Real-time high dynamic range 3D measurement using fringe projection

Liang Zhang,1,2,3 Qian Chen,1,2,4 Ⓞ Chao Zuo,1,2,3 Ⓞ and Shijie Feng1,2,3,5

1School of Electronic and Optical Engineering, Nanjing University of Science and Technology, No. 200 Xiaolingwei Street, Nanjing, Jiangsu Province 210094, China
2Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing, Jiangsu Province 210094, China
3Smart Computational Imaging (SCI) Laboratory, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China
4chenqian@njust.edu.cn
5shijiefeng@njust.edu.cn

Abstract: Fringe projection profilometry (FPP) is a widely used technique for real-time three-dimensional (3D) shape measurement. However, it tends to compromise when measuring objects that have a large variation range of surface reflectivity. In this paper, we present a FPP method that can increase the dynamic range for real-time 3D measurements. First, binary fringe patterns are projected to generate grayscale sinusoidal patterns with the defocusing technique. Each pattern is then captured twice with different exposure values in one projection period. With image fusion, surfaces under appropriate exposure are retained. To improve the real-time performance of high dynamic range (HDR) 3D shape measurements, we build a binocular fringe projection profilometry system that saves the number of patterns by geometry constraint. Further, to ensure the accuracy and robustness of HDR 3D measurements, we propose a mixed phase unwrapping method that can reduce phase unwrapping errors for dense fringe patterns. Experiment results show that the proposed method can realize accurate and real-time 3D measurement for HDR scenes at 28 frames per second.

© 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

With the development of technology, 3D measurement technology has found wide applications in industry, medicine, and the protection of historical relics. Fringe projection profilometry (FPP), as a structured light based 3D measurement technique, has obtained attention widely due to the advantages of high speed, high accuracy, and non-contact [1–3]. A typical FPP system usually consists of one projector and one/several camera(s). The projector is used to project the coded fringe patterns, and the cameras are used to capture the images reflected by the measured objects. The light intensity values of captured images depend on the reflectivity of the measured objects. When the measured objects have large variation ranges of surface reflectivity, it is hard for ordinary cameras to provide the required dynamic range within a single exposure. Normally, high reflection areas are prone to be saturated, while low reflection areas lead to low signal-to-noise ratio (SNR). The resultant measurement results would contain significant errors.

In recent years, many HDR 3D measurement methods (hereafter called HDR methods) have been proposed to solve the HDR problem in FPP. These HDR methods can generally be divided into two broad categories [4]: algorithm-based techniques and equipment-based techniques. The algorithm-based techniques can be considered as one kind of post-processing compensation algorithms. Take the case of the phase shifting fringe images, each image pixel has three unknowns (the average intensity, the intensity modulation, and the phase) to be calculated. Therefore, each pixel needs at least three unsaturated images to solve its corresponding unknowns. As is well
known, when the saturation phenomenon occurs, the exact grayscale values of the saturated pixels will be overridden by the saturation value (255 for an 8-bit camera). To obtain the required number of unsaturated images, one has to project more fringe patterns [5,6]. In other words, the algorithm-based techniques improve the dynamic range at the cost of the measurement efficiency. On the other hand, for algorithm-based techniques, the patterns with different phase shifting steps usually require different phase calculation algorithms, which increases the computational complexity.

The equipment-based techniques eliminate the saturation error by adjusting the hardware parameters [7,8] or using some special devices [9,10]. The multi-exposure method [7] is the most representative equipment-based technique. Its principle is similar with that of HDR photography. When measuring HDR objects, the multi-exposure method first takes a set of images with different exposures, and then combines them to a new HDR image. It fundamentally solves the HDR problem and is easy to implement with no additional hardware. But its shortcoming is also very obvious: it can only measure static objects. As mentioned before, one needs to take several sets of images with different exposures to obtain the desired dynamic range. During the above image acquisition process, any movement of the measured objects will bring the motion error [11]. In recent years, many HDR methods [12–15] have been proposed to improve the multi-exposure method. Although these methods effectively improve the measurement efficiency, their measurement speeds are still insufficient for real-time measurement. To realize the HDR measurement in real-time, Suresh et al. proposes an improved multi-exposure method based on the digital-light-processing (DLP) technology [16]. Compared with the multi-exposure method, Suresh’s method can double the dynamic range without manual intervention by using the characteristic of binary patterns. However, limited by the monocular system, Suresh’s method still needs additional fringe patterns to calculate the unwrapped phase (absolute phase). To further improve the real-time performance, we build a binocular system to replace the conventional monocular system. Compared with the monocular system, the binocular system can use fewer fringe patterns to calculate the unwrapped phase by the geometric constraint between two cameras. Moreover, a novel mixed phase unwrapping method is designed to work with the binocular system, which can effectively improve the reliability of the phase unwrapping. Experiment results show that the proposed method can realize accurate and real-time 3D measurement for HDR scenes at 28 frames per second.

2. Principle

2.1. High dynamic image acquisition

In FPP system, the relationship between the captured light intensity \( I_c \) and the projected light intensity \( I_p \) can be simplified to

\[
I_c = \alpha rt I_p.
\]

(1)

where \( \alpha \) is the camera sensitivity, \( r \) is the reflectivity of the measured objects, and \( t \) is the camera exposure. Normally one camera exposure only corresponds to one projection period. Suresh et al. [16] first attempted to get images with two exposures in one projection period by the DLP technology. As is known to all, the projection period is essentially an integration process of light intensity. For a DLP projector, the above integration process is implemented by its core component, the digital micromirror device (DMD). In one projection period, each micromirror toggles ‘on’ and ‘off’ very quickly, and the ratio of ‘on’ time to ‘off’ time determines grayscale values. Different with grayscale patterns, binary patterns are images with only two colors: black and white (0 and 255 for 8-bit grayscale). For binary patterns, one only needs to get the contrast of bright versus dark instead of the accurate grayscale values. This is the basic principle for the DLP technology.
Figure 1 illustrates the timing diagram of the DLP technology in one projection period [16]. T1 is the external signal to trigger the projector, T2 is the pattern exposure time, and T3 is the blank period between two pattern exposures. The cameras are triggered by the falling edge of external signal T4. The first capture has its exposure T5 within T2, while the second capture has part of its exposure T6 within T2 and the other part of its exposure T7 within T3. Since the blank period T3 is completely black, the image brightness of the second capture is smaller than that of the first capture. The brightness difference depends on the ratio of T5 to T6. Assuming that I1 and I2 represent the images in the first capture (with the exposure of T5) and the second capture (with the exposures of T6 and T7) respectively. When measuring HDR objects, we can use I1 to improve the SNR of low-reflection regions, and use I2 to avoid the saturation of high-reflection regions. The fusion image I' can be written as follows:

\[
I'(x',y') = \begin{cases} 
I_1(x',y'), & I_1(x',y') \leq L_{th} \\
I_2(x',y'), & I_1(x',y') > L_{th}
\end{cases},
\]

where \((x',y')\) is the pixel coordinate in the camera plane, \(L_{th}\) is the light intensity threshold. For an 8-bit camera, \(L_{th}\) should be less than 255 (grayscale). Here we use a HDR scene as example to validate the feasibility and effectiveness of the DLP technology. As shown in Fig. 2(a), the HDR scene consists of a green plastic block, a black plastic block, a white plastic block, and a brown carton. A pattern with uniform intensity of 255 (grayscale) is projected onto the HDR scene. The camera exposures are set as: \(T_6 = T_7 = \frac{1}{2}T_5\). Figure 2(b) is the captured image with T5, and Fig. 2(c) is the captured image with T6 and T7. These two images can be combined to obtain a new HDR image shown in Fig. 2(d).

**Fig. 1.** Timing Diagram of the DLP technology.

It is important to note that we cannot directly project the binary patterns onto the measured objects. They should first be converted into 8-bit grayscale sinusoidal fringe patterns by the defocusing technique [17]. The sinusoidal property of the defocusing patterns depends on the fringe frequency. The defocusing effect can be approximated as a Gaussian filtering [18]. Here we use MATLAB to simulate the effect of the projector defocusing by a Gaussian filter with a size of 39 x 39 pixels and a standard deviation of 3 pixels. The simulation results are shown in Fig. 3. It can be seen that the low-frequency binary pattern is hard to generate the sinusoidal pattern shown in Figs. 3(a)-(b). By contrast, the high-frequency binary pattern is more easily changed into the sinusoidal pattern shown in Figs. 3(c)-(d). This implied condition of the defocusing technique has a detrimental effect on phase unwrapping for the monocular system. The monocular system only has one camera, so it usually needs low-frequency fringe patterns to assist in phase unwrapping when using temporal phase unwrapping methods [19]. While the poor sinusoidal property of low-frequency binary patterns may cause great measurement error. On the other hand, projecting additional low-frequency binary patterns affects the real-time performance. To
solve this problem, we build a binocular system and propose a mixed phase unwrapping method which are detailed in the next section.

2.2. Mixed phase unwrapping method

The diagram of our binocular system is shown in Fig. 4. In our system, Camera1 is the chief camera, and Camera2 is the auxiliary camera which is used to assist Camera1 in calculating the unwrapped phase. Assuming the defocusing binary patterns captured by Camera1 can be expressed as follows:

$$I^{c1}(x, y) = \alpha rt \{ A + B \cdot \cos[\phi(x^{p}, y^{p})] \}. \quad (3)$$

where $A$ is the average intensity, $B$ is the intensity modulation, $\phi$ is the phase to be measured, and $(x^{p}, y^{p})$ is the pixel coordinate in the projector plane. For Camera1, its corresponding world coordinate can be calculated as

$$z^{w}(x^{c1}, y^{c1}) = \frac{M_{1}(x^{c1}, y^{c1})}{M_{2}(x^{c1}, y^{c1}) \cdot x^{2}(x^{c1}, y^{c1}) + 1} + M_{3}(x^{c1}, y^{c1}), \quad (4)$$

$$x^{w}(x^{c1}, y^{c1}) = M_{4}(x^{c1}, y^{c1}) \cdot z^{w}(x^{c1}, y^{c1}) + M_{5}(x^{c1}, y^{c1}), \quad (5)$$
\[ y^w(x^c1, y^c1) = M_6(x^c1, y^c1) \cdot z^u(x^c1, y^c1) + M_7(x^c1, y^c1), \]  

(6)

where \( M_1(x^c1, y^c1) \sim M_7(x^c1, y^c1) \) are the parameter matrices calculated from calibration parameters [20], and \( x^c2 \) is Camera2’s horizontal coordinate corresponding to \((x^c1, y^c1)\). If we want to get \( x^c2 \), we first need to know the unwrapped phase of Camera1 (labeled as \( \Phi^c1 \)), which can be written as:

\[ \Phi^c1(x^c1, y^c1) = \phi^c1(x^c1, y^c1) + k(x^c1, y^c1) \cdot 2\pi, \quad k(x^c1, y^c1) \in [0, N - 1], \]  

(7)

where \( \phi^c1(x^c1, y^c1) \) is the wrapped phase of Camera1, \( k(x^c1, y^c1) \) is the fringe order, and \( N \) is the period number of fringes. How to determine \( k(x^c1, y^c1) \) quickly and accurately is the essential problem of phase unwrapping.

Because of the additional camera view, the binocular system can retrieve the absolute phase on the basis of the geometric relationship between different views [21]. Such kind of method is also known as the stereo phase unwrapping (SPU) method. The principle of SPU is shown in Fig. 5(a). Assuming \( P^c1 \) is a pixel in Camera1, its wrapped phase is \( \phi^c1 \) and its unwrapped phase is \( \Phi^c1 \). \( \phi^c1 \) can be easily calculated by some phase calculation algorithms. Only knowing \( \phi^c1 \) cannot help us get the exact corresponding point in the world coordinate directly, but we can calculate all the possible points (labeled as \( P^c2_k \)) by using different fringe orders. \( P^c2_k \) are the 3D candidates with same wrapped phase but different fringe orders. They can be mapped to the camera plane of Camera2, then we can get a set of 2D candidates (labeled as \( P^c2_m \)) in Camera2. If the match pixel of \( P^c1 \) in Camera2 is \( P^c2_m \), they should have the similar properties. In this paper, \( P^c2_k \) with the closest wrapped phase to \( P^c1 \) is considered to be \( P^c2_m \). Then we can get the fringe order and unwrapped phase of \( P^c1 \). It can be seen that SPU has no restrictions on the shape of the measured objects and no requirements of extra projection. Compared with spatial/temporal phase unwrapping, SPU has wider application scope and higher efficiency.

In actual measurement, the reliability of SPU depends on the used fringe frequency. As shown in Fig. 5(a), when using low-frequency fringe patterns, there are only a few candidates in the measurement range. In this case, it is easy to find out the exact fringe order of \( P^c1 \). In other words, low-frequency fringe patterns have less phase ambiguity in SPU, tending to make the phase unwrapping more reliable [22]. However, as mentioned before, low-frequency binary patterns are difficult to be converted into sinusoidal patterns by the defocusing technique. By contrast, high-frequency patterns have better sinusoidal property and higher accuracy. However, as shown in Fig. 5(b), when projecting high-frequency patterns, there exists too many candidates...
in the measurement range which may cause phase ambiguity. Of course, we can further narrow the measurement range by using $Z_{\text{min}}$ and $Z_{\text{max}}$. But in this way, we can hardly guarantee the measured objects are located in such narrow range, especially when the objects are moving.

To overcome the dilemma of the frequency selection, we propose a novel mixed phase unwrapping method. The framework of the mixed phase unwrapping method is illustrated in Fig. 6. It can be divided into three steps as follows:

Step 1: Wrapped phase calculation. The mixed phase unwrapping method needs fringe patterns with two different frequencies. If we adopt the N-step phase shifting algorithm [23] to calculate the wrapped phase, at least three fringe patterns are required for each frequency. These six patterns are too many for real-time measurement. Therefore, we choose the Fourier transform algorithm [24] to calculate the wrapped phase. Taking the Fourier transform, Eq. (3) gives

$$
\hat{F}(f_x, f_y) = \hat{A}(f_x, f_y) + \hat{C}(f_x - f_0, f_y) + \hat{C}^*(f_x + f_0, f_y),
$$

where $\hat{A}(f_x, f_y)$ is the Fourier transforms of $A(x^c, y^c)$, $f_0$ is the frequency of the fringe pattern, $\hat{C}(f_x - f_0, f_y)$ is the Fourier transforms of $\frac{1}{2}B(x', y') \cdot e^{-i\Phi(x', y')}$, and $*$ is the complex conjugate. A properly designed filter can isolate $\hat{C}(f_x - f_0, f_y)$, and then $\Phi(x', y')$ is retrieved by calculating a complex logarithm of the result of the inverse Fourier transform. This is the principle of the
conventional Fourier transform algorithm. In actual measurement, it does not work well when measured objects have steep slopes [25]. Therefore, in this paper, we use a modified Fourier transform algorithm called background normalized Fourier transform profilometry (BNFTP) [26] to calculate the wrapped phase. Different from the conventional Fourier transform, we insert a pattern with uniform intensity of 255 (grayscale) between two fringe patterns shown in Fig. 6(b). According to Eq. (1), the captured fringe images can be expressed as

\[
I_1(x', y') = arT_5\left\{ A + B \cdot \cos[\phi_1(x', y')] \right\},
\]

\[
I_2(x', y') = arT_5(A + B),
\]

\[
I_3(x', y') = arT_5\left\{ A + B \cdot \cos[\phi_2(x', y')] \right\},
\]

\[
I_4(x', y') = arT_5\left\{ A + B \cdot \cos[\phi_1(x', y')] \right\},
\]

\[
I_5(x', y') = arT_5(A + B),
\]

\[
I_6(x', y') = arT_5\left\{ A + B \cdot \cos[\phi_2(x', y')] \right\},
\]

where \( \phi_1 \) is the wrapped phase corresponding to the first frequency, \( \phi_2 \) is the wrapped phase corresponding to the second frequency, \( I_1 \sim I_6 \) are the images in the first capture (when the camera exposure is T5), and \( I_4 \sim I_6 \) are the images in the second capture (when the camera exposures are T6 and T7). The fusion images \( I_1' \sim I_6' \) can be calculated as

\[
I_1'(x', y') = \begin{cases} 
  I_1(x', y'), function \leq L_{th} \\
  I_4(x', y'), function > L_{th}
\end{cases},
\]

\[
I_2'(x', y') = \begin{cases} 
  I_2(x', y'), function \leq L_{th} \\
  I_5(x', y'), function > L_{th}
\end{cases},
\]

\[
I_3'(x', y') = \begin{cases} 
  I_3(x', y'), function \leq L_{th} \\
  I_6(x', y'), function > L_{th}
\end{cases},
\]

\[
function = \max[I_1'(x', y'), I_2'(x', y'), I_3'(x', y')]\]

where \( \max() \) is the function which can find the maximum value. With the subtraction and normalization of the uniform intensity pattern, the effect of zero-order can be removed before the Fourier transform, and the spectrum overlap in the frequency domain can be significantly alleviated:

\[
I_1'(x', y') = \frac{2I_1(x', y') - I_2(x', y')}{{I_2(x', y')} + \gamma},
\]

\[
I_5'(x', y') = \frac{2I_5(x', y') - I_6(x', y')}{{I_6(x', y')} + \gamma},
\]

where \( \gamma \) is a small constant to prevent divide-by-zero error. Then we apply Fourier transform on \( I_1' \) and \( I_5' \) to obtain the wrapped phase \( \phi_1 \) and \( \phi_2 \).

Step 2: Phase synthesisization. In order to ensure the measuring accuracy and sinusoidal property, the frequencies of the fringe patterns in Step 1 are relatively high. To avoid phase ambiguity, their frequencies should be reduced for the subsequent phase unwrapping. Here we
use the multi-wavelength temporal phase unwrapping (MWTPU) method [27] to synthesize the high-frequency phase $\phi_1$ and $\phi_2$ to a low-frequency phase:

$$\phi_{\text{syn}} = (\phi_1 - \phi_2) \text{mod}(2\pi),$$

(16)

where $\phi_{\text{syn}}$ is the synthetic phase, and $\text{mod}()$ is the modulus operation. Assuming the frequencies of $\phi_1$ and $\phi_2$ are $f_1$ and $f_2$ respectively, the frequency of $\phi_{\text{syn}}$ can be calculated as

$$f_{\text{syn}} = |f_1 - f_2|.$$  

(17)

Equation (17) makes out that properly selecting $f_1$ and $f_2$ can effectively reduce the phase frequency. But it should be noted that MWTPU reduces the phase accuracy at the same time [19]. Therefore, $\phi_{\text{syn}}$ is only used as a reference phase to calculate the refined wrapped phase $\phi_{\text{ref}}$:

$$\lambda_{\text{syn}} = \frac{\lambda_1 \cdot \lambda_2}{\lambda_1 - \lambda_2},$$

(18)

$$k_{\text{syn}}(x^c, y^c) = \text{Round}\left[\frac{\lambda_{\text{syn}}}{\lambda_1} \cdot \phi_{\text{syn}}(x^c, y^c) - \phi_1(x^c, y^c)\right],$$

(19)

$$\phi_{\text{ref}}(x^c, y^c) = \phi_1(x^c, y^c) + k_{\text{syn}}(x^c, y^c) \cdot 2\pi,$$

(20)

where $\lambda_1$ is the wavelength of the fringe patterns with $f_1$, $\lambda_2$ is the wavelength of the fringe patterns with $f_2$, $\lambda_{\text{syn}}$ is the synthetic wavelength, $k_{\text{syn}}(x^c, y^c)$ is the fringe order corresponding to $f_{\text{syn}}$, and Round$(X)$ rounds each element of $X$ to the nearest integer. According to the above equations, we can calculate the refined wrapped phase for Camera1 and Camera2 (labeled as $\phi_{\text{ref}}^1$ and $\phi_{\text{ref}}^2$).

For our projector, its resolution is $912 \times 1140$ (pixel). The wavelengths of $\lambda_1$ and $\lambda_2$ are set to 16 pixels and 12 pixels respectively. Then $f_1$, $f_2$ and $f_{\text{syn}}$ can be calculated as

$$f_1 = \frac{912}{16} = 57,\quad f_2 = \frac{912}{12} = 76,$$

$$f_{\text{syn}} = 57 - 76 = 19.$$  

(21)

For $\phi_{\text{ref}}$, its frequency ($f_{\text{ref}}$) is low enough to avoid phase ambiguity for large measurement range, while maintaining high phase accuracy shown in Fig. 7(e).

Step 3: Stereo phase unwrapping. For a given pixel $P^1(x^c, y^c)$ in Camera1, we do not know its exact fringe order $k(x^c, y^c)$. We first plug all the possible values of $k(x^c, y^c)$ into Eq. (7), then we can get all the possible unwrapped phase of $P^1(x^c, y^c)$:

$$\Phi^1_n(x^c, y^c) = \phi^1_{\text{ref}}(x^c, y^c) + k'_n \cdot 2\pi, \quad k'_n \in [0, N - 1].$$

(22)

Then we can get a set of points in the world coordinate (labeled as $P^w_n(x^w_n, y^w_n, z^w_n)$), and the correct $P^w$ is included in these candidates. In actual measurement, there is no need to calculate all these candidates. Once the binocular system is built, its measurement range can be estimated according to the system parameters. We can rule out some candidates by the depth constraint as follows:

$$Z_{\text{min}} < z^w_n(x^c, y^c) < Z_{\text{max}}.$$

(23)

where $Z_{\text{min}}$ and $Z_{\text{max}}$ represent the minimum and maximum depth boundaries in the world coordinate respectively. As shown in Fig. 5(a), we can rule out $P^w_1$, $P^w_4$, and $P^w_5$, only preserve $P^w_2$ and $P^w_3$. For the remaining candidates, we need the assistance of Camera2. Now that we
have gotten $P^m_n$, we can map them to the camera coordinate of Camera2 by using the calibration parameters:

$$P^m_n(x^m_n, y^m_n, z^m_n) \rightarrow P^{c2}_n(x^{c2}, y^{c2}).$$

The phase difference between $P^{c1}(x^{c1}, y^{c1})$ and $P^{c2}(x^{c2}, y^{c2})$ can be calculated as

$$\Delta \phi = \left| \phi^{c1}_{c1}(x^{c1}, y^{c1}) - \phi^{c2}_{c2}(x^{c2}, y^{c2}) \right|,$$

$P^{c2}$ with the minimum $\Delta \phi$ is considered to be the matching point (labeled as $P^{c2}_m$) of $P^{c1}$, and the corresponding $k'_n$ is the fringe order of $P^{c1}$. In general, $P^{c2}_m$ has an offset to the ideal one, we can further calculate the sub-pixel match point by the linear interpolation to improve the matching accuracy [28]. Then we can calculate the world coordinate by Eqs. (4)–(6). After calculating all the $P^n(x^n, y^n, z^n)$ of the pixels in Camera1, we can obtain the 3D reconstruction results.

Finally, it should be noted that in actual measurement, the imperfect system calibration and the system noise inevitably bring the measurement error. Besides geometry constraint, we also need a simplified left-right consistency check method [22] to further rule out some wrong candidates. Moreover, an edge detection algorithm [29] is used to avoid false calculation of the points at the edge of fringes.

3. Experimental results

3.1. Experimental setup

A binocular FPP system is built to evaluate the proposed method shown in Fig. 8. The system consists of two high-speed CMOS cameras (model: Basler acA640-750um) with a resolution of $640 \times 480$ and a DLP projector (model: TI LightCrafter 4500) with a resolution of $912 \times 1140$. To precisely synchronize the cameras with the DLP projector, a field programmable gate array (FPGA) development board (model: Zircon-A4) is used to provide the external triggering signals. The Zircon-A4 board includes one oscillator that produces a 50 MHz clock signal. A clock buffer is used to distribute 50 MHz clock signal with low jitter to FPGA. According to the clock signal, we can precisely control the pulse widths of trigger signals and duration ratio. Three linear polarizers are put in front of the cameras and the DLP projector respectively to eliminate the saturation caused by specular reflection [30].
In our experiment, the wavelengths of the used fringe patterns are 16 and 12 pixels respectively. The values of $T_1 \sim T_7$ are set as follows: $T_1 = T_2 = 9000 \mu s$, $T_3 = 3000 \mu s$, $T_4 = 20.5 \mu s$, $T_5 = 4000 \mu s$, and $T_6 = T_7 = 2000 \mu s$. The projection period is

$$T_2 + T_3 = 9000 \mu s + 3000 \mu s = 12000 \mu s.$$  \hspace{1cm} (26)

Because we only need three projection periods to reconstruct the 3D structure, the 3D frame rate of our method is about

$$\frac{1 \text{s}}{3 \times 12000 \mu s} = \frac{1 \times 10^6 \mu s}{3 \times 12000 \mu s} \approx 28 \text{ Hz}.$$ \hspace{1cm} (27)

This speed is fast enough to meet the requirement of real-time measurement.

### 3.2. Static measurement

In the static measurement, we measured some stationary objects to verify the precision and measurement dynamic range of our method.

First, we measured a pair of standard ceramic spheres shown in Fig. 9(a). These spheres have the diameter of 50.8 mm within micron-sized manufacturing error. Figures 10(a) and (d) show their 3D reconstruction results calculated by our method. In order to verify the measurement precision, the 3D point cloud data were used to perform the spherical fitting, and the fitted spheres were set as the ground truth. The differences between the measurement data and the ground truth are shown in Figs. 10(b) and (e), and the error distributions of Sphere A and Sphere B are shown in Figs. 10(c) and (f). According to the spherical fitting results, we calculated their root-mean-square error (RMSE). The RMSE of Sphere A is about 0.084 mm and the RMSE of Sphere B is about 0.081 mm. This experiment demonstrates that our method can get high accurate measurement results.

Then we measured a HDR scene which consists of a black paperboard and a white plastic toy shown in Fig. 9(b). Because of the color difference between these two measured objects, it is difficult to measure them well within a single exposure. Figures 11(a) and (d) show that using the high exposure $T_5$ is favourable to the SNR improvement of the paperboard, but makes the plastic toy saturated. The saturation leads to incorrect measurement areas (depicted as large holes) in the 3D reconstruction results shown in Fig. 12(c). By contrary, using the low exposure $T_6$ can avoid the saturation shown in Figs. 11(b) and (e), but increases the noise impact on the measurement accuracy of the paperboard. According to the plane fitting, the RMSE of the paperboard shown
Fig. 9. (a) Photograph of the standard ceramic spheres; (b) Photograph of the HDR scene.

Fig. 10. (a) 3D reconstruction result of Sphere A; (b) Measurement error of Sphere A; (c) Error distribution of Sphere A; (d) 3D reconstruction result of Sphere B; (e) Measurement error of Sphere B; (f) Error distribution of Sphere B.

in Fig. 12(d) is about 0.22 mm. Figures 11(c) and (f) show the fusion images generated by our method. We can use them to avoid the saturation and improve the SNR at the same time. The corresponding measurement results are shown in Figs. 12(g)-(i). It can be seen that the plastic toy is measured well, and the RMSE of the paperboard with our method is about 0.10 mm. This experiment demonstrates that our method can increase the measurement dynamic range of FPP.

3.3. Real-time measurement

In the real-time measurement, we first measured a statue of a captain shown in Fig. 13(a). It can be seen that the details are rich in the face region of the captain statue. Its 3D result shown in Fig. 14 is calculated by our own program based on C++. The visual interface was developed with Qt and all the core calculation algorithms were written based on CUDA. Our program has different display models shown in Figs. 14(a)-(b). Moreover, with the help of the program, we can rotate the 3D reconstruction result of the statue and zoom in on the areas of interest shown in Figs. 14(c)-(d). The related operations can be referred to in Visualization 1. Figure 15(a) shows the real-time 3D measurement result of the statue of a captain. The whole real-time measurement process can be seen in Visualization 2. This experiment demonstrates that our method has the capacity of measuring complex objects in real-time.

Then we measured a HDR scene which consists of a statue of David and a piece of black paperboard shown in Fig. 13(b). The 3D measurement result of the HDR scene is shown in Fig. 15(b). It can be seen that both the plaster statue and the black paperboard are measured well. For the statue of David, our method can well recover its details, such as the curly hair and facial features. The real-time measurement process and related 3D reconstruction result can be seen in
Fig. 11. (a) $I_{c1}$ of Camera1 and the corresponding cross section; (b) $I_{c4}$ of Camera1 and the corresponding cross section; (c) $I_{f1}$ of Camera1 and the corresponding cross section; (d) $I_{c1}$ of Camera2 and the corresponding cross section; (e) $I_{c4}$ of Camera2 and the corresponding cross section; (f) $I_{f1}$ of Camera2 and the corresponding cross section.

Fig. 12. (a) Fitting result of the dotted line in (b); (b) 3D reconstruction result of the HDR scene with the exposure of $T_5$; (c) Enlarged detail of the plastic toy in (b); (d) Fitting result of the dotted line in (e); (e) 3D reconstruction result of the HDR scene with the exposures of $T_6$ and $T_7$; (f) Enlarged detail of the plastic toy in (e); (g) Fitting result of the dotted line in (h); (h) 3D reconstruction result of the HDR scene with our method; (i) Enlarged detail of the plastic toy in (h).
Fig. 13. (a) Photograph of the statue of a captain; (b) Photograph of the HDR scene; (c) Photograph of the HDR scene with specular reflection.

Fig. 14. The measurement result of the statue of a captain. (a) Normal display mode; (b) Depth display mode; (c) Zoom in on the statue; (d) Zoom out on the statue.

Fig. 15. (a) Real-time measurement process and result of the statue of a captain; (b) Real-time measurement process and result of the HDR scene.
Visualization 3. This experiment demonstrates that our method has the capacity of measuring HDR objects in real-time.

At last, we measured a HDR scene with specular reflection which consists of a white porcelain bottle and a blue porcelain bowl shown in Fig. 13(c). Different with the diffuse reflection, the specular reflection mechanism is much more complex. The most significant difference is that the specular saturation cannot be eliminated by reducing the camera exposure or the projected light intensity [31]. That is the reason why we add three linear polarizers into our binocular system. As shown in Figs. 16(a) and (c), if we do not use the polarizers, there are a lot of specular saturated pixels due to the surface property of porcelain. With the help of the linear polarizers, we can effectively eliminate the specular saturated pixels shown in Figs. 16(b) and (d).

![Camera1 and Camera2 images](image)

Fig. 16. (a) Captured image of Camera1 without polarizer; (b) Captured image of Camera1 with polarizer; (c) Captured image of Camera2 without polarizer; (d) Captured image of Camera2 with polarizer.

After eliminating the specular reflection, we can measure the porcelain bottle and bowl like an ordinary HDR scene shown in Fig. 17. In the real-time measurement, we held the bowl still and rotated the white bottle. The whole measurement process and related 3D reconstruction result can be seen in Visualization 4. This experiment demonstrates that our method has the capacity of measuring HDR objects with specular reflection in real-time.

![Camera1 and Camera2 images](image)

Fig. 17. (a)-(f) $I_c^1 \sim I_c^6$ of Camera1; (g)-(i) $I_f^1 \sim I_f^3$ of Camera1; (j)-(o) $I_c^1 \sim I_c^6$ of Camera2; (p)-(r) $I_f^1 \sim I_f^3$ of Camera2; (s) 3D measurement result.
4. Conclusion

This paper has proposed a novel real-time HDR 3D measurement method for FPP. It can increase the dynamic range of the FPP system by two times. Benefiting from the binocular system and the mixed phase unwrapping method, we only need three projection periods to reconstruct a 3D structure. Compared with conventional HDR methods, the proposed method uses fewer fringe patterns, thus improving the measurement efficiency and providing better real-time performance. For the binocular system, there are some other methods [32–34] can reduce the number of fringe patterns. However, they are usually implemented with the assist of complicated and time-consuming spatial domain processing algorithms, which increases the amount of calculation remarkably. By contrast, our method costs lower and it is easier to realize.

There are several aspects that need to be further improved in the future work. According to the real-time measurement results, there are still some ripples on the edges of 3D reconstructions when the measured objects move fast. In the next work, we will focus on the motion compensation to eliminate the motion error. Second, our method only increases the dynamic range by a factor of two. How to further increase the dynamic range of real-time measurement is another important direction for future investigation.

Funding

National Key Research and Development Program of China (2017YFF0106403); National Natural Science Foundation of China (111574152, 61705105, 61722506); Fundamental Research Funds for the Central Universities (30916011322, 30917011204); ‘333 Engineering’ research project of Jiangsu Province, China (BRA2015294, BRA2016407); ‘Six Talent Peaks’ project of Jiangsu Province, China (2015-DZXX-009); National Defense Science and Technology Foundation of China (0106173); Outstanding Youth Foundation of Jiangsu Province of China (BK20170034); The Key Research and Development Program of Jiangsu Province, China (BE2017162); Equipment Advanced Research Fund of China (61404150202); Final Assembly ‘13th Five-Year Plan’ Advanced Research Project of China (30102070102).

Disclosures

The authors declare no conflicts of interest.

References