Sign language learning based on high-speed fringe projection profilometry employing defocused binary fringe*

WANG Jian-hua (王建华)[1,2]**, **ZHOU Yu-guo** (周玉国)[1], **and YANG Yan-xi** (杨延西)[2]
1. School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China
2. School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China

(Received 13 March 2019; Revised 25 April 2019)

A high-speed fringe projection profilometry employing defocused binary fringe is presented to record the 3-D sign language, including gesture and mouth movement, to help people learn the sign language. It employs a number of advanced approaches, such as fringes binarization and defocus method, high precision phase calculation based on the phase-shifting method and the three pitches heterodyne unwrapping (TPHU) method, and how to combine multiple binary fringes into a 24-bit fringe. Experiments have shown that the proposed system can acquire and display high-quality 3D gesture and mouth movement at a speed of 500 frames per second.

Document code: A **Article ID:** 1673-1905(2020)01-0065-10
DOI https://doi.org/10.1007/s11801-020-9040-2

Sign language is the language of deaf mutes. It is an important tool for communication between deaf mutes, healthy people and deaf mutes. Unfortunately, the deaf and dumb people have some shortcomings in listening and speaking. They will have some difficulties in communicating with the voice world, but sign language can solve this difficulty effectively. In order to communicate with deaf mutes and help them integrate into society, not only deaf mute people need to learn sign language, but also healthy people need to learn. Sign language is usually composed of gesture and mouth movement. In China, the basic unit is 26 letters' sign language, as shown in Fig.1.

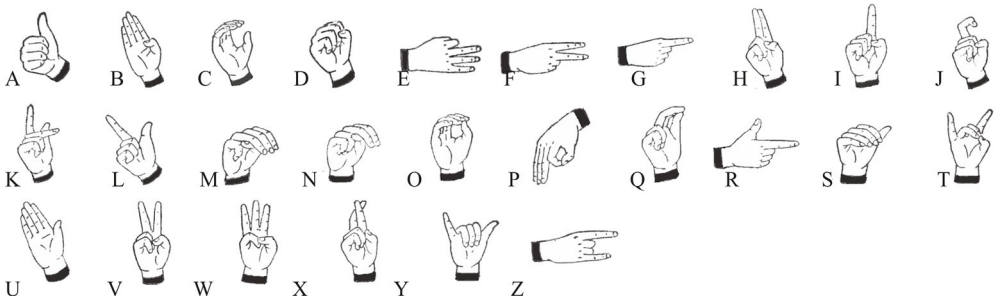


Fig.1 Sign language of 26 letters

Sign language learning is usually done through video learning. However, because the traditional video is two-dimensional, some details of the gesture and mouth movement are not easy to observe. Recently, in the fields of continuous deformation detection, high speed online detection, human 3-D detection and bionic design, fringe projection dynamic 3-D shape measurement has been widely and deeply studied[1-3]. Therefore, we can use dynamic fringe projection profilometry to record the basic gesture and mouth movements of 26 letters, because the 3-D measurement video is three-dimensional, the various angles of the mouth movement are easy to observe, and then the sign language learning will be more accurate and convenient.

According to the number of projected fringes needed to reconstruct a frame 3-D image, dynamic fringe projection profilometry can be divided into two categories: dynamic 3-D shape measurement based on single-fringe and multi-fringes. 3-D shape measurement method based on fringe projection generally includes two steps: wrapped phase calculation and unwrapped phase calculation. Wrapped phase calculation methods include phase-shifting profilometry (PSP)[4,5], Fourier transform profilometry (FTP)[6,7], and wavelet transform profilometry

* This work has been supported by the National Natural Science Foundation of China (No.51275405).
** E-mail: wjh051130@163.com

(WTP)[8-13]. FTP and WTP only need one fringe to obtain the wrapped phase, so it can be used for the dynamic 3-D shape measurement of continuously changing objects. However, the measurement accuracy may be affected by noise, shadow and height jump. PSP has high precision and strong anti-interference ability, but its measurement speed is limited due to the multi-fringes projection.

Since PSP, FTP and WTP use the arctangent function to extract phase information, the phase is wrapped into a range from $-\pi$ to π, which needs to be unwrapped in the next step. There are two categories of methods exist for phase unwrapping: temporal phase unwrapping method and spatial phase unwrapping method. Spatial phase unwrapping method only needs one wrapped phase to get the unwrapped phase, if we combine it with FTP or WTP, the dynamic three-dimensional measurement based on single-fringe can be realized. However, it requires that the captured fringes satisfy the Nyquist sampling theorem. Due to the noise, shadow, or height jump, however, this condition is difficult to meet, which gives rise to unwrapped phase error[14]. Temporal phase unwrapping method projects multiple sets of fringes with different frequencies along the time coordinate and unwraps the phase of each pixel individually. This method has the advantages of robustness, high accuracy and good noise immunity, it can also be used to measure complex objects with height jump[15,16]. However, it requires multi-wrapped phases to get the unwrapped phase.

So far, there are typically three coding methods for temporal phase unwrapping, as shown in Fig.2. In this paper, multi-frequency fringe coding is adopted, as shown in Fig.3(a). The gray-code mode is a typically binary coding method, as shown in Fig.3(b). In this approach, fringe orders of the wrapped phase are encoded within a serial binary gray-code pattern over time. Since N patterns can only code 2^N fringe orders, the minimum number of additional binary patterns required is INT($\log_2 f$)+1, where f is the total number of fringe orders within the full phase map and INT is the 'integer part' function.

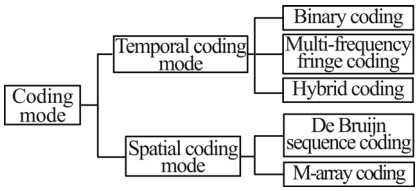


Fig.2 Coding mode

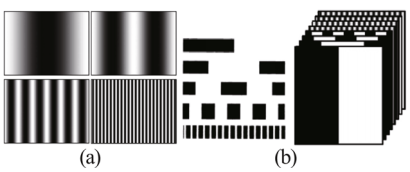


Fig.3 Multi-frequency fringe coding and gray-code method: (a) Multi-frequency fringe; (b) Gray-code

Wang and Zhang[17] presented a high speed multi-frequency phase-shifting technique where nine patterns were projected to recover a dynamic scene. Then, to project fewer patterns, Zuo et al[18] firstly proposed a four-frame pattern projection strategy to measure dynamic isolated objects in real time and then developed a bi-frequency phase-shifting algorithm based on tri-polar pulse width modulation with which one kilohertz measuring speed was reached[19]. In addition, Zhang et al[20] reported a color fringe projection technique with which fringe images were encoded into different color channels of a same image for the increase of the projection efficiency. Except to use the phase information as a cue for surface reconstruction, some researchers utilize random statistical patterns to label a pixel where the codification is performed based on the uniqueness of captured intensities over time. Schaer et al[21] proposed to project randomly distributed laser speckles to encode a surface. Then a temporal correlation was used to determine the corresponding pixels.

For the continuously changing objects with height jump, it is better to use 3-D shape measurement based on multi-fringes. However, multi-fringes projection will seriously affect the measurement results, which requires that the speed of digital light procession (DLP) projection and camera acquisition should be high enough. In the fringe projection profilometry, the camera's acquisition speed has reached a high level, but the refreshing frequency of DLP has been limited. The refreshing frequency of DLP is related to the depth of projected fringes. Such as DLP LightCraft 4500 of Texas Instruments (TI), its shortest refreshing frequency of sinusoidal fringe with a 8 bit depth is 120 Hz (i.e. 8 333 μs), but its shortest refreshing frequency of sinusoidal fringes with 1 bit depth is up to 4 225 Hz (i.e. 235 μs). Therefore, in order to improve the speed of three-dimensional measurement and to realize the dynamic three-dimensional measurement of continuously changing objects, the binary fringe is needed. However, its disadvantage is that the higher harmonic component is easily introduced into the binary fringe, which seriously affects the precision of the three-dimensional measurement. Therefore, we need to defocus the DLP to filter out the higher harmonics and to approximate the 256 level grayscale sinusoidal fringe to the greatest extent. The advantage of the binary fringes is that there is no phase error caused by the nonlinear response of the measurement system[22].

Fig.4 shows the typical structure of the fringe projection profilometry. In Fig.4, E_C and E_P are optical centers of camera lens and projector lens, respectively. $Oxyz$ is reference coordinate, d_0 is the distance between E_C and E_P, l_0 is the distance between E_P and reference plane. PP' is height of a point P on object surface. Since the ΔPAB and ΔPE_PE_C are similar triangles, PP' can be expressed as

$$\frac{AB}{E_C E_P}=\frac{PP'}{PP'+l_0}\quad , \tag{1}$$

$$PP' = \frac{AB}{d_0 - AB} l_0 . \tag{2}$$

Assuming that the phase information at point O is 0, the phase information at point A can be expressed as

$$\varphi_A = \frac{2\pi \overline{OA}}{\lambda_0} , \tag{3}$$

where λ_0 is the pitch of grating fringe.

Similarly, the phase information at point B can be expressed as

$$\varphi_B = \frac{2\pi \overline{OB}}{\lambda_0} . \tag{4}$$

According to Eqs.(3) and (4), AB is obtained and expressed as

$$AB = \overline{OB} - \overline{OA} = \frac{(\varphi_B - \varphi_A)\lambda_0}{2\pi} . \tag{5}$$

We substitute Eq.(5) in Eq.(2)

$$PP' = \frac{(\varphi_B - \varphi_A)}{2\pi d_0 / \lambda_0 - (\varphi_B - \varphi_A)} l_0 . \tag{6}$$

φ_A represents the phase information of the point A on reference plane, φ_B represents the phase information of the point B on object surface. Therefore, if the absolute phases φ_A and φ_B are calculated, the object height can be obtained as well.

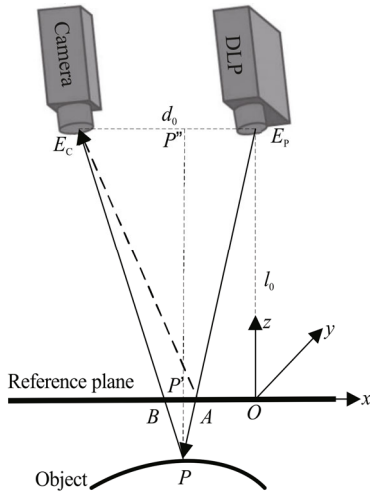


Fig.4 The structure of the fringe projection profilometry

When phase information is calculated by WTP and spatial phase unwrapping method, the accuracy of phase calculation is vulnerable to the noise, shadow, or height jump. Therefore, the PSP and the temporal phase unwrapping method are used to calculate phase information. However, a large number of projection fringes are needed. In order to realize the 3-D measurement of dynamic objects, it is necessary to realize the projection and acquisition of multiple images in a short time. At present, the acquisition speed of industrial cameras is very high, but the refresh frequency of projectors is limited. To this end, the high-speed projection technology based on fringe binarization technology and projector defocusing technology are innovatively proposed. After fringe binarization, the refresh frequency of DLP Lightcraft 4500 can be increased from 120 Hz to 4 225 Hz. However, the binarization of 8-bit gray fringes will introduce high-order harmonic components, resulting in phase calculation errors. Theoretical analysis and experiments show that the defocusing technology of projector can be used to solve the problem.

The proposed method consists of four steps, as shown in Fig.5.

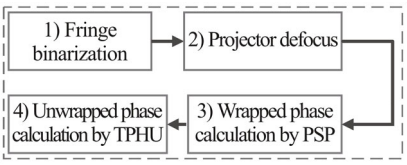


Fig.5 The flow chart of the proposed method

There are four commonly used fringe binarization methods.

The principle of square wave binary fringe is as follows. By setting the threshold, the 256 level grayscale sinusoidal fringe will be is converted into to a binary fringe, and then defocused by DLP, so as to approximate the original sinusoidal fringe.

$$f_{\text{Square}}(x, y) = \begin{cases} 0, & (y \bmod T) \le \dfrac{T}{2} \\ 1, & (y \bmod T) > \dfrac{T}{2} \end{cases} , \tag{7}$$

where f_{Square} is the binary fringe, T is the period of fringe. It is worth noting that the value of T is a multiple of N when the N-step phase-shifting method is adopted.

Fig.6(a) is binary fringe based on the square wave when the sinusoidal fringe period is 76 pixels, and Fig.6(b) is a row of pixels in Fig.6(a). Fig.6(c) is binary fringe based on the square wave when the sinusoidal fringe period is 16 pixels, and Fig.6(d) is a row of pixels in Fig.6(c).

The binary projection fringe based on the square wave is obviously deviated from the sinusoidal character. The degree of the approximate sinusoidal fringe depends on the DLP defocus degree. The larger the defocus, the closer the sinusoidal fringe is, but the contrast decreases with the increase of defocus.

The principle of SPWM binary fringe is to compare the original sinusoidal fringe with a triangular wave fringe, when the grayscale value of the sinusoidal fringe is greater than that of the triangular wave fringe, the value of SPWM binary fringe is 1, on the other hand, its value is 0. Therefore the SPWM binary fringe can be generated.

$$f_{\text{SPWM}}(x, y) = \begin{cases} 0, & I_{\text{sin}}(x, y) < I_{\text{tri}}(x, y) \\ 1, & I_{\text{sin}}(x, y) \ge I_{\text{tri}}(x, y) \end{cases} , \tag{8}$$

where f_{SPWM} is the binary fringe, I_{sin} is sinusoidal fringe, and I_{tri} is triangular wave fringe.

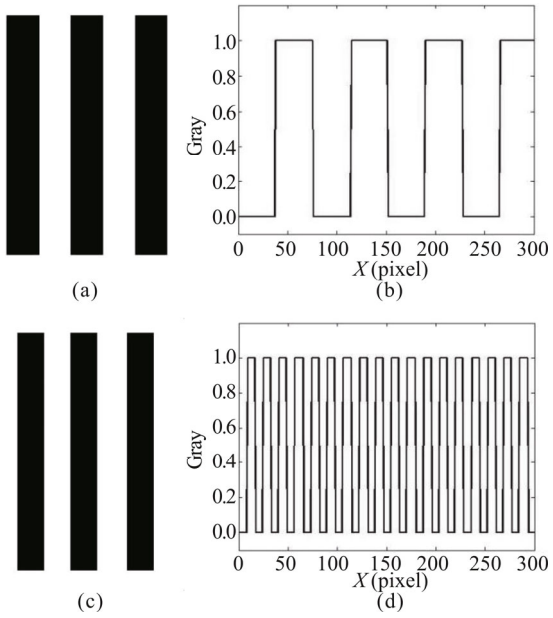


Fig.6 Square wave binary fringes: (a) Period of 76 pixels; (b) A row in (a); (c) Period of 16 pixels; (d) A row in (c)

Fig.7(a) is SPWM binary fringe when the sinusoidal fringe period is 76 pixels. Fig.7(b) is a row of pixels in Fig.7(a). Fig.7(c) is SPWM binary fringe when the sinusoidal fringe period is 16 pixels. Fig.7(d) is a row of pixels in Fig.7(c). The SPWM binary fringe is not too bad when the sinusoidal fringe period is relatively large, as shown in Fig.7(a). It is unacceptable when the sinusoidal fringe period is small, as shown in Fig.7(c). However, in order to ensure the accuracy of three-dimensional measurement, the period of projection sinusoidal fringe is generally small, so the application of SPWM binary fringe is less.

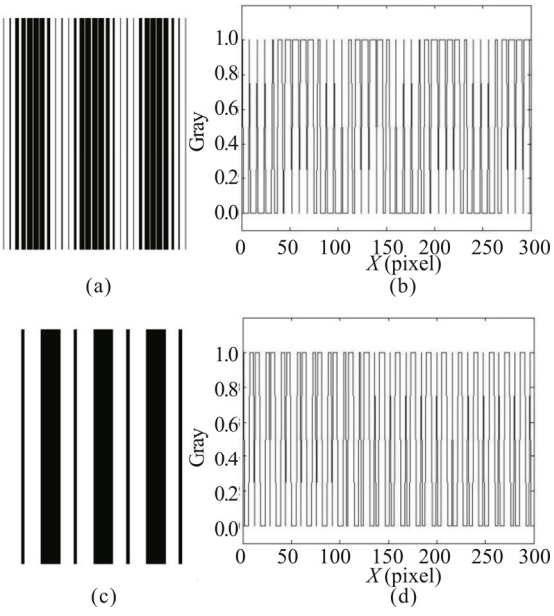


Fig.7 SPWM binary fringes：(a) Period of 76 pixels; (b) A row in (a); (c) Period of 16 pixels; (d) A row in (c)

The principle of Bayer dithered binary fringe is as follows. Firstly, the normalized Bayer matrix and sinusoidal fringe are generated. Then, the normalized sinusoidal fringe is compared with the corresponding element values in the normalized Bayer matrix. If the gray value of a pixel point in the sinusoidal fringe is larger than the corresponding element value in the Bayer matrix, the value is 1, otherwise it is 0.

The 2×2 Bayer matrix is expressed as:

$$\boldsymbol{M}_1 = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} . \tag{9}$$

This is the basic Bayer matrix, and the higher-order Bayer matrix can be obtained from the basic matrix iteration. It is expressed is:

$$\boldsymbol{M}_{n+1} = \begin{bmatrix} 4\boldsymbol{M}_n & 4\boldsymbol{M}_n + 2\boldsymbol{E}_n \\ 4\boldsymbol{M}_n + 3\boldsymbol{E}_n & 4\boldsymbol{M}_n + \boldsymbol{E}_n \end{bmatrix} , \tag{10}$$

where $\boldsymbol{E}_n$ is an n identity matrix. The 4×4 Bayer matrix and 8×8 Bayer matrix are expressed as:

$$\boldsymbol{M}_2 = \begin{bmatrix} 0 & 8 & 2 & 10 \\ 12 & 4 & 14 & 6 \\ 3 & 11 & 1 & 9 \\ 15 & 7 & 13 & 5 \end{bmatrix} ,$$

$$\boldsymbol{M}_3 = \begin{bmatrix} 0 & 32 & 8 & 40 & 2 & 34 & 10 & 42 \\ 48 & 16 & 56 & 24 & 50 & 18 & 58 & 26 \\ 12 & 44 & 4 & 36 & 14 & 46 & 6 & 38 \\ 60 & 28 & 52 & 20 & 62 & 30 & 54 & 22 \\ 3 & 35 & 11 & 43 & 1 & 33 & 9 & 41 \\ 51 & 19 & 59 & 27 & 49 & 17 & 57 & 25 \\ 15 & 47 & 7 & 39 & 13 & 45 & 5 & 37 \\ 63 & 31 & 55 & 23 & 61 & 29 & 53 & 21 \end{bmatrix} . \tag{11}$$

The normalization of Bayer matrix is to divide all elements by 4^n. Fig.8(a) is Bayer dithered binary fringe when the sinusoidal fringe period is 76 pixels, and Fig.8(b) is Bayer dithered binary fringe when the sinusoidal fringe period is 16 pixels.

The principle of dithered binary fringe based on error diffusion is to assign the error produced in the quantization process to a certain proportion of the pixels around it. Firstly, the sinusoidal fringe is normalized and image I is obtained; then a threshold T (such as T=0.5) is set, and the binary fringe B is obtained. Finally, the error between I and B is calculated. By using a diffusion core, the error is assigned to the unprocessed pixels on the right, below, lower left, and lower right side. There are many kinds of diffusion cores, and we choose the Floyd-Steinberg diffusion core here, then the formula is:

$$\boldsymbol{FS} = \begin{bmatrix} - & * & \frac{7}{16} \\ \frac{3}{16} & \frac{5}{16} & \frac{1}{16} \end{bmatrix} , \tag{12}$$

where symbols - is the left side pixel that do not require error allocation, symbols * is the pixel being processed, the right pixel allocation error is 7/16, the lower left pixel distribution error is 3/16, the lower pixel distribution error

is 5/16, and the right lower pixel distribution error is 1/16.

Fig.8(c) is the dithered binary fringe based on the error diffusion when sinusoidal fringe period is 76 pixels, and Fig.8(d) is the dithered binary fringe based on the error diffusion when sinusoidal fringe period is 16 pixels.

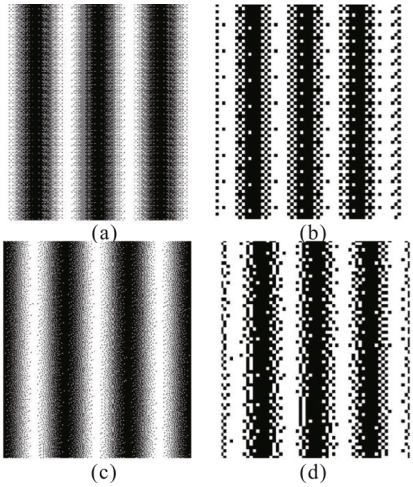


Fig.8 Dithered binary fringes: (a) Period of 76 pixels based on Bayer dithered; (b) Period of 16 pixels based on Bayer dithered; (c) Period of 76 pixels based on error diffusion dithered; (d) Period of 16 pixels based on error diffusion dithered

The binary fringe is $f(x, y)$, if the fringe after projector defocusing is expressed as $g(x, y)$, the point spread function of the projector after defocusing can be written as a two-dimensional Gauss function:

$$h(x,y)=\frac{1}{2\pi\sigma^2}\exp(-\frac{x^2+y^2}{2\sigma^2}) \quad , \tag{13}$$

where $h(x,y)$ is the two-dimensional Gauss function, and σ is the standard deviation of Gauss function.

The binary fringe after the DLP defocusing can be expressed as:

$$g(x,y)=f(x,y)*h(x,y)\,, \tag{14}$$

where symbols * is convolution operator.

Fourier transform is applied to the Eq.(7), and the expression in frequency domain is:

$$G(u,v)=F(u,v)H(u,v)\,, \tag{15}$$

$$H(u,v)=\exp(-\frac{1}{2}(u^2+v^2)\sigma^2)\,. \tag{16}$$

Since $H(u, v)$ is a Gauss function, it is known from the characteristics of Gauss's function that the defocused optical system is equivalent to a low pass filter, which can filter out the high harmonic components, and the width of the filter is mainly determined by σ. $\sigma = kR$, where k is scale coefficient which is greater than zero, R is the radius of the fuzzy circle, and the fuzzy circles on the image surface after the object is defocused.

$$R=\frac{D}{2}s[f^{-1}-u^{-1}-s^{-1}] \quad , \tag{17}$$

where f is the focal length of the projector, u is the object distance, s is the distance from the fuzzy circle to the projector lens (i.e. the distance of the defocus), and D is the diameter of the projector lens.

According to the imaging formula of thin lens:

$$f^{-1}=u^{-1}+v^{-1}\,. \tag{18}$$

Substituting Eq.(18) in Eq.(17), we can obtain:

$$R=\frac{D}{2}s[v^{-1}-s^{-1}] \quad . \tag{19}$$

We can see that σ is proportional to R and s. The bigger the defocus distance s is, the larger the σ, and the narrower the filter, the more high order harmonics of the filter. The smaller defocus distance s is, the smaller the σ, the wider the filter, the less high order harmonics of the filter.

Figs.9(a)—(d) are non-defocused fringe and defocused fringe captured by camera, Figs.9(e)—(h) are the 3-D measurement results comparison between them. It can be seen from the comparison that the measured results after fringe defocused are obviously better than that by non-defocused.

In order to achieve a quantitative comparison, we adopt the twenty-step phase-shifting algorithm and temporal phase unwrapping method to calculate the ideal phase information, and utilize root mean square error (*RMSE*) to evaluate phase precision. The comparison is shown in Tab.1. *RMSE* can be expressed as:

$$RMSE=\sqrt{\frac{\sum_{y=0}^{N-1}\sum_{x=0}^{M-1}(\varphi_i(x,y)-\varphi_r(x,y))^2}{MN}}\,, \tag{20}$$

where $\varphi_i(x,y)$ is the ideal phase information, $\varphi_r(x,y)$ is the real phase information, and M and N are the numbers of pixels in the x-axis and y-axis, respectively.

In PSP, multiple phase-shifting sinusoidal fringes are projected sequentially onto the object surface, and the distorted fringes are captured by a CCD camera. The ith fringe pattern, $g_i(x, y)$, can be presented as:

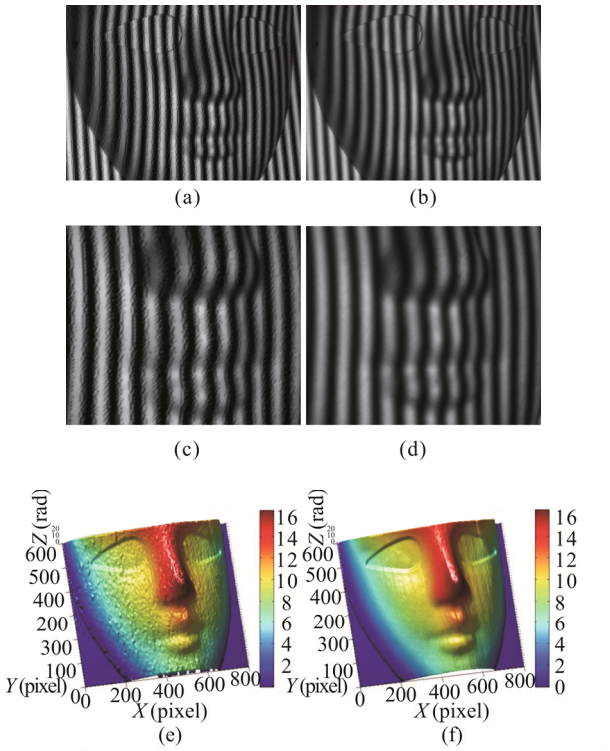
$$g_i(x,y)=a(x,y)+b(x,y)\cos[\varphi(x,y)+2\pi(i-1)/N]\,, \tag{21}$$

where $a(x,y)$ is the average intensity relating to pattern brightness and background illumination, $b(x,y)$ is the intensity modulation relating to pattern contrast and surface reflectivity, $\varphi(x,y)$ is the phase modulation of the fringes (the required phase distribution), i is a phase-shifting index and N is the total number of phase-shifting steps.

The wrapped phase can be extracted with the following equation:

$$\psi(x,y)=-\arctan\left[\frac{\sum_{i=1}^{N}g_i(x,y)\sin(2\pi(i-1)/N)}{\sum_{i=1}^{N}g_i(x,y)\cos(2\pi(i-1)/N)}\right], \tag{22}$$

where $\psi(x,y)$ is the wrapped phase and limited between $-\pi/2$ and $\pi/2$ due to the application of the arctangent function, but because in our case we consider the sign of real and imaginary parts, $\psi(x,y)$ will be limited between $-\pi$ and π. There are two kinds of modulo operator in MATLAB, that is, mod and rem, rem is adopted here.



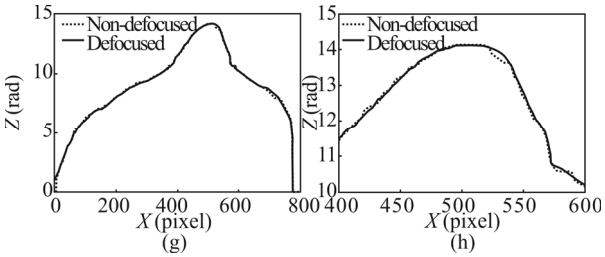


Fig.9 Measurement results comparison between non-defocused fringe and defocused fringe: (a) Non-defocused fringe; (b) Defocused fringe; (c) Local non-defocused fringe; (d) Local defocused fringe; (e) Measurement result of non-defocused fringe; (f) Measurement result of defocused fringe; (g) Comparison in a row; (h) Local comparison in a row

When N=3, the PSP is called a three-step PSP, when N=4 and N=5, the PSPs are called a four-step PSP and a five-step PSP, respectively. The wrapped phase can be represented as:

$$\psi_{3\text{-step}}(x,y)=\arctan\left[\frac{\sqrt{3}(g_2(x,y)-g_3(x,y))}{2g_1(x,y)-g_2(x,y)-g_3(x,y)}\right], \quad (23)$$

$$\psi_{4\text{-step}}(x,y)=\arctan\left[\frac{g_4(x,y)-g_2(x,y)}{g_1(x,y)-g_3(x,y)}\right], \quad (24)$$

$$\psi_{5\text{-step}}(x,y)=\arctan\left[\frac{\sin(\frac{2\pi}{5})(g_2(x,y)-g_5(x,y))+\sin(\frac{4\pi}{5})(g_3(x,y)-g_4(x,y))}{g_1(x,y)+\cos(\frac{2\pi}{5})(g_2(x,y)+g_5(x,y))+\cos(\frac{4\pi}{5})(g_3(x,y)+g_4(x,y))}\right]. \quad (25)$$

The three pitches heterodyne unwrapping method combines the three pitches method with the heterodyne method. It has high efficiency while guaranteeing the higher reliability of the unwrapping phase. The frequencies of the grating fringes of the three pitches heterodyne unwrapping method are $f_1=S+\sqrt{S}+1$, $f_2=S$, $f_3=S-\sqrt{S}$. The principle of the three pitches heterodyne method is that the phase-shifting method is used to obtain the wrapped phases of the three sets of grating fringes first, then the wrapped phase of $f_{12}=\sqrt{S}+1$ is obtained from $f_1=S+\sqrt{S}+1$ and $f_2=S$ by using the heterodyne principle, the wrapped phase of $f_{23}=\sqrt{S}$ is obtained from $f_2=S$ and $f_3=S-\sqrt{S}$ by using the heterodyne principle and, finally, the wrapped phase of $f_{123}=1$ is obtained from $f_{12}=\sqrt{S}+1$ and $f_{23}=\sqrt{S}$ by using the heterodyne principle. Since $\varphi(f_{123})=\psi(f_{123})$, that is, the wrapped phase of $f_{123}=1$ is also the unwrapped phase, an unwrapped phase of higher frequency can be obtained by using Eq.(20).

Tab.1 The phase accuracy comparison between defocusing and non-defocusing fringes

Fringes	Non-defocusing fringes	Defocusing fringes
RMSE	0.152 4	0.078 5

$$\begin{cases}\varphi(1)=\psi(1)\\ \varphi(\sqrt{S}+1)=\psi(\sqrt{S}+1)+2\pi\times round\left[\dfrac{(\sqrt{S}+1)\varphi(1)-\psi(\sqrt{S}+1)}{2\pi}\right]\\ \varphi(\sqrt{S})=\psi(\sqrt{S})+2\pi\times round\left[\dfrac{\sqrt{S}\varphi(1)-\psi(\sqrt{S})}{2\pi}\right]\\ \varphi(S+\sqrt{S}+1)=\psi(S+\sqrt{S}+1)+2\pi\times round\left[\dfrac{(S+\sqrt{S}+1)\varphi(\sqrt{S}+1)/(\sqrt{S}+1)-\psi(S+\sqrt{S}+1)}{2\pi}\right]\\ \varphi(S)=\psi(S)+2\pi\times round\left[\dfrac{S\varphi(\sqrt{S}+1)/(\sqrt{S}+1)-\psi(S)}{2\pi}\right]=\psi(S)+2\pi\times round\left[\dfrac{S\varphi(\sqrt{S})/\sqrt{S}-\psi(S)}{2\pi}\right]\\ \varphi(S-\sqrt{S})=\psi(S-\sqrt{S})+2\pi\times round\left[\dfrac{(S-\sqrt{S})\varphi(\sqrt{S})/\sqrt{S}-\psi(S-\sqrt{S})}{2\pi}\right]\end{cases}, \quad (26)$$

where $\varphi(f_i)$ is the unwrapped phase of the phase-shifting fringes with a frequency of f_i, $\psi(f_i)$ is the wrapped phase of phase-shifting fringes with a frequency of f_i.

$\varphi(S+\sqrt{S}+1)$, $\varphi(S)$ and $\varphi(S-\sqrt{S})$ can all be regarded as the final unwrapped phases. In addition, it is also possible to utilize the above intermediate phase information to calculate the phase slope by using a least squares fitting method, and thus the final unwrapped phase can be represented as:

$$\varphi(S)=s\frac{\sum_{i=1}^{6}S_i\varphi(S_i)}{\sum_{i=1}^{6}(S_i)^2}. \tag{27}$$

When the N-step phase-shifting method is used to obtain the wrapped phase, the square wave binary fringe method requires the square wave period to be N multiples. Obviously, it can not be applied to the three pitches heterodyne method. The period of triangular wave fringe in the SPWM binary fringe method must be a multiple of 2, and the period of sinusoidal fringe must be an integer multiple of the triangular wave fringe. Therefore, the SPWM binary fringe method can also not be applied to the three pitches heterodyne method. There are no requirements for the period of sinusoidal fringe by the Bayer dithered binary fringe and the dithered binary fringe based on error diffusion, especially in the lower period of the period, which can still provide high quality fringe pattern. Therefore, dithered binary fringe can be used in the three pitches heterodyne method.

Using the three pitches heterodyne method to obtain the phase information, we compare the measurement accuracy of three kinds of projected fringes, i. e. the 256 level grayscale sinusoidal fringe, the dithered binary fringe based on error diffusion and the Bayer dithered binary fringe, as shown in Fig.10. By contrast, the measurement result of using the Bayer dithered binary fringe as the projected fringe is poor, but the result of the dithered binary fringe based on error diffusion is close to that of the original sinusoidal fringe. So we choose the dithered binary fringe based on error diffusion to realize the high-speed dynamic 3-D shape measurement.

The structure of the 3-D measurement system is shown in Fig.11, including the DLP LightCraft 4500, the camera MER-050-560U3M, the single-chip microcomputer (SCM) and the personal computer. The projection fringes are stored in the DLP LightCraft 4500. The SCM is used to control the beginning and the end of the DLP projection and the camera acquisition. The computer reconstructs the 3-D shape of the object after the analysis and processing of the captured fringes. The synchronous trigger signal is sent to the I/O of the camera through the external output of DLP LightCraft 4500. At the same time, DLP Lightcraft 4500 adopts pattern sequence, its pattern exposure and pattern period must be with the exposure time of the camera.

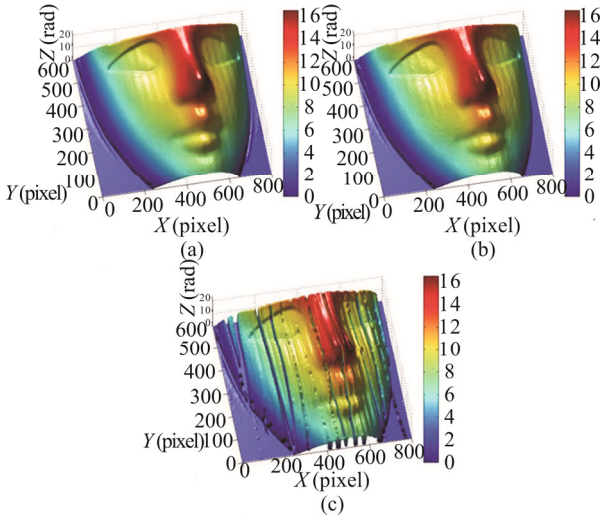


Fig.10 Phase information calculation by using three pitches heterodyne method employing different kinds of fringes: (a) 256 level grayscale sinusoidal fringe; (b) Dithered binary fringe based on error diffusion; (c) Bayer dithered binary fringe

The maximum frequency of the camera (MER-050-560 U3M) is 560 frames per second. In this experiment, the acquisition speed of the camera is set to 500 frames per second, and the refreshing frequency of DLP LightCraft 4500 is set to 500 Hz, the DLP and camera work synchronously, therefore, the time consumption of a fringe image is only 0.002 s.

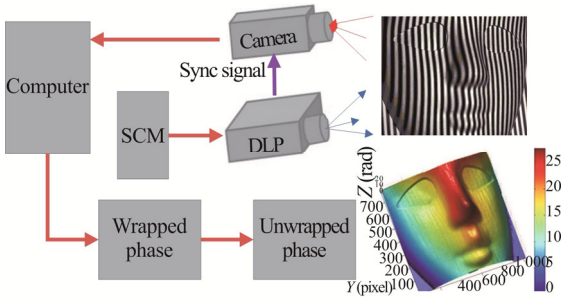


Fig.11 The hardware structure of the measurement system

When the DLP Lightcraft 4500 has the highest refresh frequency, two 24-bit deep fringes can be projected at most and the image size must be 912×1 140 pixels. If the four-step phase-shifting method is used to get the wrapped phase and the three pitches heterodyne method is used to get the unwrapped phase, the number of projected fringes is 12, so we should combine 12 binary fringes into a 24-bit fringe.

Fig.12 shows the 12 binary fringes and a 24-bit fringe synthesized by them. Formula is expressed as:

$$\begin{cases} f_r(x,y) = 2^0\left[f_{\text{binary\_1}}(x,y)\right] + 2^1\left[f_{\text{binary\_2}}(x,y)\right] + 2^2\left[f_{\text{binary\_3}}(x,y)\right] + 2^3\left[f_{\text{binary\_4}}(x,y)\right] + \\ \qquad 2^4\left[f_{\text{binary\_5}}(x,y)\right] + 2^5\left[f_{\text{binary\_6}}(x,y)\right] + 2^6\left[f_{\text{binary\_7}}(x,y)\right] + 2^7\left[f_{\text{binary\_8}}(x,y)\right] \\ f_g(x,y) = 2^0\left[f_{\text{binary\_9}}(x,y)\right] + 2^1\left[f_{\text{binary\_10}}(x,y)\right] + 2^2\left[f_{\text{binary\_11}}(x,y)\right] + 2^3\left[f_{\text{binary\_12}}(x,y)\right] \end{cases}, \tag{28}$$

where f_r is the red component of 24-bit color fringe, f_g is the green component of 24-bit color fringe, and $f_{\text{binary\_1}}$, $f_{\text{binary\_2}}$, …, $f_{\text{binary\_12}}$ are 12 binary fringes.

Sign language of 26 letters includes gesture and mouth movement. We randomly select letters B, F, Q and V for gestures 3-D measurement, and letter V for the dynamic 3-D measurement of mouth movement. The 3-D measurement results of four gestures are shown in Fig.13, the dynamic 3-D measurement results of mouth movement are shown in Fig.14. We can see from Fig.13 that the 3-D measurement results of gestures are of high accuracy and more detailed information than 2-D photographs. We can also see from Fig.14 that for the mouth movement of the letter V pronunciation, the proposed method can well measure its dynamic process and help people learn sign language better.

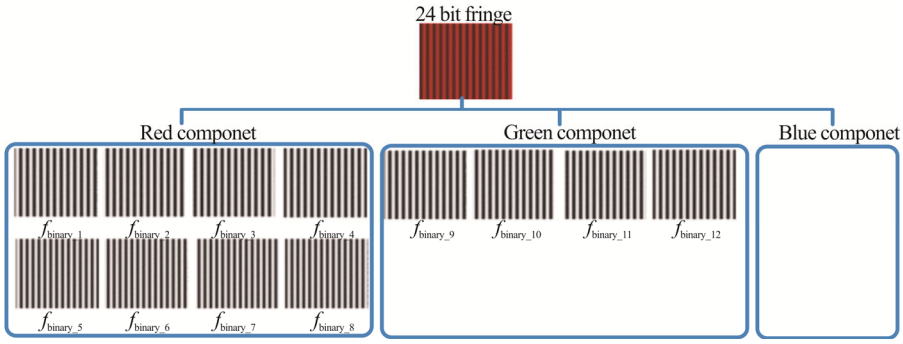


Fig.12 12 binary fringes and a 24-bit fringe synthesized by them

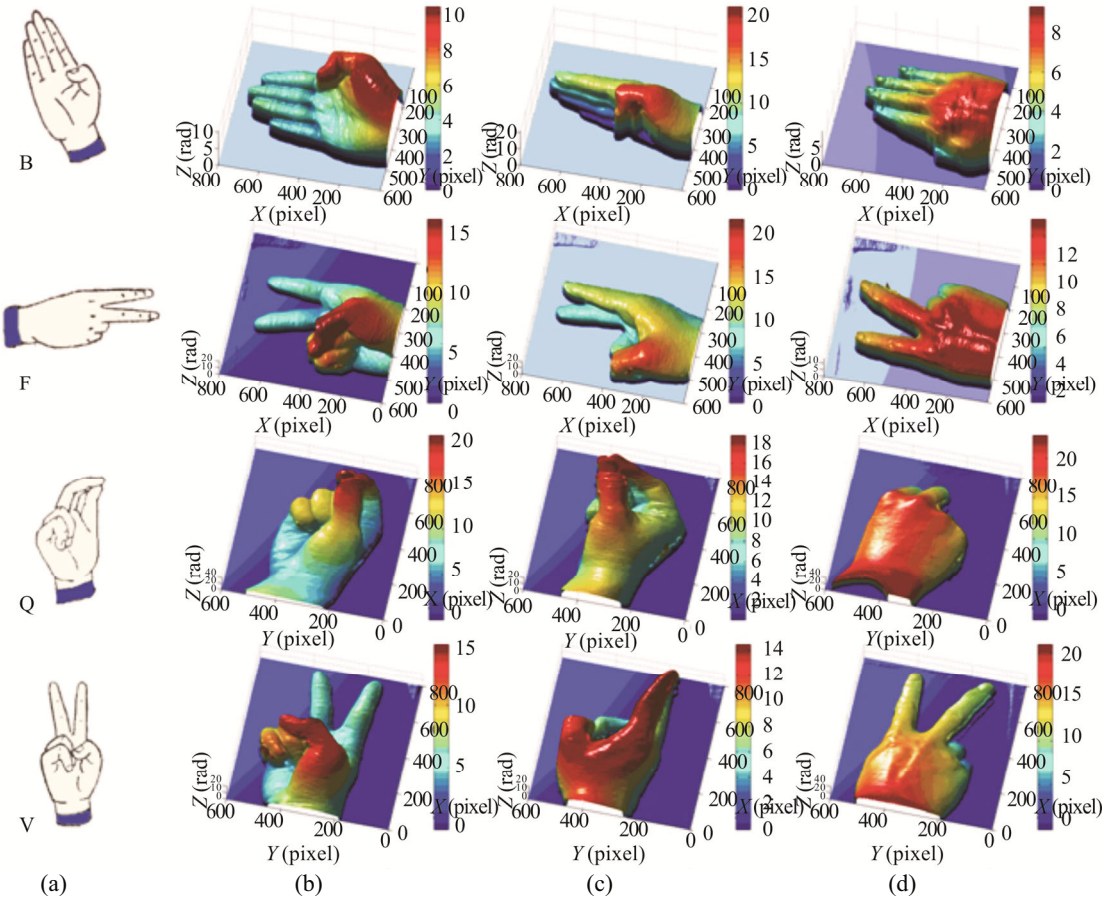


Fig.13 The 3-D measurement results of four gestures: (a) Four 2-D gestures; (b) Front view of four 3-D gestures; (c) Side view of four 3-D gestures; (d) Back view of four 2-D gestures

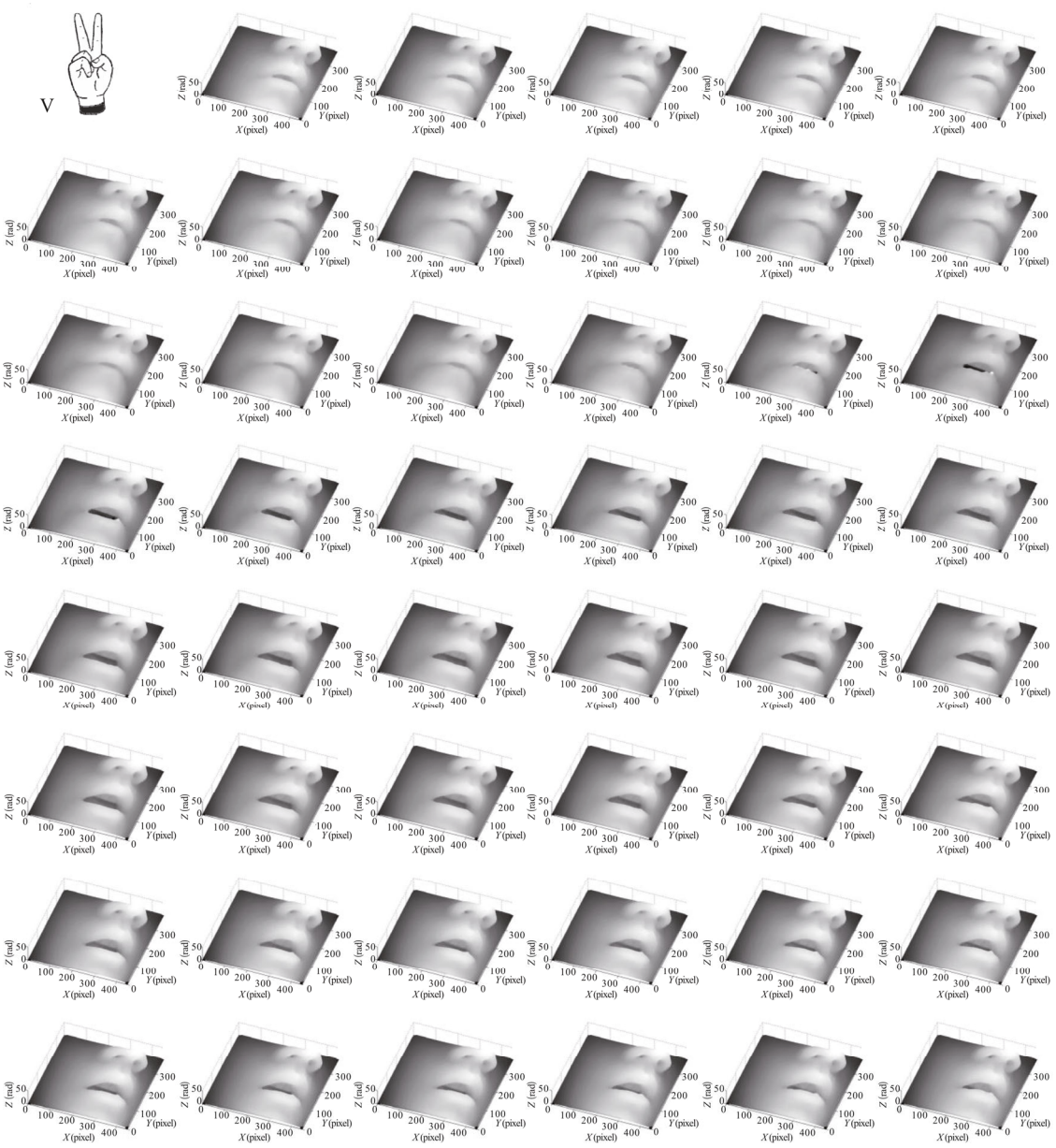


Fig.14 The dynamic 3-D measurement results of the mouth movement of the letter V pronunciation

This paper has presented a high-speed fringe projection profilometry employing binary fringe to record the 3-D sign language, including gesture and mouth movement, to help people learn the sign language. Firstly, the 256 level grayscale sinusoidal fringes are transformed to binary fringes by the appropriate method to achieve high-speed projection. Secondly, the method of defocusing of binary fringes is analyzed. Thirdly, the phase-shifting and the three pitches heterodyne method are used to calculate phase information with high accuracy. At last, the composition of the high-speed measurement system is introduced, and we present how to combine multiple binary fringes into a 24-bit fringe. Experiments have shown that the proposed system can acquire and display high-quality 3D gesture and mouth movement at a speed of 500 frames per second, which can help people learn the sign language with more accurate and convenient.

References

[1] S. J. Feng, C. Zuo, T. Y. Tao, Y. Hu, M. L. Zhang, Q. Chen and G. H. Gu, Optics and Lasers in Engineering **103**, 127 (2018).

[2] S. J. Feng, Q. Chen, C. Zuo, T. Y. Tao, Y. Hu and A. Asundi, Optics Express **25**, 540 (2017).

[3] T. Y. Tao, Q. Chen, J. Da, S. J. Feng, Y. Hu and C. Zuo, Optics Express **24**, 20253 (2016).

[4] S. Zhang and S.-T. Yau, Optics Express **14**, 2644 (2006).

[5] S. Y. Gai and F. P. Da, Optics and Lasers in Engineering

4, 205 (2010).
[6] M. Takeda, H. Ina and S. Kobayashi, Journal of the Optical Society of America **72**, 156 (1982).
[7] M. Takeda and K. Mutoh, Applied Optics **22**, 3977 (1983).
[8] Z. B. Zhang and J. G. Zhong, Optics Express **21**, 18777 (2013).
[9] A. Abid, M. Gdeisat, D. Burton, M. Lalor and F. Lilley, Applied Optics **46**, 6120 (2007).
[10] M. Zhong, F. Chen, C. Xiao and Y. C. Wei, Optics and Lasers in Engineering **88**, 243 (2017).
[11] M. A. Gdeisata, A. Abid, D. R. Burton, M. J. Labor and F. Lilley, Optics and Lasers in Engineering **47**, 1348 (2009).
[12] Z. Y. Wang, J. Ma and V. Minh, Optics and Lasers in Engineering **50**, 1052 (2012).
[13] J. Ma, Z. Y. Wang and T. Y. Pan, Optics and Lasers in Engineering **55**, 205 (2014).
[14] G. L. Du, M. M. Wang, C. L. Zhou, S. C. Si, H. Li, Z. K. Lei and Y. J. Li, Journal of Modern Optics **64**, 231 (2017).
[15] C. Zuo, L. Huang, M. L. Zhang, Q. Chen and A. Asundi, Optics and Lasers in Engineering **85**, 84 (2016).
[16] S. Xing and H. Guo, Applied Optics **56**, 1591 (2017).
[17] Y. Wang and S. Zhang, Optics Express **19**, 5149 (2011).
[18] C. Zuo, Q Chen, G. Gu, S. Feng and F. Feng, Optics Express **20**, 19493 (2012).
[19] Z. Zhang, C. Towers and D. P. Towers, Optics Express **14**, 6444 (2006).
[20] M. Schaer, M. Grosse and R. Kowarschik, Applied Optics **49**, 3622 (2010).
[21] C. Zuo, Q. Chen, G. Gu, S. Feng, F. Feng, R. Li and G. Shen, Optics and Lasers in Engineering **51**, 953 (2013).
[22] P. S. Huang, Q. J. Hu and F. P. Chiang, Applied Optics **41**, 4503 (2002).