches [75–77], we first gather neighboring pixels within a continuous region of the phase map into groups. Then the isolated pixels or pixels falling into small groups are considered as fringe order errors, and their phase values will be corrected sequentially according to an order ranked by a predefined reliability function (we adopt the minimum projection distance in PDM algorithm as the reliability function, and larger distance means lower reliability). After RGC compensation, we can obtain the refined absolute phase map that is free from fringe order errors (Φ_{2C} shown in Fig. 4)

(4) **Mapping from phase to 3D coordinates**: The final step of μFTP reconstruction is to establish the projector-camera pixel correspondences based on the refined absolute phase map (Φ_{2C} shown in Fig. 4) and to reconstruct the 3D coordinates of the object surfaces at time t_0 based on the calibration parameters of the projector and the camera. In this process, the effect of projection-imaging distortion of lenses is explicitly considered and effectively corrected by an iterative scheme based on lookup table (LUT) implementation (described in **Appendix D**). For more details about the implementation of μFTP, one can refer to the **MATLAB source code** available on our website [78].

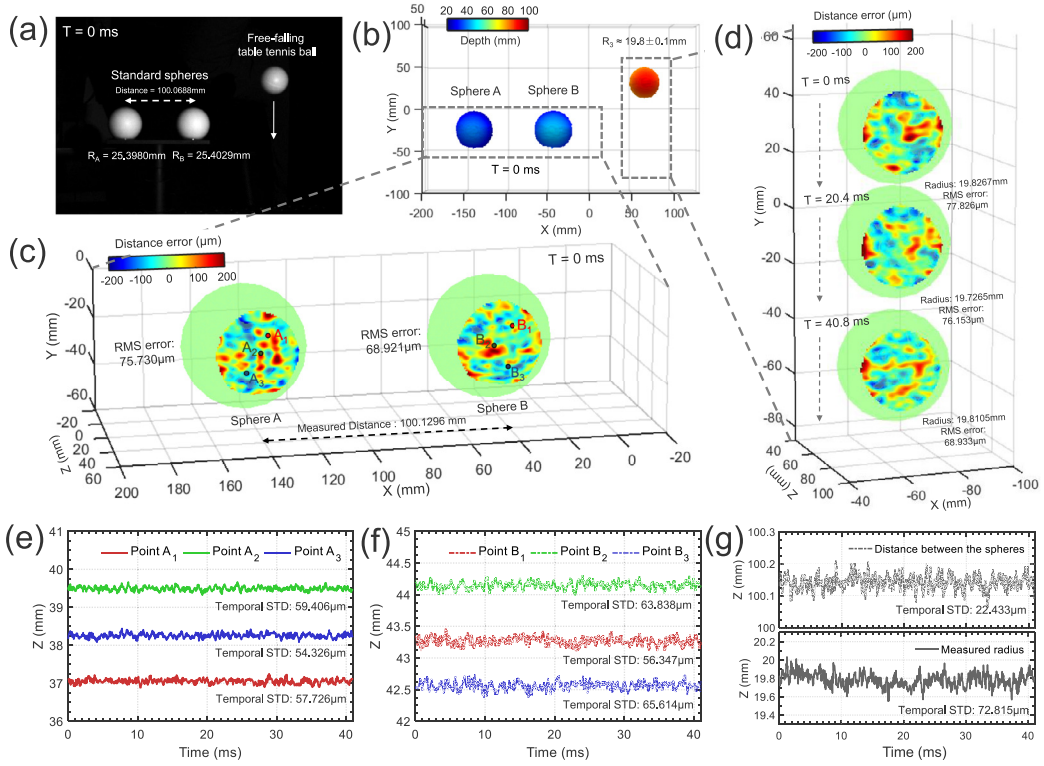

Fig. 5. 3D reconstruction accuracy of μFTP measurement. (a) Test scene consisting of two standard ceramic spheres and a free-falling table tennis ball at T = 0 ms. (b) Corresponding color-coded 3D reconstruction. (c) The error distribution of the measured standard spheres at T = 0 ms. (d) The deviation map (the difference between the measured points and the fitted spheres) of the free-falling table tennis ball at three different time points (T = 0 ms, 20.4 ms, 40.8 ms). (e) Repeatability of 820 measurements over a 41 ms period of 3 points (A_1, A_2, and A_3) on the standard Sphere A. (f) Repeatability of 820 measurements of 3 points (B_1, B_2, and B_3) on the standard Sphere B. (g) Repeatability of 820 measurements of the center-to-center distance between the two spheres as well as the radius of the free-falling table tennis ball.

With the μFTP framework, an accurate, unambiguous, and distortion-free 3D point cloud can be recovered with every two projected patterns. Furthermore, since the phase information is mainly encoded within a single high-frequency sinusoidal pattern, μFTP can be considered "single-shot" in essence, allowing for motion-artifact-free 3D reconstruction with fine temporal resolution (2 patterns per 3D reconstruction, corresponding to 10,000 3D fps). It is also worth mentioning that since μFTP is based on sliding-window reconstruction, any newly added image can be combined with its preceding $2n-1$ images to produce a new 3D result. Although in this case the consecutive two 3D reconstructions may use the same high-frequency fringes for phase evaluation, one can achieve a pseudo frame rate of 20,000 3D fps, which is same as the projector and camera speed.

3. Results

3.1. Quantitative analysis of 3D reconstruction accuracy

To quantitatively determine the accuracy of μFTP measurement, we conduct an experiment on a test scene consisting of two standard ceramic spheres and a free-falling table tennis ball, as shown in Fig. 5(a). The measured objects are put in the measurement volume, which is approximately 400 mm × 275 mm × 400 mm, established by using a calibration panel (see **Appendix D3** for details). According to the calibration based on a coordinate measurement machine (CMM), the radii of the two standard spheres are 25.3980 mm and 25.4029 mm, respectively, and their center-to-center distance is 100.0688 mm. These two spheres with accurate calibrated dimensions are used to quantify the measurement accuracy and repeatability of μFTP system. The table tennis ball is used to test the performance of the system for measuring moving object, whose radius is about 19.8 ± 0.1 mm (dimension uncalibrated, radius is simply measured by a vernier caliper). Fig. 5(b) shows the color-coded 3D reconstruction by μFTP at T = 0 ms and Fig. 5(c) shows the corresponding error distribution of the two measured standard spheres. The accuracy is distinguished by fitting the standard sphere to the point cloud representing the spherical surface and calculating the difference between the measured points and the fitted sphere. As shown in Fig. 5(c), the root mean square (RMS) errors corresponding to the two standard spheres are 75.730 μm and 68.921 μm, respectively. The measured center-to-center distance between the two spheres is 100.1296 mm. In Fig. 5(d), we further show the deviation maps of the free-falling table tennis ball at three different time points (T = 0 ms, 20.4 ms, 40.8 ms). Note that since the dimension of the table tennis ball is uncalibrated, we determine the best spherical fit on the 3D point cloud to estimate the spherical diameter, and the deviation maps shown in Fig. 5(d) are the difference between the measured points and the fitted spheres. These results indicate that the measurement accuracy of μFTP is better than 80 μm for both static and dynamic objects. Repeatability of the μFTP measurement is also analyzed by performing 820 measurements over a 41 ms period (at a pseudo frame rate of 20,000 fps) for 3 different points on respective standard sphere [$A_1 \sim A_3$ and $B_1 \sim B_3$, as labelled in Fig. 5(c)], the center-to-center distance between the two standard spheres, and the radius of the free-falling table tennis ball, as shown in Fig. 5(e)–(g), respectively. The temporal movie of color-coded 3D reconstruction of the test scene and the corresponding error analysis over the 41 ms period is further provided in **Supplementary Video 2**. One can clearly observe the excellent repeatability of μFTP system: the center-to-center distance measurement exhibited a very low temporal standard deviation (STD) of 22.433 μm; the temporal STD at one given point on the standard sphere is typically around 60 μm; the

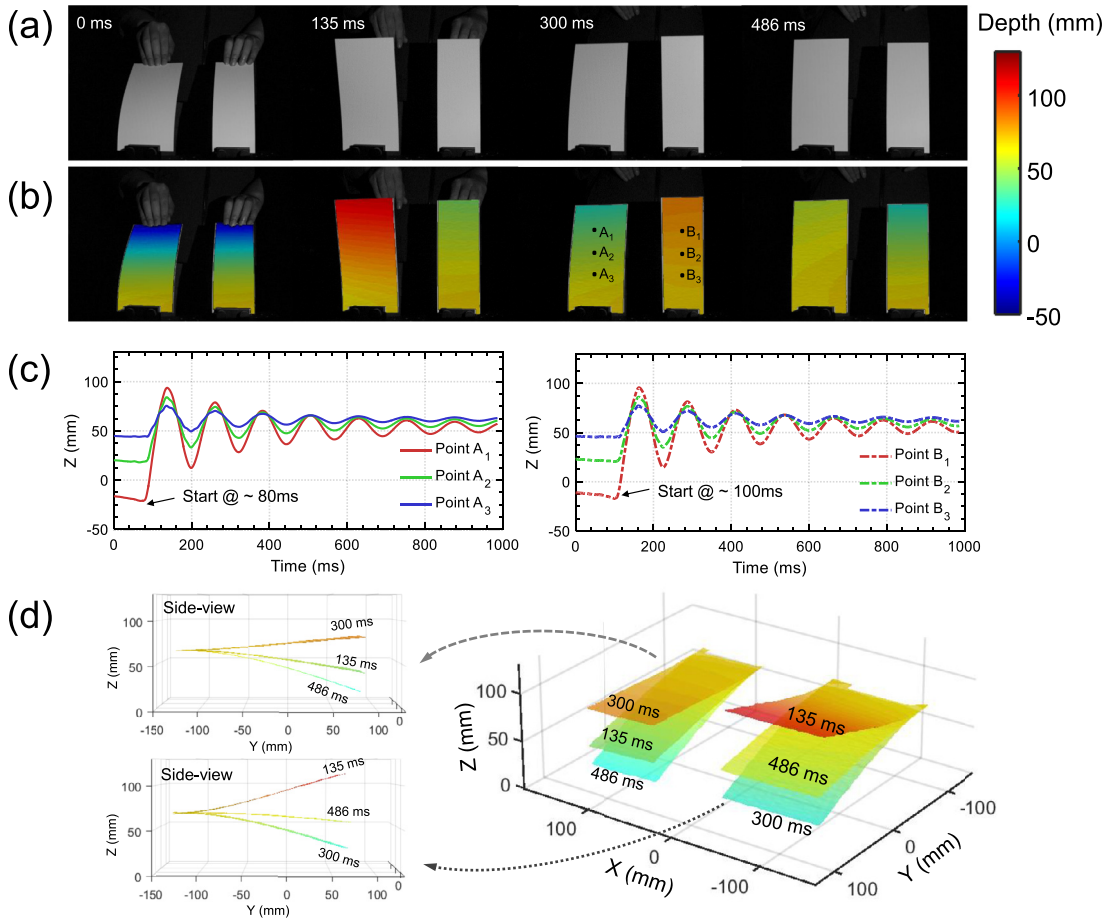

Fig. 6. Measurement of two vibrating cantilever beams. (a) Representative camera images at different time points. (b) Corresponding color-coded 3D reconstructions. (c) Displacement of 3 points on each cantilever [$A_1 \sim A_3$ and $B_1 \sim B_3$, as labelled in (b)] as a function of time. (d) The 3D geometric field of the two cantilevers at three different time points. The two insets show the side-views (y-z plane) of the respective cantilever.

radius of the free-falling table tennis ball has a slightly higher temporal STD (72.815 μm) due to the object motion. These results show that the presented system achieves a measurement accuracy better than 80 μm and a temporal STD below 75 μm within 400 mm × 275 mm × 400 mm measurement volume size.

3.2. Vibrating cantilevers

In this subsection, vibrating cantilever beams are used to validate μFTP fast 3D imaging capability and compare its performance with the state-of-the-arts. Firstly, two fixed-end homogeneous plastic cantilever beams are imaged. The dimensions of the two cantilever beams are 215 mm × 110 mm × 1.5 mm [for the larger one shown on the left of Fig. 6(a)] and 215 mm × 80 mm × 1.5 mm [for the smaller one shown on the right of Fig. 6(a)]. One end of each cantilever is clamped on the optical table while the other end is excited by human hand. Fig. 6(a) and (b) show representative camera images (white pattern) and corresponding color-coded 3D reconstructions by μFTP at different time points. Initially, the two cantilevers are heavily bent with manual pressure, with their ends aligned at about the same depth. When hands release, the stored elastic potential energy converts to kinetic energy to enable vibrations. To study the vibrating process quantitatively, the out-of-plane (z) displacement of 3 points on each cantilever [$A_1 \sim A_3$ and $B_1 \sim B_3$, as labelled in Fig. 6(b)] are plotted as a function of time in Fig. 6(c). The plots show that the largest vibration amplitudes occur at points A_1 and B_1, since they are closer to the free ends of the cantilevers. Their vibration amplitudes gradually reduced from about 50 mm to 10 mm (for point A_1 and B_1) with a frequency of ~ 8 Hz. Besides, there is about 20 ms time difference between the two hand releases, making the vibrations of the two cantilever out-of-phase. In Fig. 6(d), we show the reconstructed 3D shapes of the two cantilevers at three different time points, with the two insets showing the side-views (y-z plane) of the respective cantilever. The movie of color-coded 3D rendering of the two cantilevers surfaces as well as the corresponding side-views is provided in **Supplementary Video 3**. These results verify that the proposed μFTP enables high-speed 3D reconstruction of object vibration and provides high-accuracy quantitative evaluation of any characteristic points on the object surface.

Next, we compare our μFTP with the well-known three-wavelength phase shifting profilometry (PSP) [47,53] by using only the larger cantilever beam. One end of the cantilever is vertically fixed on the optical table, and the free end is heavily excited to create large amplitude vibration. In order to achieve a fair comparison, we project the required 15 patterns for both μFTP (6 patterns) and three-wavelength PSP (9 patterns) sequentially onto the same scene, according to the pattern sequence illustrated in Fig. 7(a). Both algorithms use the same fringe wavelengths $\{\lambda_1, \lambda_2, \lambda_3\} = \{14, 16, 18\}$ pixels, and the retrieved phases are processed with the same algorithms (PDM unwrapping and RGC), respectively. The final 3D point clouds are both reconstructed from the respective unwrapped phases corresponding to λ_2. Since the two groups of patterns are sequentially projected onto the moving surface, there exists a 400 μs time difference between the two 3D reconstructions from these two methods. Fig. 7(b) shows representative camera images (white patterns from μFTP) at different time points. In Fig. 7(c) and (d), we

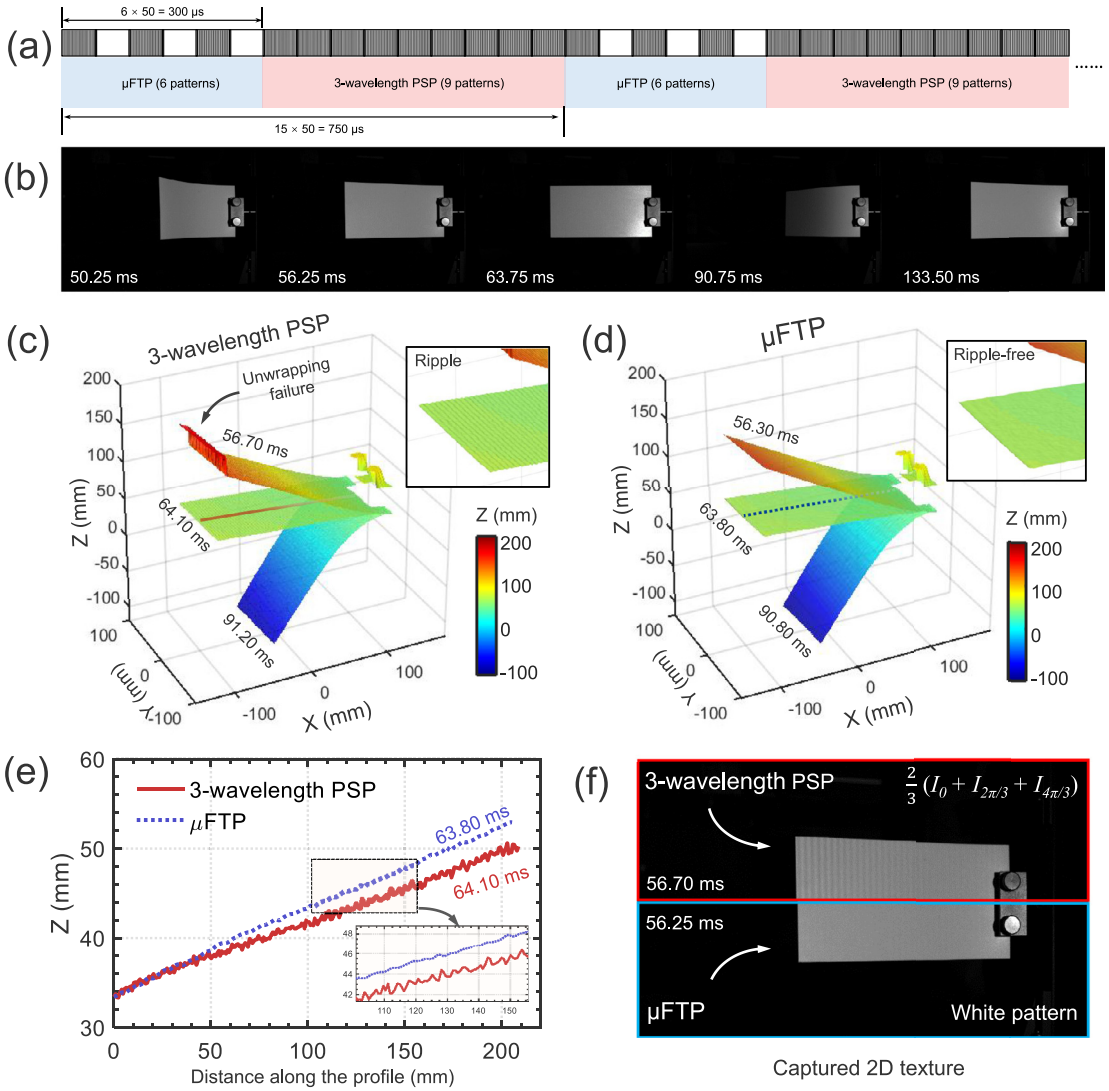

Fig. 7. Comparison of μFTP with three-wavelength PSP. (a) Projected pattern sequence (6 patterns for μFTP and 9 pattern for 3-wavelength PSP). (b) Representative camera images (white pattern from μFTP) at different time points. (c) 3D reconstructions of the cantilever surface by 3-wavelength PSP at three different time points. (d) 3D reconstructions of the cantilever surface by μFTP at three different time points. (e) Line profiles along the central lines of cantilever surface [corresponding to the dashed lines in (c) and (d)]. (f) Comparison of 2D texture obtained by three-wavelength PSP and μFTP.

compare the reconstructed 3D surfaces of the cantilever at three different time points. As clearly highlighted in the zoom area and the line profile [Fig. 7(e)], apparent ripples can be observed in the results of 3-wavelength PSP, which also cause unwrapping errors around the end of cantilever (where the vibration amplitude is high). Whereas the μFTP produces decent 3D reconstruction without notable motion ripples (see also the **Supplementary Video 4** for a comparative movie). It should be pointed out that the flat image in μFTP is inherently a normal 2D image that can be used for texture mapping. However, as shown in Fig. 7(f), the fringe images have to be averaged to create a texture for PSP, which is also easily distorted by the object motion (see also the **Supplementary Video 5** for a comparative movie of the 3D cantilever surfaces with texture mapping). These results verify that μFTP is completely immune to motion ripples [46,50], leading to distortion-free 3D reconstruction along with high-quality 2D texture even though the motion is fast and the out-of-plane displacement is large.

Finally, we compare our μFTP with conventional FTP [36,38] and modified FTP [41]. The measured scene is similar with the previous demonstration but an additional sticker (our lab logo) is pasted on the cantilever surface to create large reflectivity variations [Fig. 8(a)]. All three algorithms use the same raw image data (the white images are not used in conventional FTP), and the retrieved phases are processed with the same algorithms (PDM unwrapping and RGC). The reconstructed 3D surfaces of the cantilever at three different time points by using the three methods are shown in Fig. 8(b)–(d), respectively. As can be seen, the results generated by conventional FTP suffer from obvious artifacts and unwrapping errors around sharp edges and dark regions. Modified FTP eliminates the unwrapping error and reduces the artifacts through background substraction. But there still remain severe fluctuations (highlighted in the zoom area), which is mainly caused by large reflectivity variations. In contrast, μFTP produces a much smoother reconstruction without notable artifacts because the influence of surface reflectivity variations is significantly suppressed. These results suggest that μFTP notably improves the performance of state-of-the-arts in terms of accuracy and robustness when measuring textured surfaces.

3.3. Rotating fan blades

The next test object is a commercially available desk fan with 3 blades made of ~ 1.6 mm thick plastic. The fan is fixed on the optical table with its front protect shell removed so that the fan blades can be directly exposed to the measurement system. As shown in Fig. 9(a),

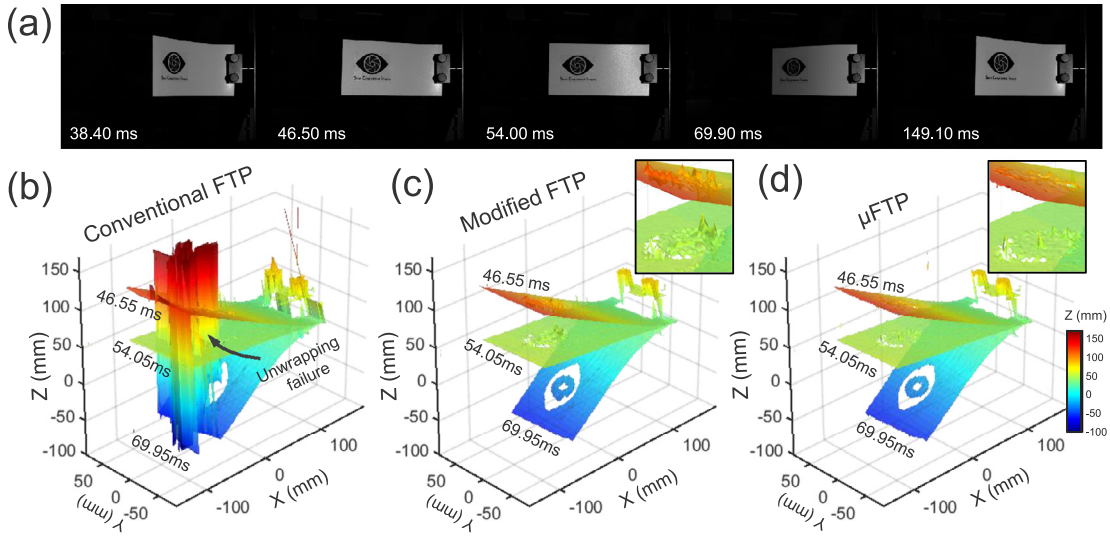

Fig. 8. Comparison of conventional FTP, modified FTP (background subtraction), and μFTP (background subtraction and normalization). (a) Representative camera images at different time points. (b) 3D reconstructions of the cantilever surface with conventional FTP at three different time points. (c) 3D reconstructions with modified FTP. (d) 3D reconstructions with μFTP.

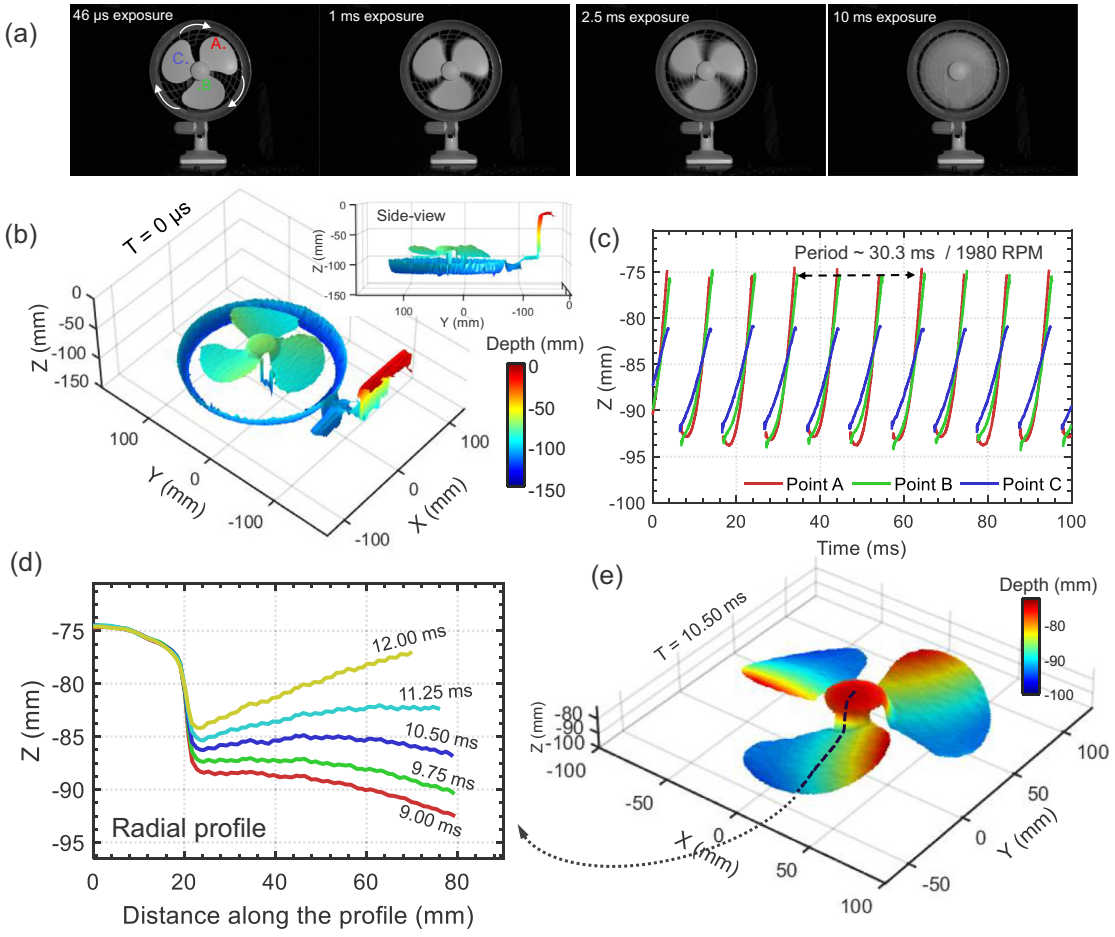

Fig. 9. Measurement of rotating blades. (a) Images captured by the camera at different exposure time (46 μs, 1 ms, 2.5 ms, and 10 ms). (b) The reconstructed 3D shape of the whole fan at the start of the observation time (T = 0 ms), the inset shows the side-view (y-z plane) of the 3D reconstruction. (c) Displacement in the z direction at 3 chosen point locations [A, B, and C, shown in (a)] as a function of time over a 100 ms period. (d) 5 line profiles drawn along the radial direction [corresponding to the dashed line in (e)] at time intervals of 0.75 ms. (e) The color-coded 3D rendering of the fans surface at T = 10.5 ms.

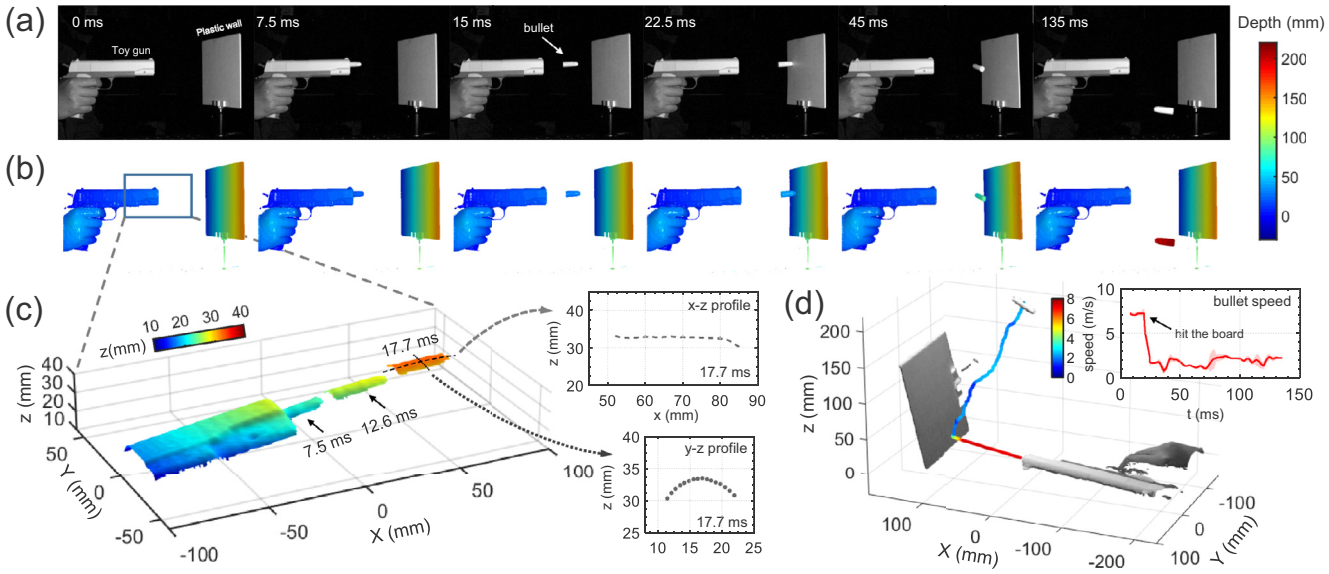

Fig. 10. 3D measurement and tracking a bullet fired from a toy gun. (a) Representative camera images at different time points. (b) Corresponding color-coded 3D reconstructions. (c) 3D reconstruction of the muzzle region [corresponding to the boxed region shown in (b)] as well as the bullet at three different points of time over the course of flight (7.5 ms, 12.6 ms, and 17.7ms). The insets show the horizontal (x–z) and vertical (y-z) profiles crossing the body center of the flying bullet at 17.7 ms. (d) The 3D point cloud of the scene at the last moment (135 ms), with the colored line showing the 130 ms long bullet trajectory. The inset plots the bullet velocity as a function of time.

though the fan is rotating at its highest possible speed, the 46 μs exposure time of μFTP system is short enough to freeze the high-speed motion and record a clear image of the whirling fan blades. To illustrate the rotating speed of the fan blades more intuitively, we increase the camera exposure from 46 μs to 1 ms, 2.5 ms, and 10 ms. The blade edges become increasingly blurred. And finally at 10 ms exposure, we are unable to identify the fan blades since the motion blur makes them to appear as one streak. Despite of this challenging rotation speed, the 3D shape of the whole fan, including the base, center hub, side cover, and three blades are well reconstructed by μFTP, as shown in Fig. 9(b). One point on each blade is chosen to demonstrate the cyclic displacements of the fan blades [points A, B, and C, shown in Fig. 9(a)]. Displacement in the z (out-of-plane) direction at the chosen point locations are plotted as a function of time over a 100 ms period, as shown in Fig. 9(c). The plot shows that the fan has a rotation period of approximately 30.3 ms, corresponding to a speed of 1980 rotations per minute (rpm). The plot also demonstrates a good repeatability of the μFTP measurement. Besides, by applying a proper threshold on the captured white pattern, a binary mask can be generated to extract the moving fan blades region of interest from the static background. The color-coded 3D rendering of the fans surface at one point in time (10.5 ms) is illustrated in Fig. 9(e). Fig. 9(d) further gives 5 line profiles drawn along the radial direction out from the center hub [corresponding to the dashed line in Fig. 9(e)] at time intervals of 0.75 ms. Within the short 3 ms, the fan blade quickly sweeps through the radical profile for about 1/10 of a revolution, resulting in a maximum variation over 15 mm in the z direction. The results also demonstrate that the length of fan blade is $\sim$ 80 mm in the radial direction out from the center hub which has a radius of $\sim$ 20 mm. The corresponding 3D movie is further provided in **Supplementary Video 6**. It is important to mention that unlike the previous study based on stroboscopic structured illumination [79], here we truly recorded the entire process of fan rotation without any stroboscopic time gap between two successive 3D frames.

3.4. Bullet fired from a toy gun

Next, we apply μFTP to image one-time transient event: a bullet fired from a toy gun and then rebounded from a plaster wall. Fig. 10(a)-(b) show representative camera images (white pattern) and corresponding color-coded 3D reconstructions at different time points. T = 0 ms is the start of the observation time, and the bullet begins to occur in the vicinity of the gun muzzle at about T = 7.5 ms. After travelling in free-flight for about 15 ms, the bullet hits the plaster and rebounds towards the camera. In Fig. 10(c), we show the 3D reconstruction of the muzzle region [corresponding to the boxed region in Fig. 10(b)] as well as the bullet at three different points of time (7.5 ms, 12.6 ms, and 17.7 ms). The two insets further provide the horizontal (x–z) and vertical (y-z) profiles crossing the body center of the bullet at 17.7 ms, which indicate that the bullet has a length of $\sim$ 35.5 mm and a diameter of $\sim$ 11.8 mm. Besides, the 3D data can be used to quantitatively analyze the process with regards to the ballistic trajectory and velocity. By tracing the center of the bullet body, we can obtain the bullet trajectory in 3D space. The instantaneous speed of the bullet can then be estimated by taking the derivative of the position function with respect to time. The calculated muzzle velocity (velocity of the bullet when it leaves the barrel) is around 7.3 m/s. At that speed, the bullet moves about one pixel between each camera frames. Since the phase information is encoded in single fringe pattern in μFTP, the frame-by-frame motion does not introduce any visible artifacts. In Fig. 10(d), we further show the 3D point cloud of the scene at the last moment (T = 135 ms), with the colored line showing the 130 ms long bullet trajectory (the bullet velocity is encoded by the line color, see **Supplementary Video 7** for time evolution of the trajectory). The inset of Fig. 10(d) provides the bullet velocity as a function of time, indicating that the bullet speed keeps almost constant when travelling in free-flight, and suddenly reduces to about 2 m/s during the collision. The fluctuation of speed after collision mainly results from the rolling over of the bullet, which also increases the estimation uncertainty [red shaded region in Fig. 10(d)] due to the difficulties in accurate tracking of the bullet body center. A more detailed illustration of the transient event is provided in **Supplementary Video 8**, which is a slow-motion 3D movie containing 2700 3D frames with a frame interval of 50 μs (corresponding to 20,000 3D fps over an observation period of 135 ms). These experimental results demonstrate the potential applications of μFTP for tracking 3D trajectory of fast moving object within a wide observation volume.

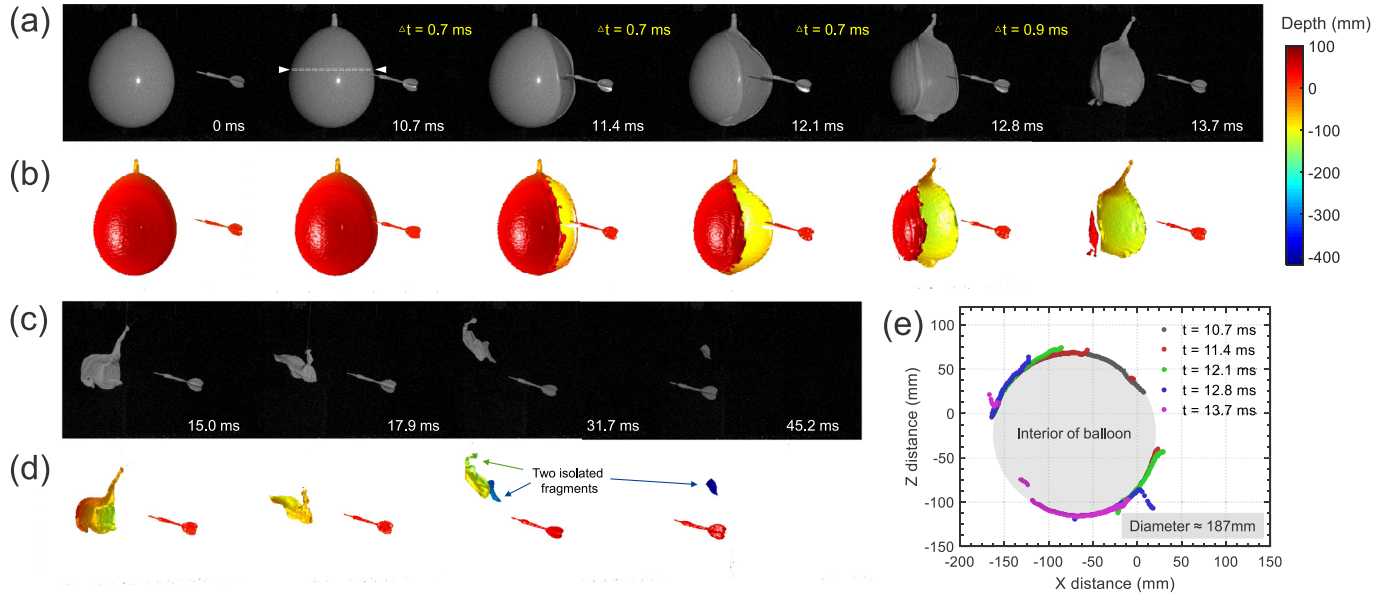

Fig. 11. Balloons explosion triggered by a flying dart. (a) Representative camera images at different time points. (b) Corresponding color-coded 3D reconstructions. (c) and (d) continue to (a) and (b), respectively. (e) Line profiles across the dashed line in (a) at the time points of 10.7 ms, 11.4 ms, 12.1 ms, 12.8 ms and 13.7 ms.

3.5. Balloons explosion triggered by a flying dart

In the last demonstration, the proposed μFTP is applied to capturing a very-high-frequency event air balloon bursting punctured by a flying dart. Fig. 11(a)–(d) show representative 2D camera images (white pattern) and corresponding color-coded 3D reconstructions at different time points. The corresponding movie is provided in **Supplementary Video 9**. The balloon is suspended in air by threads and keeps still until its surface touchs the tip of the flying dart (T = 10.7 ms). The pricked hole propagates into a radial crack towards to two poles of the balloon, slicing the balloon into a piece of rubber membrane (T = 13.7 ms. The membrane then shrinks into a crumpled form rapidly (T = 15 ms) and finally breaks into 2 fragments (T = 31.7 ms, indicated by the yellow and blue arrows). The whole process lasts about 47 ms, while the key event (balloon blowing up) takes place only within about 4 ms (note that the frame intervals from the second to sixth 3D snapshots in Fig. 11(b) are less than 1 ms). The 3D data can be used to quantitatively analyze the explosion process. Fig. 11(e) shows 5 line profiles across the dashed line shown in Fig. 11(a), corresponding to the time points of 10.7 ms, 11.4 ms, 12.1 ms, 12.8 ms and 13.7 ms. When the balloon is intact, only the top surface can be imaged (shown in gray). With the crack expanding and propagating, the bottom (inner) surface of the balloon is revealed. It is interesting that, except for the eversion around the crack boundaries, the main balloon surface still demonstrates a good axi-symmetry during explosion (3 ms), characterized by a longitudinal diameter of $\sim$ 187 mm [shaded region in Fig. 11(d)]. In **Supplementary Video 9**, we can see that the depth information of the sudden explosion within time spans on the order of tens of microseconds is fully recovered by μFTP. The 3D results show good image quality without depth ambiguities (note in Fig. 11(c) and (d), the two "overlapping" fragments in 2D camera images are actually separated in 3D space due to their large difference in depth). Some depth artifacts noticeable are attributed to the insufficient scene overlapping between adjacent camera frames because μFTP relies on certain spatio-temporal redundancy for phase unwrapping.

4. Discussion

In this study, we have demonstrated μFTP, which is able to reconstruct dense and precise 3D shapes of complex scenes of several independently moving objects at 10,000 fps. Within the entire measurement volume of 400 mm $\times$ 275 mm $\times$ 400 mm, we achieve a depth accuracy better than 80 μm and a temporal uncertainties less than 75 μm. μFTP has several advantages over conventional high-speed 3D imaging methods. Not only can the high-resolution, unambiguous depth information be retrieved from two images, but also high-quality 2D textures (white patterns) provided simultaneously with the 3D geometry. This allows filling in the speed gap between high-speed 2D photography and fast 3D sensing, pushing the speed limits of unambiguous, motion-artifact-free 3D imaging to the range of tens of kilo-hertz (half of the native camera frame rate). The effectiveness of μFTP has been verified by several experiments on various types of transient events, including objects that are rapidly moving or undergoing sudden shape deformation. Experimental results suggest the potential applications of μFTP in various fields, such as solid mechanics, material science, fluid dynamics, and biomedical research. Furthermore, μFTP is highly flexible: the fringe pitches, number of wavelengths, sliding window length can be adjusted according to the surface characteristics and motion speed of objects, for instance using more than 3 wavelengths to achieve higher reconstruction reliability for more complex objects when the motion speed is lower than the camera frame rate.

Being a recording and post-processing technique, the processing speed of μFTP has not yet been fully optimized. We have implemented the μFTP reconstruction in MATLAB (the RGC algorithm is written in C++ language and called from MATLAB using the Mex "MATLAB Executable" dynamically linked subroutine). The reconstruction code can be accessed from our website [78]. The time required for reconstructing one 3D frame is approximately 870 ms on a desktop computer (Intel Core i7-4790 CPU 3.6 GHz, 16 GB RAM). The processing speed can be significantly improved by using graphics processing units (GPUs), as the involved algorithms, such as 2D fast Fourier transform and pixel-wise PDM phase unwrapping are highly parallelizable. This could further enable μFTP to execute real-time 3D video reconstruction with low latency, if a high-speed camera with synchronous data acquisition capability is employed. Moreover, due to the very similar architecture and general versatility, μFTP is possible to be extended to other computational illumination based imaging techniques for high-speed imaging tasks, such as structured illumination microscopy [18,80,81] and computational ghost imaging [82,83].

Finally, it should also be mentioned that due to the spatio-temporal nature of μFTP reconstruction, the reconstructed point clouds are not

strictly independent. Though the principal value of the phase is retrieved from two patterns (mainly from a single high-frequency sinusoidal pattern), the PDM phase unwrapping procedure relies on all frames within the sliding window (6 frames in our case). Since phase unwrapping only remove the modulus 2π ambiguities without affecting the principal value of the phase, we take advantage of the inherent spatio-temporal correlation between successively captured phase images. As we have shown, when the object motion does not induce significant discrepancies between consecutive phase images, μFTP can still produce accurate and reliable 3D reconstruction based on robust temporal phase unwrapping and spatial phase error compensation. However, for scenes with very fast object motion, it is prone to fail due to large phase discrepancies or insufficient overlap of the object between consecutive phase images, which does impose certain restrictions on its practical application. In future work, we hope to further improve the robustness of phase unwrapping by introducing multiple cameras or perspectives. The use of multi-viewpoint geometric constraints increases the data redundancy and allows the determination of absolute phase with a shorter sliding temporal window. This improvement should make our system more feasible for the measurement of faster moving objects.

Acknowledgments

This work was supported by the National Key Technologies R&D Program of China (2017YFF0106403), National Natural Science Fund of China (61722506, 61705105, 111574152), Final Assembly '13th Five-Year Plan' Advanced Research Project of China (30102070102), Outstanding Youth Foundation of Jiangsu Province of China (BK20170034), 'Six Talent Peaks' project of Jiangsu Province, China (2015-DZXX-009), '333 Engineering' Research Project of Jiangsu Province, China (BRA2016407, BRA2015294), Fundamental Research Funds for the Central Universities (30917011204, 30916011322), and Open Research Fund of Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense (3091601410414).

Appendix A. Background Normalized Fourier Transform Profilometry (BNFTP)

A1. Theory

In fringe projection profilometry technique, the sinusoidal fringe pattern designed in projector space can be represented as

$$I^p(x^p, y^p) = a^p + b^p \cos\left(2\pi f_0^p x^p\right) \tag{A.1}$$

where a^p is the mean value, and b^p is the amplitude (or projector modulation), (x^p, y^p) is the pixel coordinate of the projector, f_0^p is the frequency of the sinusoidal fringe. All generated fringe patterns are sent to a projector and then projected onto the object. Without loss of generality, we assume that the fringes are oriented perpendicular to the x^p axis, as we did in this work. When the fringe pattern is projected onto the object surface, the light reflected from the object can be attributed to two sources: projector light as well as the ambient light. The projected fringe pattern combined with the ambient light β_1 is modulated and reflected by the object. The reflected light is then captured by the camera with some additional ambient light β_2 directly entering the camera. So, the fringe image actually captured by the camera is:

$$\begin{aligned} I(x^c, y^c) = {} & \alpha(x^c, y^c)\{a^p + b^p cos[2\pi f_0 x^c + \phi(x^c, y^c)] \\ & + \beta_1(x^c, y^c)\} + \beta_2(x^c, y^c) \end{aligned} \tag{A.2}$$

where (x^c, y^c) is the pixel coordinate in the camera space, $\alpha(x^c, y^c)$ is the reflectivity of the measured object, and $\phi(x^c, y^c)$ is the phase containing the depth information of the object. For simplicity, $I(x^c, y^c)$ is commonly expressed as follows

$$I\left(x^c, y^c\right) = A\left(x^c, y^c\right) + B\left(x^c, y^c\right) \cos\left[2\pi f_0 x^c + \phi\left(x^c, y^c\right)\right] \tag{A.3}$$

where $A(x^c, y^c)$ is the average intensity which equals $\alpha(x^c, y^c)\left[a^p + \beta_1(x^c, y^c)\right] + \beta_2(x^c, y^c)$, $B(x^c, y^c)$ represents the intensity modulation which equals $\alpha(x^c, y^c)b^p$, f_0 is the carrier frequency of the captured fringe image. The 2D Fourier transform of $I(x^c, y^c)$ gives

$$\hat{I}\left(f_x, f_y\right) = \hat{A}\left(f_x, f_y\right) + \hat{C}\left(f_x - f_0, f_y\right) + \hat{C}^*\left(f_x + f_0, f_y\right) \tag{A.4}$$

where (f_x, f_y) is the vector in spatial frequency domain corresponding to (x^c, y^c). $\hat{A}\left(f_x, f_y\right)$ and $\hat{C}\left(f_x - f_0, f_y\right)$ are the Fourier transforms of $A(x^c, y^c)$ and $\frac{1}{2}B(x^c, y^c)\exp\left\{-i\left[2\pi f_0 x^c + \phi(x^c, y^c)\right]\right\}$, respectively, and * denotes the complex conjugate. Note that influence of the phase factor $\exp\left(-i2\pi f_0 x\right)$ can be interpreted as a translation of f_0 in the spatial frequency domain. In conventional Fourier Transform Profilometry (FTP) [25,36–38], ϕ, A, and B are assumed to vary slowly compared with the carrier frequency f_0, and thus, the zero order ($\hat{A}\left(f_x, f_y\right)$) is separated with the $+1$ order ($\hat{C}\left(f_x - f_0, f_y\right)$) and -1 order ($\hat{C}^*\left(f_x + f_0, f_y\right)$) in the frequency domain. Then a properly designed band-pass filter can be applied to extracting the $+1$ order ($\hat{C}\left(f_x - f_0, f_y\right)$), and the phase can be retrieved by taking the angle part of the resultant inverse Fourier transform.

One major limitation of conventional FTP lies in the fact that when the measured surface contains sharp edges, discontinuities, or large surface reflectivity variations, the support of the three terms in Eq. (A.4) will be significantly extended, so that the zero frequency may overlap with the $+1$ and -1 orders[37]. The spectrum overlapping makes it difficult to filter out the $+1$ order ($\hat{C}\left(f_x - f_0, f_y\right)$), precluding high-accuracy phase reconstruction of complex objects. To address this problem, π -shift FTP [39] and modified FTP [41] have been proposed to effectively suppress the zero order by projecting an additional π -shift sinusoidal fringe image or a flat image. However, in π -shift FTP, the phase information is encoded in two sinusoidal fringes and thus the sensitivity to object motion is increased. While in modified FTP, the flat image (average gray-scale of the sinusoidal fringe pattern) is not able to be perfectly generated by the DMD operating in binary (1-bit) mode (since only "0"s and "1"s can be projected, the gray-scale "0.5" needs to be created by additional spatial dithering). Furthermore, though the effect of zero order is largely removed, neither approach can handle large surface reflectivity variations, which introduce spectrum leakage and thus influence the high-quality phase retrieval.

Therefore, in this work, we propose a improved FTP-based approach so-called Background Normalized Fourier Transform Profilometry (BNFTP) that is specially designed for high-speed 3D measurement with binary patterns. Similar to modified FTP [41], BNFTP uses a sinusoidal fringe image and an additional "white" image with all "1"s in the projection pattern (all mirrors in the DMD are in the "on" state). The fringe image is created by binarizing an ideal spatial sinusoid, therefore the parameter a^p and b^p in Eq. (A.1) should be both 0.5:

$$I_1^p(x^p, y^p) = \frac{1}{2} + \frac{1}{2}\cos\left(2\pi f_0^p x^p\right) \tag{A.5}$$

The additional while image can be simply represented as:

$$I_2^p(x^p, y^p) = 1 \tag{A.6}$$

In high-speed imaging conditions, the effect of the ambient light (β_1 and β_2) can be neglected compared with the strong projector light within the very limited exposure time (46 μs for our μFTP system), especially when all other light sources are turned off in measurement environment. Then the corresponding captured images can be simplified as

$$I_1\left(x^c, y^c\right) = \frac{1}{2}\alpha\left(x^c, y^c\right) + \frac{1}{2}\alpha\left(x^c, y^c\right)\cos\left[2\pi f_0 x^c + \phi\left(x^c, y^c\right)\right] \tag{A.7}$$

$$I_2\left(x^c, y^c\right) = \alpha\left(x^c, y^c\right) \tag{A.8}$$

By taking the normalized difference between I_1 and I_2, the zero-frequency term as well as the effect of surface reflectivity variations can be effectively removed:

$$I_d\left(x^c, y^c\right) = \frac{2I_1 - I_2}{I_2 + \gamma} = \cos\left[2\pi f_0 x^c + \phi\left(x^c, y^c\right)\right] \tag{A.9}$$

where γ is a small constant to prevent divide-by-zero error. Then Fourier transform is applied on the normalized image I_d to extract the phase information. With the subtraction and normalization of the white image, the effect of zero-order as well as surface reflectivity variations is removed before the Fourier transform, and the spectrum overlap in the frequency domain can be prevented or significantly alleviated. Furthermore, the white image I_2 is also used to define a binary mask. The pixel with sufficient large reflectivity ($I_2(x^c, y^c) > threshold$) is considered reliable in determining that pixels phase, and only these pixels are used to reconstruct a 3D surface.

A2. Experimental comparison with state-of-the-art FTP methods

To demonstrate the performance of BNFTP in terms of phase error reduction, we measure an A4-size (21.0 mm × 29.7 mm) textured flat board with 12-by-8 checkerboard pattern printed on its surface. The corresponding sinusoidal fringe image and white image captured by the μFTP system (camera exposure 46 µs, 20,000 fps, see Section **Projection and capture synchronization** for details) are shown in Fig. A.1(a) and (b), respectively. Fig. A.1(c)–(e) show the results obtained by the conventional FTP method (without using the white image), background subtracted FTP (similar to modified FTP[41] but use $2I_1 - I_2$ for phase retrieval), and BNFTP (use Eq. (A.9) for phase retrieval). The images from left to right are processed fringe patterns, Fourier spectra, filtered Fourier spectra (180 × 300 Hanning window), recovered wrapped phases (carrier phases removed by spectrum centering for better illustration of the phase error), 3D reconstructions, and the corresponding line profiles (for phase to 3D conversion, see **Appendix D** for details). The root mean square (RMS) error shown here is calculated by fitting the line profile to a straight line and calculating the difference between the measured points about the fitted line. In the Fourier spectrum of conventional FTP approach, three terms corresponding to the 0 and ± 1 orders are clearly visible. However, they are significantly extended and overlapped because of the background and surface reflectivity variations, leading to prominent reconstruction artifacts (RMS error 664.15 µm). Due to the spectrum leakage, such artifacts are not just limited to the dark square regions but propagate to bright regions and degrade the reconstruction accuracy prevailingly. Subtracting the background can suppress the zero-order and alleviate the spectrum overlapping to some extent, reducing the RMS error to 169.66 µm. However, the reconstruction error is still obvious because of the negligible spectrum overlapping between ± 1 orders. All these errors are almost eliminated in the results of BNFTP, as shown in Fig. A.1(e). With the subtraction and normalization of the white image, the zero-order is removed and the support of ± 1 orders is significantly shrunk into a compact star-shaped region. The 3D result illustrates the improved smoothness as we expected. The RMS error along the line profile is only 61.78 µm, which is reduced by 11 times and 2.74 times as compared with conventional FTP and background subtracted FTP (modified FTP), respectively.

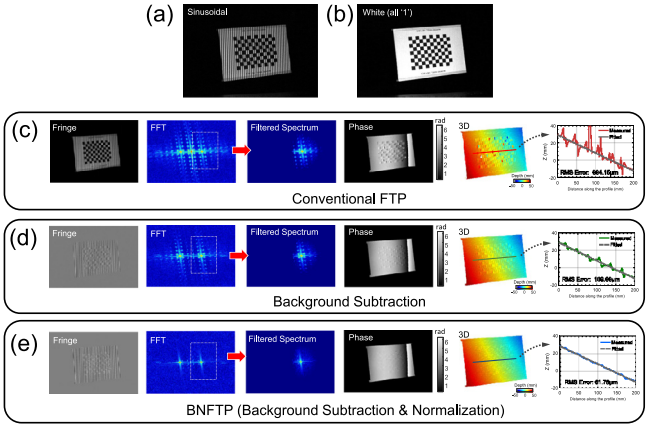

Fig. A.1. Measurement of a textured flat board. (a) Sinusoidal fringe image. (b) White image. (c) Results obtained by conventional FTP method without using the white image. (d) Results obtained by background subtracted FTP (similar to modified FTP but use $2I_1 - I_2$ for phase retrieval). (e) Results obtained by BNFTP. The images from left to right are: processed fringe patterns, Fourier spectra, filtered Fourier spectra (180 × 300 Hanning window), recovered wrapped phases (carrier phases removed by spectrum centering for better illustration of the phase error), 3D reconstructions, and the corresponding line profiles (labelled by the red, green, and blue solid lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Appendix B. Temporal Phase Unwrapping with Projection Distance Minimization (PDM)

B1. Theory

The initial phase retrieved by BNFTP is wrapped to principle values of the arctangent function, and consequently, the phase discontinuities occur when the unknown true phase changes by 2π. Thus, phase unwrapping is required to remove the ambiguities and correctly extract the object depth. With only single wrapped phase map, spatial phase unwrapping algorithms cannot uniquely determine the period numbers for the cases of large discontinuities or spatially isolated surfaces. Thus, temporal phase unwrapping (TPU) approaches with more than one relative phase maps have to be used to remove such depth ambiguities [52]. In this work, a new algorithm so-called projection distance minimization (PDM) is proposed to address the multi-frequency temporal phase unwrapping in an optimum way. Without loss of generality, for n relative phase maps represented by a vector $\varphi = [\phi_1,\ \phi_2,\ \cdots,\ \phi_n]^T$ and characterized by the fringe wavelengths (fringe pitches) $\boldsymbol{\lambda} = [\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n]^T$, the corresponding unwrapped phase maps $\boldsymbol{\Phi} = [\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n]^T$ can be represented as

$$\boldsymbol{\Phi} = \varphi + 2\pi\mathbf{k} \tag{B.1}$$

where $\mathbf{k} = [k_1,\ k_2,\ \cdots,\ k_n]^T$ is the integer fringe order vector. Since TPU is performed pixel-wisely over the whole image, the pixel coordinate (x^c, y^c) is omitted here for simplicity. The task of phase unwrapping is to determine the fringe orders $\mathbf{k}$ from the knowledge of wrapped phase vector $\boldsymbol{\varphi}$ only, and the continuous phase maps $\boldsymbol{\Phi}$ can be recovered by Eq. (B.1). To achieve this goal, the fringe wavelengths λ should be properly chosen so that the phase $\boldsymbol{\varphi}$ can be successfully unwrapped without ambiguities within the desired measurement range [84]. This relies on the fact that given a set of fringe wavelengths $\boldsymbol{\lambda} = [\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n]^T$, their least common multiple $LCM(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n)$ determines the maximum range on the absolute phase axis within which each combinations for wrapped phase values are unique [85–87]. Considering the projection pattern has W pixels along the horizontal axis wherein the sinusoidal intensity value varies (the pixels in the same column all have equal intensity) the condition

$$LCM(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n) \geq W \tag{B.2}$$

should be satisfied to exclude ambiguity.

Given the projector coordinate $0 \leq x_p < W$ and the condition of Eq. (B.2) holds, all unwrapped phase values can be connected with the corresponding projector coordinate x_p through the following relation:

$$\boldsymbol{\Phi} \circ \boldsymbol{\lambda} = 2\pi x_p \equiv t \tag{B.3}$$

where $\circ$ is the Hadamard product (entrywise product). Eq. (B.3) suggests that the trajectory of continuous phase values $(\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n)$ form a straight line passing through the origin in dimension n . The direction vector of the line is $\boldsymbol{\lambda}^{-1} = [\frac{1}{\lambda_1},\ \frac{1}{\lambda_2},\ \cdots,\ \frac{1}{\lambda_n}]^T$ and can be parameterized by parameter t. For a given set of wrapped phase values $\varphi = [\phi_1,\ \phi_2,\ \cdots,\ \phi_n]^T$, the problem of TPU is recast to finding the integer fringe order vector $\mathbf{k} = [k_1,\ k_2,\ \cdots,\ k_n]^T$ so that the final unwrapped phase values $\boldsymbol{\Phi} = [\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n]^T$ calculated from Eq. (B.1) can precisely fall on the straight line described by Eq. (B.3). Since the ambiguity in the projector space is ruled out by condition Eq. (B.2), there should be only one qualified fringe order vector $\mathbf{k}$ within the range $0 \leq x_p < W$. For simplicity of explanation, we consider an example with

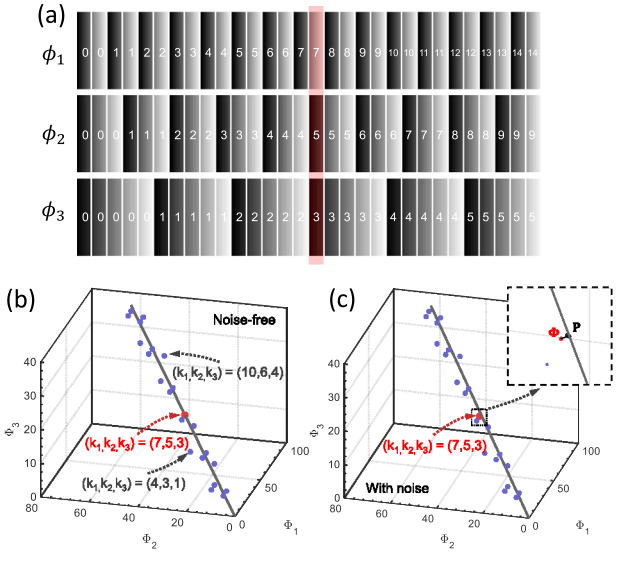

Fig. B.1. A simple example explaining the basic idea of PDM $[(\lambda_1,\ \lambda_2,\ \lambda_3)$=(2, 3, 5) pixels, and $LCM(\lambda_1,\ \lambda_2,\ \lambda_3) = 30$ pixels]. (a) Changes of three wrapped phase maps within the unambiguous range. (b) Calculated the unwrapped phase values of the red-shaded pixel with all possible combinations of fringe orders shown in 3D space described by $(\Phi_1,\ \Phi_2,\ \Phi_3)$. Only the correct fringe orders (7,5,3) produces unwrapped phase values (red dot) that can precisely fall on the straight line described by Eq. (B.3). (c) In the presence of noise, the calculated unwrapped phase values may never precisely fall on the straight line even when the fringe order vector is correct. But the optimum fringe orders should produce unwrapped phase values (red dot) that is closest to the straight line. The inset shows the magnified boxed region where the point **P** representing the projection of unwrapped phase values **Φ** onto the line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

three wrapped phase maps $\varphi = [\phi_1,\ \phi_2,\ \phi_3]^T$, characterized by the fringe pitches $\lambda_1 = 2$ pixels, $\lambda_2 = 3$ pixels, and $\lambda_3 = 5$ pixels. Fig. B.1(a) shows changes of three wrapped phase maps within the unambiguous range $LCM(\lambda_1,\ \lambda_2,\ \lambda_3)$= 30 pixels. In Fig. B.1(a), each small rectangular block represents one pixel, labeled by its fringe order. Careful observation reveals that for a given pixel coordinate, each combination of three wrapped phase values is unique, therefore the fringe orders of the three phase maps $\mathbf{k} = [k_1,\ k_2,\ k_3]^T$ can be uniquely determined by examining their wrapped phases. For example, considering the red shaded pixel with x^p= 16, we can calculate the corresponding unwrapped phase values with all possible combinations of fringe orders $\mathbf{k} = [k_1,\ k_2,\ k_3]^T$. Though there are 22 possible combinations of $(k_1,\ k_2,\ k_3)$ in this case: (0,0,0), (1,0,0), (1,1,0), (2,1,0), (2,1,1) , (3,2,1), ⋯, (14,9,5), only (7,5,3) produces the unwrapped phase values **Φ** that can precisely fall on the straight line described by Eq. (B.3) in 3D space, as shown in Fig. B.1(b).

However, in practice, there are many factors such as non-sinusoidal fringe intensity, random noise of the projector and the camera, object motion in the measurement process that may induce errors in obtained wrapped phase maps. In such cases, the unwrapped phase values $\mathbf{\Phi} = [\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n]^T$ calculated from Eq. (B.1) may not precisely fall on the straight line described by Eq. (B.3) for all possible fringe order vectors [even when the fringe order vector is correct, as illustrated in Fig. B.1(c)]. To solve the phase unwrapping problem in the presence of noise optimally, we should choose the fringe order which produces the unwrapped phase values $\mathbf{\Phi} = [\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n]^T$ that is closest to the straight line. Thus, the distance from the point $(\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n)$ to the line in Euclidean geometry should be calculated to quantify how close they are. We first project the point onto the line to get a projected point $\mathbf{P} = [P_1,\ P_2,\ \cdots,\ P_n]^T$, which should satisfy the following simultaneous equations

$$\begin{cases} \mathbf{P}\circ\boldsymbol{\lambda} \equiv \mathbf{t} \\ (\mathbf{P}-\mathbf{\Phi})^T\boldsymbol{\lambda}^{-1}=0 \end{cases} \tag{B.4}$$

The first line of Eq. (B.4) indicates that the projected point **P** must lie on the line described by Eq. (B.3)), and the second line of Eq. (B.4) suggests that the projected vector $\mathbf{P}-\mathbf{\Phi}$ must be perpendicular to the direction vector of the line. Solving Eq. (B.4) we obtain the parameter

$$t = \left(\left\|\boldsymbol{\lambda}^{-1}\right\|^2\right)^{-1}\left(\boldsymbol{\lambda}^{-1}\right)^T\mathbf{\Phi} = \left(\sum_{i=1}^{n}\left(\frac{1}{\lambda_i^2}\right)\right)^{-1}\sum_{i=1}^{n}\frac{\Phi_i}{\lambda_i} \tag{B.5}$$

and the projected point

$$\mathbf{P} = \boldsymbol{\lambda}^{-1}t \tag{B.6}$$

where $\|\|$ is the Euclidean norm. The distance d from the point to the line is just the distance between the point $\mathbf{\Phi} = [\Phi_1,\ \Phi_2,\ \cdots,\ \Phi_n]^T$ and the projected point $\mathbf{P} = [P_1,\ P_2,\ \cdots,\ P_n]^T$:

$$d^2 = \|\mathbf{P}-\mathbf{\Phi}\|^2=(\mathbf{P}-\mathbf{\Phi})^T(\mathbf{P}-\mathbf{\Phi}) \tag{B.7}$$

With all the above equations at hand, the basic procedures of the PDM algorithm can be summarized as follows:

Step 1. Choose a proper set of fringe wavelengths $\boldsymbol{\lambda} = [\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n]^T$ satisfying condition Eq. (B.2) to generate wrapped phase vector $\varphi = [\phi_1,\ \phi_2,\ \cdots,\ \phi_n]^T$ with unique phase values over the maximum projector coordinate W.

Step 2. Generate all fringe order vector $\mathbf{k}_i$ that contains all possible combinations of integer fringe orders$[k_1,\ k_2,\ \cdots,\ k_n]^T$ over the maximum projector coordinate W , as we did in Fig. B.1(a).

Step 3. For each fringe order vector $\mathbf{k}_i$, calculate the corresponding unwrapped phase values $\mathbf{\Phi}_i$ [Eq. (B.1)], projected unwrapped phase values $\mathbf{P}_i$ [Eq. (B.5) and (B.6)], and finally the distance d_i^2 between them [Eq. (B.7)]. (Note that if the approximate depth range of the measured scene can be estimated, geometric constraint [58,61–63] can be applied to limiting the search range of $\mathbf{k}_i$ and ruling out several false candidates beforehand. See **Appendix D2** for details)

Step 4. Select the fringe order vector which produces the minimum d_i^2 (denoted as $d_{\min}^2$) as the optimum solution $\mathbf{k}_{\text{opt}}$. Meanwhile, the corresponding unwrapped phase values $\mathbf{\Phi}_{\text{opt}}$ can be obtained [Eq. (S10)]. It should be also noted that the minimum projection distance map $d_{\min}^2$ in PDM reflects the unwrapping reliability for each pixel (larger $d_{\min}^2$ value means lower unwrapping reliability), which is further used in the subsequent fringe order error compensation algorithm (See **Appendix C** for details).

B2. Proof about optimality

Let us assume the additive phase noise model of the phase measurement

$$\varphi_m = \varphi + \Delta\varphi \tag{B.8}$$

where $\Delta\boldsymbol{\varphi}$ is zero-mean Gaussian distributed noise vector with a variance vector of $\boldsymbol{\sigma}^2$. This assumption is valid for typical image sensor in which thermal or shot noise is the main noise type. The importance of the Gaussian distribution arises from the fact that many distributions can be approximated by the Gaussian one. Furthermore, a combination of noise sources of the same kind usually behaves like a Gaussian noise source as a result of central limit theorem.

If the wrapped phase maps $\boldsymbol{\varphi}_m$ are recovered from the same phase retrieved algorithm (e.g., BNFTP or standard phase shifting algorithm with the same phase-shift step, which is true for most practical conditions), their phase error variance should be identical $\sigma_1^2=\sigma_2^2=\cdots=\sigma_n^2=\sigma^2$ (which is independent with fringe wavelength, see our recent TPU review paper [52] for detailed explanations). After phase unwrapping, the phase values in the phase maps are only modified by multiples of 2π, so the noise variances for all of unwrapped phase maps should be identical as well:

$$\mathbf{\Phi}_m = \mathbf{\Phi} + \Delta\varphi \tag{B.9}$$

where $\Delta\boldsymbol{\varphi}$ is zero-mean Gaussian distributed noise vector with a variance of σ^2 , which is same as the one in Eq. (B.8). For phase unwrapping

problem, since the true unwrapped phase vector is unknown, the fringe orders **k** can only be determined from the measured values. Suppose we have a group of unwrapped phase maps $\mathbf{\Phi}_m = [\Phi_{m1},\ \Phi_{m2},\ \cdots,\ \Phi_{mn}]^T$, each phase map Φ_{mi} actually corresponds to a statistically independent measurement of the same physical quantities. In most case, the phase value is used as a unique identifier of the projector pixel coordinate x_p [or equivalently, parameter t in Eq. (B.3)], which is related to the actual height/depth of the measured objects. So we can obtain n separate and independent (but inconsistent) estimates of the parameter t

$$\mathbf{t}_m = \mathbf{\Phi}_m \circ \boldsymbol{\lambda} = t + \Delta\mathbf{t} \tag{B.10}$$

where $\Delta\mathbf{t} = \boldsymbol{\lambda} \circ \Delta\varphi$ is an zero-mean Gaussian distributed derivation vector with a variance vector of $\lambda^2\sigma^2$. Based on the principle of maximum likelihood, the best estimate of t from these actual estimates is given by the following weighted average [88]

$$\hat{t}_{ML} = \frac{\mathbf{w}^T\mathbf{t_m}}{\|\mathbf{w}\|_1} = \frac{\sum_{i=1}^{n} w_i\lambda_i\Phi_{mi}}{\sum_{i=1}^{n}|w_i|} \tag{B.11}$$

where the weights are inverse squares of the uncertainties in actual estimates

$$w_i = \frac{1}{\lambda_i^2\sigma^2} \tag{B.12}$$

The rationale of Eq. (B.10) can be understood based on the fact that using denser fringe patterns (smaller wavelengths) generally leads to more sensitive measurements and lower uncertainties, and hence the weights for smaller wavelengths should be larger. Substituting Eq. (B.12) into (B.11) and after simple reduction, it can be verified that Eq. (B.11) is exactly the same as Eq. (B.5). This exactly proves that the proposed PDM algorithm always produces the optimum fringe order pair so that the corresponding unwrapped phase values are closest to the maximum likelihood estimates.

B3. Comparison with heterodyne phase unwrapping

Similar to many other TPU algorithms, the PDM solves the phase ambiguity problem by employing multiple phase measurements with different frequencies. Compared with the state-of-the-art algorithms[52], it not only guarantees to give an optimal solution in a maximum likelihood sense but also provides an inherent metric $d_{\min}^2$ to evaluate the unwrapping reliability quantitatively (which is used in the reliability guided compensation algorithm. See **Appendix C** for details). To better understand the advantage of the PDM algorithm, here we compare it with the classic heterodyne approach [89,90], which is probably the most widely used TPU algorithm especially for the three-wavelength case [53,91]. Let us first consider the two-frequency scenario with the wavelengths $\lambda_1 < \lambda_2 < 2\lambda_1$. The heterodyne approach extends the unambiguous phase range to the synthetic wavelength λ_{12} by taking the difference of the phase measurements taken at each wavelength:

$$\lambda_{12} = \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} \tag{B.13}$$

However, for the proposed PDM, it is easy to verify that its unambiguous measurement range is always no less than that of the heterodyne approach

$$LCM(\lambda_1, \lambda_2) \geq \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} \tag{B.14}$$

The conclusion can be easily extended to the cases of three or even more wavelengths. Thus, the proposed PDM has a larger unwrapping range than classic heterodyne approach. Another advantage of the proposed approach lies in fact that it is less sensitive to noise, and thus can provide higher unwrapping reliability (success rate). For space reasons, here we only support our claim by simulation and experiment (see **Appendices B4 and B5**). More detailed theoretical analysis about the noise tolerance and the optimum choice of wavelengths will be presented in future publications.

B4. Comparison by simulations

In this section, the performance of PDM and heterodyne phase unwrapping algorithms is compared by numerical simulations. Two different situations will be considered in these simulations. The smooth surface (phase map 1, generated from the MATLAB built-in function "peaks") and a discontinuous surface (phase map 2) with sharp abrupt edges. The true wrapped and continuous phase maps (1024 × 1024 pixels) are shown in Fig. B.2(a) and (d), respectively.

Firstly, three different wavelengths $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$ are used in the simulations, and the true continuous phases are shown in Fig. B.2(b) and (e). These continuous phases are corrupted by normally distributed random noise with a mean value of zero and a standard derivation of 0.04 and then wrapped to the range of $[-\pi, \pi)$, as shown in Fig. B.2(c) and (f). Figs. B.3 and B.4 show the phase unwrapping re-

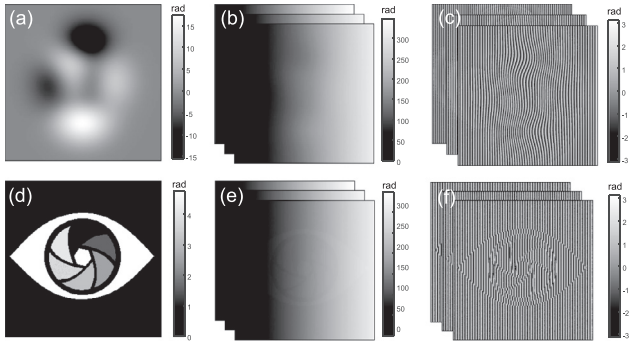

Fig. B.2. Simulated of continuous phase maps and wrapped phase maps. (a) Phase map 1. (b) True continuous phase maps when $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$. (c) Wrapped phase maps with noise (phase noise variance 0.04^2). (d) Phase map 2. (e) True continuous phase map when $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$. (f) Wrapped phase maps with noise (phase noise variance 0.04^2).

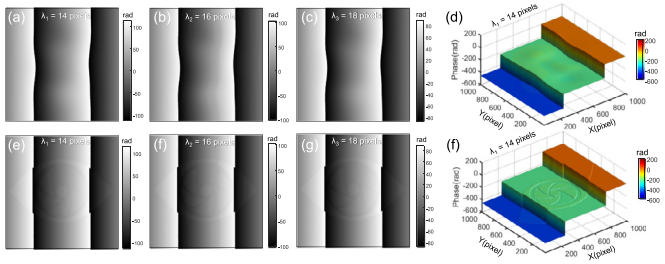

Fig. B.3. Phase unwrapping results of three-wavelength heterodyne method $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$. (a)–(c) Unwrapped phase maps of Phase 1. (d) 3D plot of unwrapping result with $\lambda_1 = 14$. (e)–(g) Unwrapped phase maps of Phase 2. (f) 3D plot of unwrapping result with $\lambda_1 = 14$. Note for (d) and (f), the linear carrier phase (phase ramp) is removed from each result to get a better view.

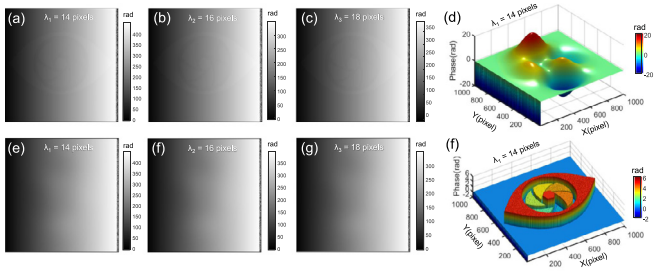

Fig. B.4. Phase unwrapping results of three-wavelength PDM method $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$. (a)–(c) Unwrapped phase maps of Phase 1. (d) 3D plot of unwrapping result with $\lambda_1 = 14$. (e)–(g) Unwrapped phase maps of Phase 2. (f) 3D plot of unwrapping result with $\lambda_1 = 14$. Note for (d) and (f), the linear carrier phase (phase ramp) is removed from each result to get a better view.

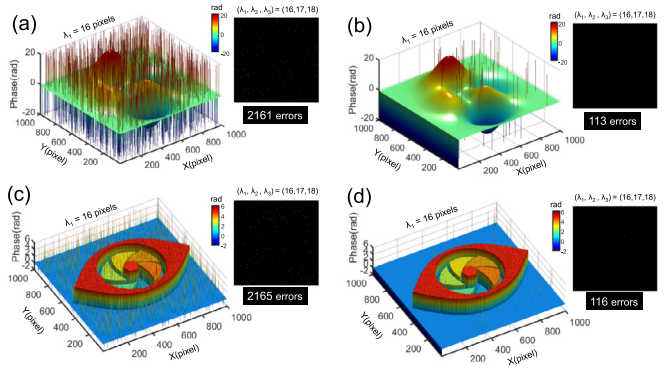

Fig. B.5. Phase unwrapping results of three-wavelength heterodyne and PDM methods $(\lambda_1,\ \lambda_2,\ \lambda_3) = (16, 17, 18)$. (a) 3D plot of unwrapping result of Phase 1 with $\lambda_1 = 16$ using heterodyne method. (b) 3D plot of unwrapping result of Phase 1 with $\lambda_1 = 16$ using PDM method. (c) 3D plot of unwrapping result of Phase 2 with $\lambda_1 = 16$ using heterodyne method. (d) 3D plot of unwrapping result of Phase 2 with $\lambda_1 = 16$ using PDM method. The insets illustrate the error maps (black correct, white- wrong) and number of errors (excluding edge regions) for corresponding unwrapping results.

sults of heterodyne and PDM algorithms, respectively. Note that the unwrapped phase maps with $\lambda_1 = 14$ from each group of results are further plotted in 3D in Figs. B.2 and B.3, where the linear carrier phase (phase ramp) is removed from each result to get a better view. It can be seen that there remain several phase discontinuities in the unwrapping results of heterodyne approach. This is due to the fact that the $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$ only leads to synthetic wavelength $\lambda_{123} = 504 < 1{,}000$ pixels, so the heterodyne method cannot unwrap the whole range of phase successfully. However, for PDM, since the unambiguous measurement range can be extended to $LCM(\lambda_1,\ \lambda_2,\ \lambda_3) = 1{,}008 > 1{,}000$ pixels, all phase maps can be successfully unwrapped, and the 3D surfaces are correctly reconstructed. Note that there are still some unwrapping errors noticeable in the right edge regions of the unwrapped phase maps. This kind of edge error is normal for TPU algorithms and only influences the edge regions of the continuous phase without spoiling the whole measurement [52]. Besides, this kind of error can be avoided by using a smaller field of view of the projected patterns for actual measurements. Excluding the limited edge parts (20 pixels in width), the success rates of the two PDM unwrapped results are both 100% in this simulation.

The second simulation compares the performance of PDM and heterodyne with a different wavelength setting $(\lambda_1,\ \lambda_2,\ \lambda_3) = (16, 17, 18)$ and with other parameters in the previous simulation unchanged. Under this condition, the theoretical unambiguous measurement ranges of heterodyne and PDM are the same and both larger than the image dimension ($\lambda_{123} = LCM(\lambda_1,\ \lambda_2,\ \lambda_3) = 2{,}448 > 1{,}000$ pixels). Fig. B.5 compares phase unwrapping results of heterodyne and PDM. The number of errors (excluding edge regions) as well as error distribution maps (black correct, white wrong) for corresponding unwrapping results are also provided in the insets of Fig. B.5. It can be seen that the PDM approach provides higher success rate and fewer phase unwrapping errors. Only about 110 out of the totally pixels are not correctly unwrapped. In contrast, unwrapping errors are far more prevailing in the results of heterodyne approach, with more than 2,100 pixels are contaminated by fringe order errors [Fig. B.5(a)–(c)].

From the above two simulations, it can be found that the proposed PDM approach offers extended unambiguous measurement range compared to classic heterodyne approach using the same wavelengths. It also offers better resistance to noise and higher unwrapping reliability. By carefully comparing the results of Figs. B.4 and B.5, it is also found that the unwrapping reliability of PDM is closely related to the selection of wavelengths. Simply changing the wavelengths $(\lambda_1,\ \lambda_2,\ \lambda_3)$ from (14, 16, 18) to (16, 17, 18) leads to obvious decline in the noise-robustness. This means that there does exist certain configurations of wavelengths which is more efficient, in terms of noise immunity, than that of other settings. In practice, the optimal set of wavelengths should depend on the noise levels, resolutions of the projector and the camera, scenes albedos, and the degree of lens defocusing. We cannot discuss this point in detail for lack of space. But a simple, empirical procedure is suggested here for obtaining a "reasonably good" (but not necessarily yields the optimal) set of wavelengths for the μFTP measurement.

(1) Determine a wavelength range so that the wavelength is sufficient small for BNFTP phase retrieval but can achieve both good fringe visibility (modulation) and sinusoidity with a slight defocused projector.

(2) Select each wavelength within the determined range properly so that $LCM(\lambda_1,\ \lambda_2,\ \lambda_3)$ is larger but close enough to the required ambiguous measurement range (the projector resolution).

B5. Experiment on a complicated scene

To demonstrate the performance of the proposed PDM approach, we measure a complicated scene which contains a plaster model on the left and a human hand separately on the right. The wavelength set $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$ is used in the experiment, and the unambiguous range is $LCM(\lambda_1,\ \lambda_2,\ \lambda_3) = 1008$ pixels. Though this value is a bit smaller than the $W = 1024$ pixels of the projector used, it is sufficient to exclude phase ambiguity especially when a smaller field of view of the projected patterns is used for actual measurements, and the geometric constraint is imposed to restrict the search range of fringe orders (see **Appendix D2** for details).

Fig. B.6(a1)–(a4) show the three fringe images with different wavelengths and the white image captured by the μFTP system (camera exposure 46 μs, 20,000 fps, see Materials and methods for details). The corresponding wrapped phase maps recovered by BNFTP algorithm (see **Appendix A** for details) are shown in Fig. B.6(b1)– (b3). These three

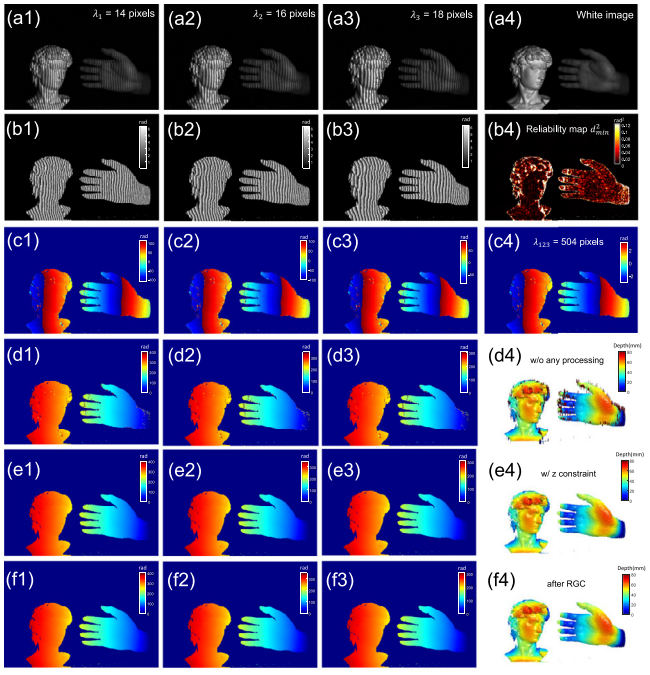

Fig. B.6. Measurement results of a complicated scene using three-wavelength PDM $(\lambda_1,\ \lambda_2,\ \lambda_3) = (14, 16, 18)$. (a1)–(a3) Fringe images with different wavelengths. (a4) White image. (b1)–(b3) Wrapped phase maps recovered by BNFTP algorithm. (b4) Minimum projection distance map (reliability map $d^2_{\min}$). (c1)–(c3) Phase unwrapping results of the classic heterodyne approach. (c4) Synthetic phase map Φ_{123} obtained by taking the differences of the phase measurements taken at each wavelength. (d1)–(d3) Phase unwrapping results of PDM. (d4) 3D reconstruction based on the unwrapped phase map corresponding to λ_2. (e1)–(e4) are the results when the geometric constraint is imposed to restrict the search range of fringe orders in Step 3 of PDM algorithm. (f1)–(f4) are the final results after the reliability guided fringe order error compensation algorithm.

wrapped phases are unwrapped by the classic heterodyne approach and the proposed PDM. The minimum projection distance map ($d^2_{\min}$) for each pixel generated by PDM is shown in Fig. B.6(b4). Fig. B.6(c1)–(c3) show the phase unwrapping results of heterodyne approach, and Fig. B.6(c4) is the synthetic phase map Φ_{123} obtained by taking the differences of the phase measurements taken at each wavelength. Phase discontinuities can obviously seen in the results as expected because the heterodyne method cannot unwrap the whole phase range (synthetic wavelength $\lambda_{123} = 504 < 1{,}000$ pixels), which is consistent with the simulation results shown in Fig. B.3. Fig. B.6(d1)–(d3) show the phase unwrapping result of PDM, and Fig. B.6(d4) is the 3D reconstruction based on the unwrapped phase map Φ_2(see **Appendix D2** for details). It can be seen from Fig. B.6(d) that most points (more than 98% of pixels) are correctly unwrapped through the PDM approach. The limited fringe order errors are mostly concentrated on the dark regions and object edges where the minimum projection distance map $d^2_{\min}$ is large [the bright regions of Fig. B.6(b4)]. These fringe order errors can be greatly reduced by imposing the geometric constraint to restrict the search range of fringe orders in Step 3 of PDM algorithm and finally completely eliminated by a subsequent reliability guided fringe order error compensation algorithm, which will be introduced in the next section.

Appendix C. Reliability Guided Compensation (RGC) of Fringe Order Error

C1. Theory

Though PDM can provide optimal unwrapping performance even under noisy conditions, fringe order errors are still inevitable especially around dark regions and object edges where the fringe quality is low [Fig. B.6(d)]. The problem is even worse for high speed imaging of transient events due to the very limited exposure time and nonnegligible frame-by-frame object motion. Here we propose an approach so called RGC for identifying and compensating those fringe order errors by exploiting additional information in spatial domain. The rationale of the approach is based on the fact that the fringe order errors are usually isolated (at least less concentrated than the correct phase values) delta-spike artifacts with a phase error of integral multiples of 2π. Inspired by the quality guided (spatial) phase unwrapping [75–77], we first gather contiguous pixels within a continuous region of the phase map into groups. Then the isolated pixels or pixels falling into the small groups are considered as fringe order errors, and their phase values will be corrected with respect to the phase value of the adjacent pixel within a larger group, according to an order ranked by a predefined reliability function.

There are two main issues in the RGC algorithm: the choice of the reliability function and the design of processing path. In PDM, temporal phase unwrapping is recast as an optimization problem of choosing the fringe orders so that the resulting unwrapped phase value combination $\mathbf{\Phi} = [\Phi_1, \Phi_2, \cdots, \Phi_n]^T$ is closest to the straight line described by Eq. (B.3). Therefore, the minimum projection distance map $d^2_{\min}$ for each pixel calculated from Eq. (B.7) is used to evaluate the reliability of phase unwrapping [larger $d^2_{\min}$ value means lower reliability, as demonstrated in Fig. B.6(b4)]. It should be noted that other criterions, like the fringe modulation and phase gradient, can also serve as the reliability function when $d^2_{\min}$ is unavailable (when a different TPU approach is employed). The processing path is determined by comparing the value of the reliability of the edges instead of pixels [75]. The reliability of an edge is defined as the summation of the reliability of the two pixels connected by the edge. The edges are stored in an array and sorted by the value of reliability. The edges with higher reliability are resolved first.

Fig. C.1 presents the flow chart of the RGC algorithm as well as an example explaining its basic principle. At the preparation stage, the reliability value of each edge is constructed based on the distance map, as illustrated in Fig. C.1(a). Then, we go through all the edges in the

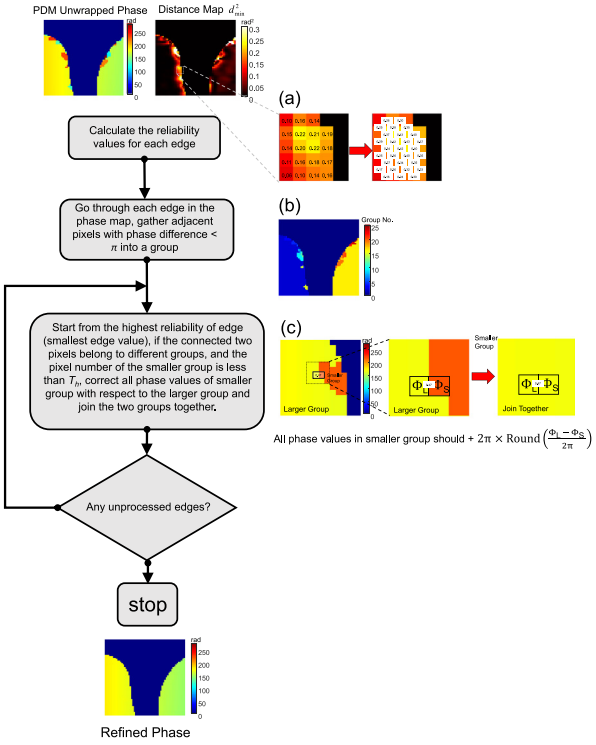

Fig. C.1. Flow chart of the RGC algorithm as well as an example explaining its basic principle. (a) Construct the reliability value based on the distance map $d^2_{\min}$. The reliability value of each edge is defined as the summation of the reliability of the two pixels connected. (b) Go through all the edges in the phase map and gather contiguous pixels within a continuous region of the PDM unwrapped phase map into groups. (c) Isolated pixels or pixels falling into the small groups are considered as fringe order errors, and their phase values will be corrected with respect to the phase value of the adjacent pixel within a larger group.

phase map and join adjacent pixels with phase difference less than π into a group. After this step, all contiguous pixels within a continuous region of the PDM unwrapped phase map can be gathered into the same group, as illustrated in Fig. C.1(b). The third step of RGC algorithm is to correct the fringe order error according to an order ranked by the value of reliability (edges with higher reliability or smaller edge value are processed first). For a given edge being processed, if the connected two pixels belong to different groups, and the number of pixel in the smaller group is less than a predefined threshold T_h (a typical value of $T_h = 200$), we need to correct all phase values of smaller group with respect to the larger group and join the two groups together. As shown in Fig. C.1(c), if the edge with the reliability value 0.47 is currently being processed, and this edge connects two pixels with phase values of Φ_L and Φ_S, which belong to two different groups. Then the $Round\left(\frac{\Phi_L - \Phi_S}{2\pi}\right)$ multiples of 2π value is added to all the pixels in the group that contains the smaller number of pixels, and the two groups are joined together. This step is repeated until all the edges are processed. Eventually, the fringe order errors can be effectively compensated.

C2. Experiments

In Fig. B.6(e), we have already demonstrated that the fringe order error in PDM unwrapped phase can be completely corrected by the RGC algorithm. However, since the measured object is quasi-static, the phase unwrapping errors are very limited, especially when the geometric constraint is applied. Thus, in Fig. C.2, we further present two experimental results to demonstrate the validity of RGC algorithm when measuring fast changing objects. The measured objects are a rotating fan and a bursting balloon punctured by a flying dart (more details about the experimental configurations can be found in Sections **Rotating fan**

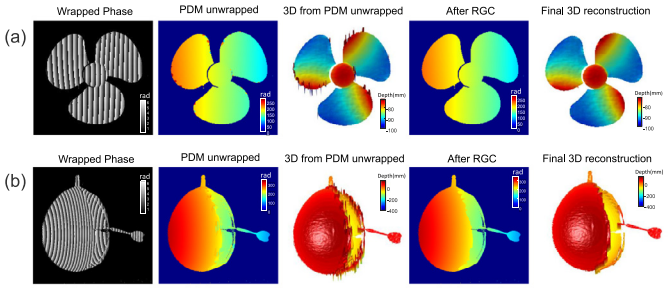

Fig. C.2. Experimental results of RGC algorithm when measuring dynamically changing objects. (a) Rotating fan blades. (b) A bursting balloon punctured by a flying dart. The images from left to right are: wrapped phases recovered by BNFTP, phase unwrapped by PDM (with geometric constraint), 3D reconstructions from the PDM unwrapped phases, phases refined by RGC, and final 3D reconstructions from the RGC refined phases, respectively.

blades and **Balloons explosion triggered by a flying dart**). As shown in Fig. C.2, the object motion introduces more prominent fringe order errors, and the geometric constraint is inadequate to suppress them. However, these fringe order errors are successfully corrected by the RGC algorithm, resulting in high-quality 3D reconstructions without any delta-spike artifacts.

Appendix D. System Calibration, Geometric Constraints, and 3D Coordinate Mapping

D1. Mapping from phase to 3D coordinates

Once the final unwrapped phase image is obtained after RGC, we can establish a unique correspondence value x^p for every camera pixel (x^c, y^c) through the linear equation

$$\Phi(x^c, y^c) = \frac{2\pi}{\lambda} x^p \tag{D.1}$$

Based on the pin-hole model of the imaging lens, the relationship between the 3D world coordinates of the measured object (x^w, y^w, z^w) and the 2D camera pixel coordinates (x^c, y^c) can be described as

$$s\begin{bmatrix} x^c \\ y^c \\ 1 \end{bmatrix} = \mathbf{A}^c\left[\mathbf{R}^c, \mathbf{t}^c\right]\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} = \mathbf{P}^c\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} = \begin{pmatrix} p^c_{11} & p^c_{12} & p^c_{13} & p^c_{14} \\ p^c_{21} & p^c_{22} & p^c_{23} & p^c_{24} \\ p^c_{31} & p^c_{32} & p^c_{33} & p^c_{34} \end{pmatrix}\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} \tag{D.2}$$

Here superscript c denotes the camera, s is a scaling factor, $[\mathbf{R}^c, \mathbf{t}^c]$ is the 3×4 extrinsic parameter matrix representing the rotation-translation from the world coordinate system to the camera coordinate system. $\mathbf{A}^c$ is the 3×3 intrinsic parameter matrix of the camera. $\mathbf{P}^c$ is the 3×4 perspective matrix, which is just the product of extrinsic parameter matrix and the intrinsic parameter matrix of the camera. Since the projector has exactly the same mathematical model as the camera, similar relationship can be established between the 3D world coordinates of the measured object (x^w, y^w, z^w) and the 2D projector pixel coordinates (x^p, y^p)

$$s\begin{bmatrix} x^p \\ y^p \\ 1 \end{bmatrix} = \mathbf{A}^p\left[\mathbf{R}^p, \mathbf{t}^p\right]\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} = \mathbf{P}^p\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} = \begin{pmatrix} p^p_{11} & p^p_{12} & p^p_{13} & p^p_{14} \\ p^p_{21} & p^p_{22} & p^p_{23} & p^p_{24} \\ p^p_{31} & p^p_{32} & p^p_{33} & p^p_{34} \end{pmatrix}\begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix} \tag{D.3}$$

Here superscript p denotes the projector. The meaning of the parameters in Eq. (D.3) are the same as those in Eq. (D.2) but for the DLP projector. If both projector and camera are calibrated under the same world coordinate system, Eq. (D.1)–(D.3) can be combined to obtain the 3D world coordinates (x^w, y^w, z^w) for each camera pixel (x^c, y^c) :

$$\begin{bmatrix} x^w \\ y^w \\ z^w \end{bmatrix} = \begin{pmatrix} p^c_{11} - p^c_{31}x^c & p^c_{12} - p^c_{32}x^c & p^c_{13} - p^c_{33}x^c \\ p^c_{21} - p^c_{31}y^c & p^c_{22} - p^c_{32}y^c & p^c_{23} - p^c_{33}y^c \\ p^p_{11} - p^p_{31}x^p & p^p_{12} - p^p_{32}x^p & p^p_{13} - p^p_{33}x^p \end{pmatrix}^{-1} \times \begin{bmatrix} p^c_{34}x^c - p^c_{14} \\ p^c_{34}y^c - p^c_{24} \\ p^p_{34}x^p - p^p_{14} \end{bmatrix} \tag{D.4}$$

Eq. (D.4) can be effectively implemented based on lookup tables (LUTs) if the matrix inversion and multiplication operations are expended and pre-calculated for each pixel [54]. After simplification, the mapping of camera pixel to world coordinates can be represented as

$$z^w(x^c, y^c) = L_{z1}(x^c, y^c) + \frac{L_{z2}(x^c, y^c)}{L_{z3}(x^c, y^c) \cdot x^p + 1} \tag{D.5}$$

$$x^w(x^c, y^c) = L_{x1}(x^c, y^c)z^w(x^c, y^c) + L_{x2}(x^c, y^c) \tag{D.6}$$

$$y^w(x^c, y^c) = L_{y1}(x^c, y^c)z^w(x^c, y^c) + L_{y2}(x^c, y^c) \tag{D.7}$$

where $L_{[\bullet]}$ are pre-calculated LUTs (constant parameters) for each camera pixel coordinate.

D2. Geometric constraints in phase unwrapping

On a different note, the simple linear 3D coordinate mapping process can be reverse-applied if the measurement volume or the approximate depth range of the measured scene can be estimated beforehand according to the fringe quality, or the size of the measured object. The geometric constraint [58,61–63] allows to reduce the search space for fringe order combinations in PDM and thus, decrease the possibility of fringe order errors. For example, if the depth of interest is within the range $[z^w_{\min}, z^w_{\max}]$, we can invert Eq. (D.5) to find the corresponding valid range of absolute phase $[\Phi_{\min}, \Phi_{\max}]$ for each camera pixel:

$$\Phi_{\min}(x^c, y^c) = \frac{2\pi}{\lambda}\left\{\frac{L_{z2}(x^c, y^c)}{L_{z3}(x^c, y^c)\left[z^w_{\min} - L_{z1}(x^c, y^c)\right]} - \frac{1}{L_{z3}(x^c, y^c)}\right\} \tag{D.8}$$

$$\Phi_{\max}(x^c, y^c) = \frac{2\pi}{\lambda}\left\{\frac{L_{z2}(x^c, y^c)}{L_{z3}(x^c, y^c)\left[z^w_{\max} - L_{z1}(x^c, y^c)\right]} - \frac{1}{L_{z3}(x^c, y^c)}\right\} \tag{D.9}$$

Then the valid range of the fringe order $k \in [k_{\min}, k_{\max}]$ for each camera pixel can be determined:

$$k_{\min}(x^c, y^c) = floor\left[\frac{\Phi_{\min}(x^c, y^c)}{2\pi}\right] \tag{D.10}$$

$$k_{\max}(x^c, y^c) = ceil\left[\frac{\Phi_{\max}(x^c, y^c)}{2\pi}\right] \tag{D.11}$$

where the *floor*[.] (*ceil*[.*toolbox*]) function returns the largest (smallest) integer less (greater) than or equal to the specified numeric expression. Eqs. (D.10) and (D.11) can be incorporated into the Step 3 of PDM algorithm to limit the search range of fringe order combinations $\mathbf{k}_i$ and rule out several false candidates beforehand. As the example shown in Fig. B.6(c), if no depth constraint is applied in PDM, most of the 3D point clouds reconstructed from the wrong candidates (fringe order error) will fall outside of the measurement volume, exhibiting very large depth deviation. Fig. B.6(d) shows the results when the depth constraint

Table D.1
Experimentally Calibrated Internal, External and Distortion Parameters of μFTP system.

Index	Camera	Projector
Focal length (pixels)	[1182.245 1180.819]	[1881.674 1880.710]
Principal point (pixels)	[331.579 230.428]	[536.013 356.105]
Skew coefficient	0.00205	0.00116
Distortion coefficients	[−0.0858 0.1837 0.00047 −0.00112]	[−0.02421 −0.1305 0.00149 0.00280]
Rotation matrix	[0.995 0.0096 0.0991 0.0064 −0.999 0.0321 0.0993 −0.0313 −0.9946]	[0.984 0.0059 −0.178 0.0115 −0.999 0.0304 −0.177 −0.0320 −0.984]
Translation vector	[24.369 −21.189 808.857]	[31.207 11.652 806.228]

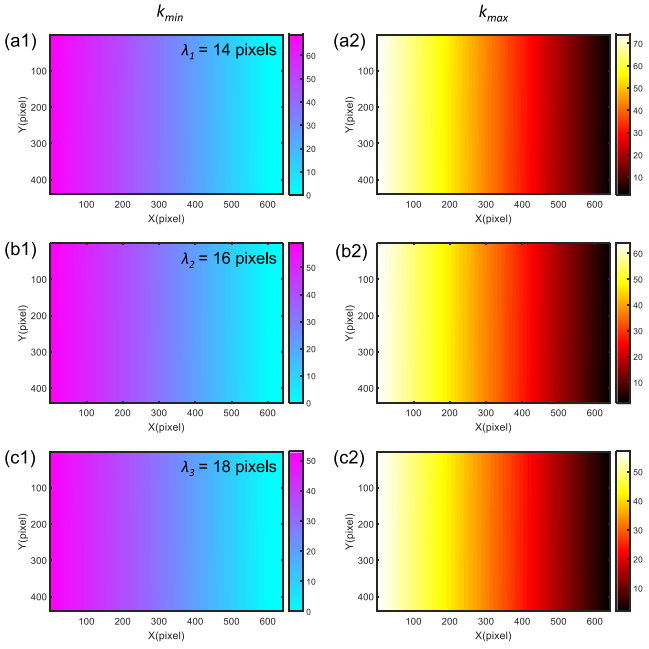

Fig. D.1. Valid range of the fringe order $[k_{\min}, k_{\max}]$ calculated based on the calibration parameters of μFTP system when the measurement depth is limited to the range $[z^w_{\min}, z^w_{\max}] = [0\text{ mm}, 100\text{ mm}]$. (a) Fringe wavelength $\lambda_1 = 14$ pixels. (b) Fringe wavelength $\lambda_2 = 16$ pixels. (c) Fringe wavelength $\lambda_3 = 18$ pixels.

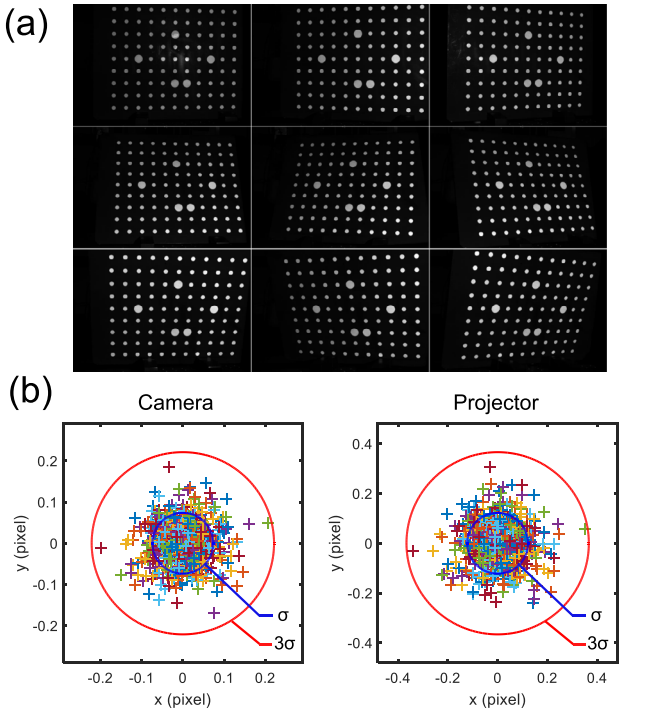

Fig. D.2. Calibration of μFTP system. (a) 9 images of the calibration board with different poses. (b) Reprojection error distributions of the camera and the projector.

is applied in the PDM ($[z^w_{\min}, z^w_{\max}] = [0\text{ mm}, 100\text{ mm}]$). Based on the calibration parameters of the μFTP system shown in Table. D.1, the valid range of the fringe order $[k_{\min}, k_{\max}]$ for different wavelengths are calculated, as shown in Fig. D.1. The results demonstrate that the geometric constraint significantly shrinks the search space of fringe orders in PDM, and considerable portion of fringe order errors can be rectified.

D3. System calibration

One important issue in the above-mentioned linear 3D coordinate mapping approach is that the projection imaging distortion of lenses used in the system is not considered. When determining the valid range of the fringe orders, the effect of lens distortion is minimum and can be simply ignored. However, it must be involved in the final 3D point cloud mapping process to give a result with higher measurement accuracy. In this work, 4 coefficients (two radial distortion parameters and two tangential parameters) are used to describe the distortion of the projector and camera lenses, and these coefficients can be obtained in conjunction with the intrinsic and extrinsic parameters during the calibration. Due to the lens distortion, the imaging points will deviate from their ideal locations, thus the correspondence between the distorted x^p_{dis} and camera pixel $\left(x^c_{dis}, y^c_{dis}\right)$ based on the absolute phase value should not be directly used for 3D reconstruction. Instead, we need to transform them to the undistorted coordinates(x^p and (x^c, y^c)) based on the calibrated distortion coefficients. Since the analytical inversion of the lens distortion model (inverse mapping relationship from the distorted coordinate to the undistorted coordinate) does not exist, an iterative scheme [92] is used to find the correspondence between the undistorted camera coordinate $\left(x^c, y^c\right)$ and the projector coordinate x^p , and the distortion-free 3D world coordinates (x^w, y^w, z^w) can then be obtained through Eq. (D.4). Note that the undistorted camera coordinates can be pre-calculated and stored in a LUT, while the projector distortion needs to be corrected based on the phase value and can only be performed at runtime. For more details about the implementation, one can refer to the **MATLAB source code** available on our website [78].

In the system calibration procedure, a calibration board with a white 9 × 11 circle array distributed uniformly on a black background is used. The distance between the centers of each circles is 30 mm. We capture 9 images of the calibration board with different poses to get a calibration volume of ∼ 400 mm × 275 mm × 400 mm, as shown in Fig. D.2(a). The projector is calibrated as an inverse camera with the help of the calibrated camera. We project two orthogonal sets of frequency-varying (fringe periods 1, 8, 64 in the x direction and 1, 8, 48 in the y direction) and phase-shifted (16-frame for the highest frequency) sinusoidal fringe patterns onto the calibration board. The absolute phase recovered by phase-shifting and multi-frequency temporal phase unwrapping techniques is used to establish a one-to-one mapping between a camera pixel and a DMD pixel. The intrinsic, extrinsic, and distortion parameters of the projector and camera are calibrated based on MATLAB Calibration

toolbox from J. Bougue [93], and optimized with bundle adjustment (BA) [94,95] in which both measurement errors and the fabrication errors of the calibration board are taken into account. The final internal, external and distortion parameters of the camera and projector are shown in Table. D.1, and the RMS of the resultant reprojection error is (0.053, 0.051) pixels for the camera and (0.087, 0.085) pixels for the projector, as shown in Fig. D.2(b).

Appendix E. Supplementary Videos

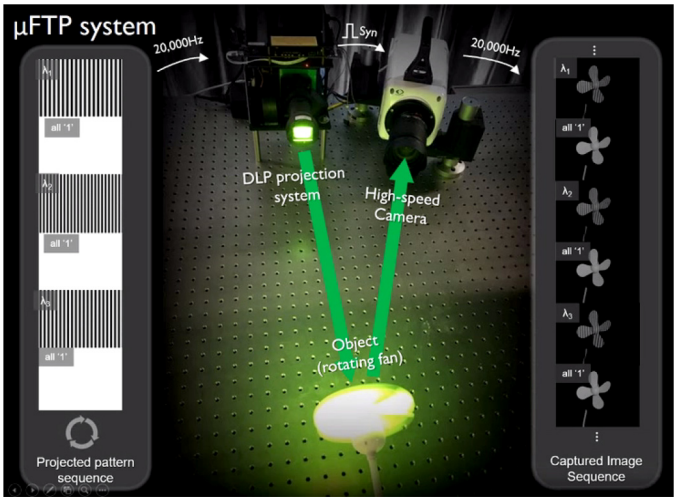

Fig. E.1. Supplementary Video 1: Projection and capture synchronization of μFTP system. The designed pattern sequences are projected on a dynamic scene (a rotating desk fan) sequentially and periodically at 20,000 Hz. The corresponding images are precisely captured by a synchronized high-speed camera.

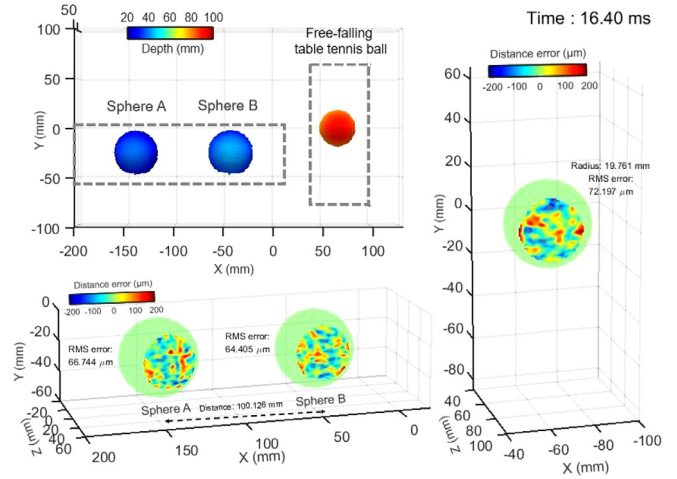

Fig. E.2. Supplementary Video 2: Color-coded 3D reconstruction of the test scene and the corresponding error analysis over the 41 ms period [see also Fig. 5(b)–(d)]. The test scene consists of two standard spheres and a free-falling table tennis ball. The color-coded 3D reconstruction is shown on the top left corner. The error distributions of enlarged areas corresponding to the dashed-boxes are shown on the bottom left and right.

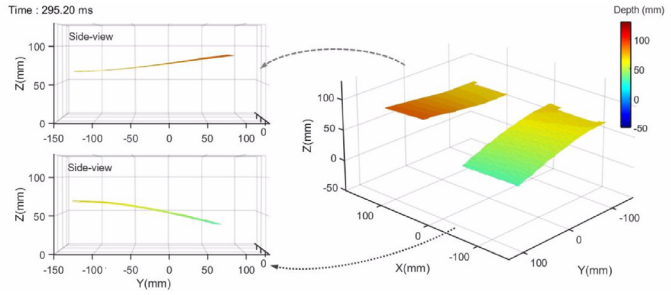

Fig. E.3. Supplementary Video 3: Color-coded 3D rendering of the two vibrating cantilevers surfaces and the corresponding side-views over the whole 986.4 ms period [see also Fig. 6(d)].

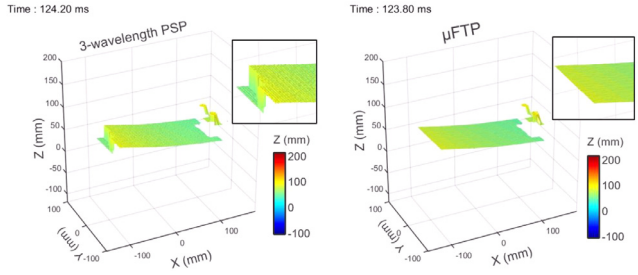

Fig. E.4. Supplementary Video 4: Comparison of μFTP with three-wavelength PSP based on a vibrating cantilever. μFTP is completely immune to motion ripples that are inevitable in conventional multi-frequency phase-shifting based techniques [see also Fig. 7(c)-(d)].

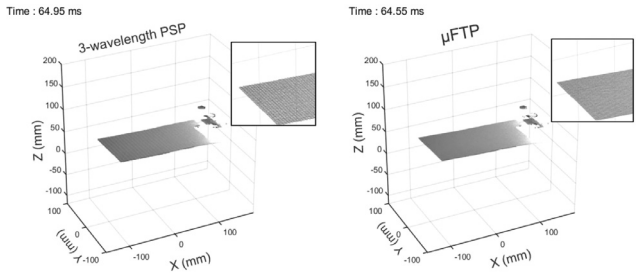

Fig. E.5. Supplementary Video 5: Comparison of μFTP with three-wavelength PSP based on a vibrating cantilever. The 3D point clouds are visualized with texture mapping.

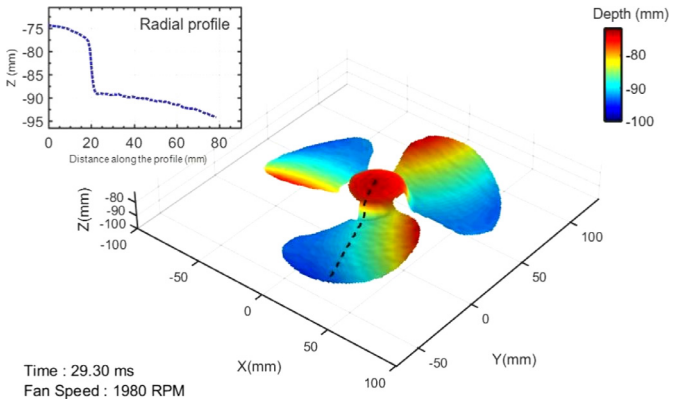

Fig. E.6. Supplementary Video 6: Color-coded 3D rendering of the rotating fans surface and the evolution of the corresponding radical profile along the dash line for approximately two full rotating cycles (60 ms) [see also Fig. 9(d)-(e)].

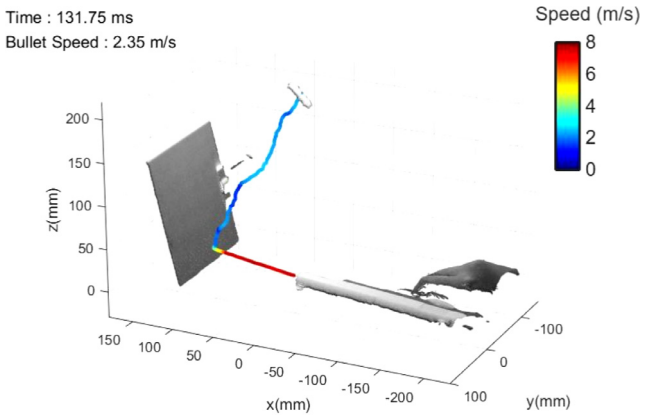

Fig. E.7. Supplementary Video 7: Temporal evolution of the bullet trajectory and the 3D point clouds of the scene over the 135 ms period. The bullet velocity at each track point is encoded by the line color [see also Fig. 10(d)].

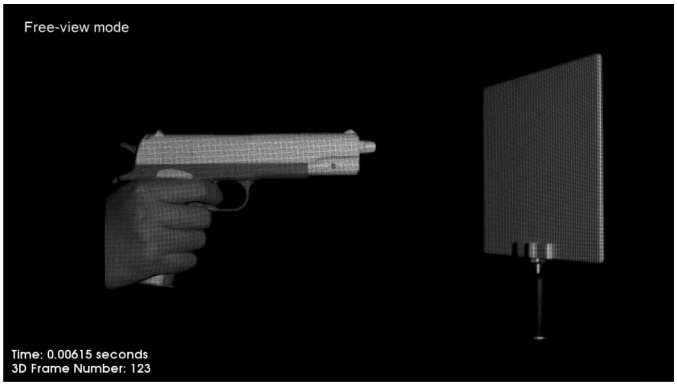

Fig. E.8. Supplementary Video 8: Slow-motion 3D movie of the transient event: a bullet fired from a toy gun and rebounded from a plaster wall. The video contains 2700 3D frames with a frame interval of 50 μs (corresponding to a frame rate of 20,000 fps over an observation period of 135 ms).

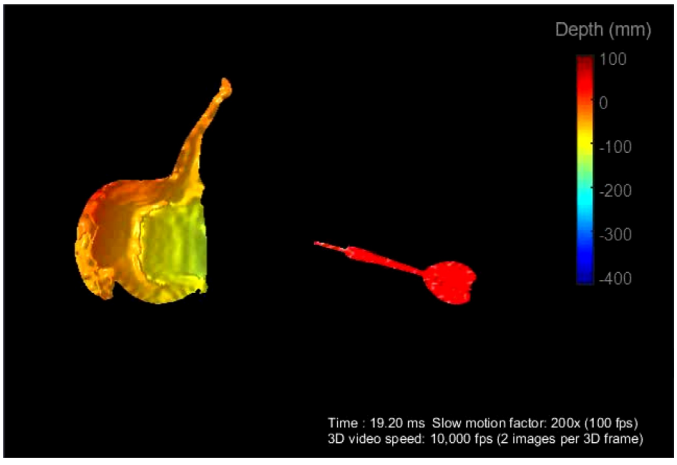

Fig. E.9. Supplementary Video 9: Air balloon bursting punctured by a flying dart. The 2D images are directly taken by the camera (corresponding to the white patterns) at 20,000 fps. The 3D reconstruction speed is 10,000 fps (2 images per 3D frame) [see also Figs. 11(a)–11(d)].

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.optlaseng.2017.10.013.

References

[1] Lunn GH. High-speed photography. Nature 1981;291:617–19.
[2] Field J. High-speed photography. Contemp Phys 1983;24(5):439–59.
[3] Fuller P. An introduction to high speed photography and photonics. Imaging Sci J 2009;57:293–302.
[4] Kleinfelder S, Lim S, Liu X, El Gamal A. A 10000 frames/s cmos digital pixel sensor. IEEE J Solid-State Circuits 2001;36(12):2049–59.
[5] El-Desouki M, Jamal Deen M, Fang Q, Liu L, Tse F, Armstrong D. Cmos image sensors for high speed applications. Sensors 2009;9(1):430–44.
[6] Beurg M, Fettiplace R, Nam J-H, Ricci AJ. Localization of inner hair cell mechanotransducer channels using high-speed calcium imaging. Nat Neurosci 2009;12(5):553–8.
[7] Hockett P, Bisgaard CZ, Clarkin OJ, Stolow A. Time-resolved imaging of purely valence-electron dynamics during a chemical reaction. Nat Phys 2011;7(8):612–15.
[8] Goda K, Tsia K, Jalali B. Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena. Nature 2009;458(7242):1145–9.
[9] Nakagawa K, Iwasaki A, Oishi Y, Horisaki R, Tsukamoto A, Nakamura A, et al. Sequentially timed all-optical mapping photography (stamp). Nat Photon 2014;8(9):695–700.
[10] Velten A, Lawson E, Bardagjy A, Bawendi M, Raskar R. Slow art with a trillion frames per second camera. In: ACM SIGGRAPH 2011 Talks. New York, NY, USA: ACM; 2011. ISBN 978-1-4503-0974-5. SIGGRAPH '11; 44:1–44:1.
[11] Gao L, Liang J, Li C, Wang LV. Single-shot compressed ultrafast photography at one hundred billion frames per second. Nature 2014;516(7529):74–7.
[12] Ford KR, Myer GD, Hewett TE. Reliability of landing 3d motion analysis: implications for longitudinal analyses. Med Sci Sports Exercise 2007;39(11):2021.
[13] Malamas EN, Petrakis EG, Zervakis M, Petit L, Legat J-D. A survey on industrial vision systems, applications and tools. Image Vis Comput 2003;21(2):171–88.
[14] Tiwari V, Sutton M, McNeill S. Assessment of high speed imaging systems for 2d and 3d deformation measurements: methodology development and validation. Exp Mech 2007;47(4):561–79.
[15] Zhang X-y, Jin X-l, Qi W-g, Guo Y-z. Vehicle crash accident reconstruction based on the analysis 3d deformation of the auto-body. Adv Eng Softw 2008;39(6):459–65.
[16] Choi W, Fang-Yen C, Badizadegan K, Oh S, Lue N, Dasari RR, et al. Tomographic phase microscopy. Nat Method 2007;4(9):717.
[17] Webb RH. Confocal optical microscopy. Rep Progress in Phys 1996;59(3):427.
[18] Shao L, Kner P, Rego EH, Gustafsson MG. Super-resolution 3d microscopy of live whole cells using structured illumination. Nat Method 2011;8(12):1044–6.
[19] Prevedel R, Yoon Y-G, Hoffmann M, Pak N, Wetzstein G, Kato S, et al. Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. Nat Method 2014;11(7):727–30.
[20] Schwarz B. Lidar: mapping the world in 3d. Nat Photon 2010;4(7):429.
[21] Foix S, Alenya G, Torras C. Lock-in time-of-flight (tof) cameras: a survey. IEEE Sens J 2011;11(9):1917–26.
[22] Gorthi SS, Rastogi P. Fringe projection techniques: whither we are? Opt Lasers Eng 2010;48:133–40.
[23] Salvi J, Fernandez S, Pribanic T, Llado X. A state of the art in structured light patterns for surface profilometry. Pattern Recognit 2010;43(8):2666–80.
[24] Geng J. Structured-light 3d surface imaging: a tutorial. Adv Opt Photon 2011;3(2):128–60.
[25] Su X, Zhang Q. Dynamic 3-d shape measurement method: a review. Opt Lasers Eng 2010;48(2):191–204.
[26] Zhang S. Recent progresses on real-time 3d shape measurement using digital fringe projection techniques. Opt Lasers Eng 2010;48(2):149–58.
[27] Van der Jeught S, Dirckx JJ. Real-time structured light profilometry: a review. Opt Lasers Eng 2016;87:18–31.
[28] Lei S, Zhang S. Flexible 3-d shape measurement using projector defocusing. Opt Lett 2009;34(20):3080–2.
[29] Zuo C, Chen Q, Feng S, Feng F, Gu G, Sui X. Optimized pulse width modulation pattern strategy for three-dimensional profilometry with projector defocusing. Appl Opt 2012;51(19):4477–90.
[30] Wang Y, Zhang S. Three-dimensional shape measurement with binary dithered patterns. Appl Opt 2012;51(27):6631–6.
[31] Salvi J, Pages J, Batlle J. Pattern codification strategies in structured light systems. Pattern Recognit 2004;37(4):827–49.
[32] Zhang Z. Review of single-shot 3d shape measurement by phase calculation-based fringe projection techniques. Opt Lasers Eng 2012;50(8):1097–106.
[33] Zhang L, Curless B, Seitz SM. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In: 3D data processing visualization and transmission, 2002. proceedings. first international symposium on. IEEE; 2002. p. 24–36.
[34] Morita H, Yajima K, Sakata S. Reconstruction of surfaces of 3-d objects by m-array pattern projection method. In: Computer vision., second international conference on. IEEE; 1988. p. 468–73.
[35] Albitar C, Graebling P, Doignon C. Robust structured light coding for 3d reconstruction. In: Computer vision, 2007. ICCV 2007. IEEE 11th international conference on. IEEE; 2007. p. 1–6.
[36] Takeda M, Mutoh K. Fourier transform profilometry for the automatic measurement of 3-d object shapes. Appl Opt 1983;22(24):3977–82.
[37] Su X, Chen W. Fourier transform profilometry:: a review. Opt Lasers Eng 2001;35(5):263–84.
[38] Lin J-F, Su X. Two-dimensional fourier transform profilometry for the automatic measurement of three-dimensional object shapes. Opt Eng 1995;34(11):3297–302.
[39] Guo L, Su X, Li J. Improved fourier transform profilometry for the automatic measurement of 3d object shapes. Opt Eng 1990;29(12):1439–44.
[40] Berryman F, Pynsent P, Cubillo J. The effect of windowing in fourier transform profilometry applied to noisy images. Opt Lasers Eng 2004;41(6):815–25.

[41] Guo H, Huang PS. 3-d shape measurement by use of a modified fourier transform method. In: Proc. SPIE, 7066; 2008.
[42] Cong P, Xiong Z, Zhang Y, Zhao S, Wu F. Accurate dynamic 3d sensing with fourier-assisted phase shifting. IEEE J Sel Top Signal Process 2015;9(3):396–408.
[43] Chen L-C, Ho H-W, Nguyen X-L. Fourier transform profilometry (ftp) using an innovative band-pass filter for accurate 3-d surface reconstruction. Opt Lasers Eng 2010;48(2):182–90.
[44] Huang PS, Zhang C, Chiang F-P. High-speed 3-d shape measurement based on digital fringe projection. Opt Eng 2003;42(1):163–8.
[45] Zhang S, Huang PS. High-resolution, real-time three-dimensional shape measurement. Opt Eng 2006;45(12):123601.
[46] Weise T, Leibe B, Van Gool L. Fast 3d scanning with automatic motion compensation. In: Computer vision and pattern recognition, 2007. CVPR'07. IEEE conference on. IEEE; 2007. p. 1–8.
[47] Srinivasan V, Liu H-C, Halioua M. Automated phase-measuring profilometry of 3-d diffuse objects. Appl Opt 1984;23(18):3105–8.
[48] Lau DL, Liu K, Hassebrook LG. Real-time three-dimensional shape measurement of moving objects without edge errors by time-synchronized structured illumination. Opt Lett 2010;35(14):2487–9.
[49] Feng S, Chen Q, Zuo C, Li R, Shen G, Feng F. Automatic identification and removal of outliers for high-speed fringe projection profilometry. Opt Eng 2013;52(1):013605.
[50] Li B, Liu Z, Zhang S. Motion-induced error reduction by combining fourier transform profilometry with phase-shifting profilometry. Opt Express 2016;24(20):23289–303.
[51] Sansoni G, Carocci M, Rodella R. Three-dimensional vision based on a combination of gray-code and phase-shift light projection: analysis and compensation of the systematic errors. Appl Opt 1999;38(31):6565–73.
[52] Zuo C, Huang L, Zhang M, Chen Q, Asundi A. Temporal phase unwrapping algorithms for fringe projection profilometry: a comparative review. Opt Lasers Eng 2016;85:84–103.
[53] Wang Y, Zhang S. Superfast multifrequency phase-shifting technique with optimal pulse width modulation. Opt Express 2011;19(6):5149–55.
[54] Liu K, Wang Y, Lau DL, Hao Q, Hassebrook LG. Dual-frequency pattern scheme for high-speed 3-d shape measurement. Opt Express 2010;18(5):5229–44.
[55] Zuo C, Chen Q, Gu G, Feng S, Feng F. High-speed three-dimensional profilometry for multiple objects with complex shapes. Opt Express 2012;20(17):19493–510.
[56] Wang Y, Liu K, Hao Q, Lau DL, Hassebrook LG. Period coded phase shifting strategy for real–time 3-d structured light illumination. IEEE Trans Image Process 2011;20(11):3001–13.
[57] Zhang Y, Xiong Z, Wu F. Unambiguous 3d measurement from speckle-embedded fringe. Appl Opt 2013;52(32):7797–805.
[58] Tao T, Chen Q, Da J, Feng S, Hu Y, Zuo C. Real-time 3-d shape measurement with composite phase-shifting fringes and multi-view system. Opt Express 2016;24(18):20253–69.
[59] Zuo C, Chen Q, Gu G, Feng S, Feng F, Li R, et al. High-speed three-dimensional shape measurement for dynamic scenes using bi-frequency tripolar pulse-width-modulation fringe projection. Opt Lasers Eng 2013;51(8):953–60.
[60] Tao T, Chen Q, Feng S, Hu Y, Da J, Zuo C. High-precision real-time 3d shape measurement using a bi-frequency scheme and multi-view system. Appl Opt 2017;56(13):3646–53.
[61] Li Z, Zhong K, Li YF, Zhou X, Shi Y. Multiview phase shifting: a full-resolution and high-speed 3d measurement framework for arbitrary shape dynamic objects. Opt Lett 2013;38(9):1389–91.
[62] Bräuer-Burchardt C, Breitbarth A, Kühmstedt P, Notni G. High-speed three-dimensional measurements with a fringe projection-based optical sensor. Opt Eng 2014;53(11):112213.
[63] An Y, Hyun J-S, Zhang S. Pixel-wise absolute phase unwrapping using geometric constraints of structured light system. Opt Express 2016;24(16):18445–59.
[64] Schaffer M, Große M, Harendt B, Kowarschik R. Statistical patterns: an approach for high-speed and high-accuracy shape measurements. Opt Eng 2014;53(11):112205.
[65] Schaffer M, Grosse M, Harendt B, Kowarschik R. High-speed three-dimensional shape measurements of objects with laser speckles and acousto-optical deflection. Opt Lett 2011;36(16):3097–9.
[66] Heist S, Mann A, Kühmstedt P, Schreiber P, Notni G. Array projection of aperiodic sinusoidal fringes for high-speed three-dimensional shape measurement. Opt Eng 2014;53(11):112208.
[67] Heist S, Lutzke P, Schmidt I, Dietrich P, Kühmstedt P, Tünnermann A, et al. High-speed three-dimensional shape measurement using gobo projection. Opt Lasers Eng 2016;87:90–6.
[68] Feng S, Chen Q, Zuo C, Tao T, Hu Y, Asundi A. Motion-oriented high speed 3-d measurements by binocular fringe projection using binary aperiodic patterns. Opt Express 2017;25(2):540–59.
[69] Heist S, Kühmstedt P, Tünnermann A, Notni G. Theoretical considerations on aperiodic sinusoidal fringes in comparison to phase-shifted sinusoidal fringes for high-speed three-dimensional shape measurement. Appl Opt 2015;54(35):10541–51.
[70] Nguyen H, Nguyen D, Wang Z, Kieu H, Le M. Real-time, high-accuracy 3d imaging and shape measurement. Appl Opt 2015;54(1):A9–A17.
[71] Grosse M, Schaffer M, Harendt B, Kowarschik R. Fast data acquisition for three-dimensional shape measurement using fixed-pattern projection and temporal coding. Opt Eng 2011;50(10):100503.
[72] Gupta M, Nayar SK. Micro phase shifting. In: Computer vision and pattern recognition (CVPR), 2012 IEEE conference on. IEEE; 2012. p. 813–20.
[73] Dudley D., Duncan W.M., Slaughter J.. Emerging digital micromirror device (dmd) applications. vol. 4985. 2003, p. 14–25.. 10.1117/12.480761
[74] Floyd RW. An adaptive algorithm for spatial gray-scale. In: Proc. Soc. Inf. Disp., 17; 1976. p. 75–7.
[75] Herráez MA, Burton DR, Lalor MJ, Gdeisat MA. Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path. Appl Opt 2002;41(35):7437–44.
[76] Su X, Chen W. Reliability-guided phase unwrapping algorithm: a review. Opt Lasers Eng 2004;42(3):245–61.
[77] Zhao M, Huang L, Zhang Q, Su X, Asundi A, Kemao Q. Quality-guided phase unwrapping technique: comparison of quality maps and guiding strategies. Appl Opt 2011;50(33):6214–24.
[78] Smart computational imaging laboratory. http://www.scilaboratory.com/h-col-123.html, (14/02/2017).
[79] Zhang Q, Su X, Cao Y, Li Y, Xiang L, Chen W. Optical 3-d shape and deformation measurement of rotating blades using stroboscopic structured illumination. Opt Eng 2005;44(11):113601.
[80] Gustafsson MG. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. J Microsc 2000;198(2):82–7.
[81] Dan D, Lei M, Yao B, Wang W, Winterhalder M, Zumbusch A, et al. Dmd-based led-illumination super-resolution and optical sectioning microscopy. Sci Rep 2013;3:1116.
[82] Sun B, Edgar MP, Bowman R, Vittert LE, Welsh S, Bowman A, et al. 3D computational imaging with single-pixel detectors. Science 2013;340(6134):844–7.
[83] Zhang Z, Ma X, Zhong J. Single-pixel imaging by means of fourier spectrum acquisition. Nat Commun 2015;6.
[84] Gushov V, Solodkin YN. Automatic processing of fringe patterns in integer interferometers. Opt Lasers Eng 1991;14(4):311–24.
[85] Lilienblum E, Michaelis B. Optical 3d surface reconstruction by a multi-period phase shift method.. JCP 2007;2(2):73–83.
[86] Pribanic T, Mrvos S, Salvi J. Efficient multiple phase shift patterns for dense 3d acquisition in structured light scanning. Image Vis Comput 2010;28(8):1255–66.
[87] Petkovic T, Pribanic T, Donlic M. Temporal phase unwrapping using orthographic projection. Opt Lasers Eng 2017;90:34–47.
[88] Taylor J.. Introduction to error analysis, the study of uncertainties in physical measurements. 1997.
[89] Cheng Y-Y, Wyant JC. Two-wavelength phase shifting interferometry. Appl Opt 1984;23(24):4539–43.
[90] Creath K. Step height measurement using two-wavelength phase-shifting interferometry. Appl Opt 1987;26(14):2810–16.
[91] Zhang Z, Towers CE, Towers DP. Time efficient color fringe projection system for 3d shape and color using optimum 3-frequency selection. Opt Express 2006;14(14):6444–55.
[92] Heikkila J. Geometric camera calibration using circular control points. IEEE Trans Pattern Anal Mach Intell 2000;22(10):1066–77.
[93] Bouguet J.. Camera calibration toolbox for matlab. https://www.vision.caltech.edu/bouguetj/calib_doc/, (2017).
[94] Triggs B, Zisserman A, Szeliski R. Vision algorithms: theory and practice: international workshop on vision algorithms Corfu, Greece, September 21–22, 1999 Proceedings. Springer; 2003.
[95] Huang L, Zhang Q, Asundi A. Camera calibration with active phase target: improvement on feature detection and optimization. Opt Lett 2013;38(9):1446–8.