# Projective Parallel Single-pixel Imaging to Overcome Global Illumination in 3D Structure Light Scanning: Supplementary Material


Yuxi Li[1[0000−0002−1922−5811]], Huijie Zhao[2], Hongzhi Jiang[1], and Xudong Li[1]

[1] School of Instrumentation and Optoelectronic Engineering, Beihang University (BUAA), Beijing, China
uniluxli@qq.com, {jhz1862,xdli}@buaa.edu.cn
[2] Institute of Artificial Intelligence, Beihang University (BUAA), Beijing, China
hjzhao@buaa.edu.cn


## A Proof of Local Maximum Constraint Proposition

Local maximum constraint, which states that the correspondence matched point has to be a local maximum on the projection function, provides necessary condition for the location of candidate correspondence matching points when projection functions are captured. To prove the local maximum constraint proposition, two lemmas are required.

**Lemma 1.** *Shifting property of Radon transform. Suppose $f(x)$ is a two dimensional function with $x = (x, y)$, $\xi$ is a vector of the form $\xi$=$(cos\theta,sin\theta)^T$, $\delta$ is a vector of the form $\delta = (\delta_x, \delta_y)$, the Rodon transform of $f(x)$ is denoted as $\Re_\theta f(x) = \int f(x) \cdot \delta(\rho - x \cdot \xi)dx$= $\overset{\vee}{f}(\rho, \theta)$, then, the Radon transform of the shifted function $f(x - \delta)$ is $\Re_\theta f(x - \delta) = \overset{\vee}{f}(\rho - \delta \cdot \xi, \theta)$.*

Proof. From the definition of Radon transform, the Radon transform of the shifted function $f(x - \delta)$ is

$$\Re_\theta f(x - \delta) = \int f(x - \delta)\delta(\rho - x \cdot \xi)dx. \tag{13}$$

Let $y = x - \delta$, and substitute into Eq. (13)

$$\begin{aligned}\Re_\theta f(y) &= \int f(y)\delta(\rho - \delta \cdot \xi - y \cdot \xi)dy \\ &= \overset{\vee}{f}(\rho - \delta \cdot \xi, \theta).\end{aligned} \tag{14}$$

Thus, the lemma is proven. □

**Lemma 2.** *Derivative properties of Radon transform. The Radon transform of the derivative of original function has the form of*

$$\Re_\theta\{\frac{\partial f}{\partial x}\} = \cos\theta \frac{\partial \overset{\vee}{f}(\rho,\theta)}{\partial \rho}, \qquad \Re_\theta\{\frac{\partial f}{\partial y}\} = \sin\theta \frac{\partial \overset{\vee}{f}(\rho,\theta)}{\partial \rho}, \tag{15}$$

*and the Radon transform of the second derivative of original function has the form of*

$$\Re_\theta\{\frac{\partial^2 f}{\partial x^2}\} = \cos^2\theta \frac{\partial^2 \overset{\vee}{f}(\rho,\theta)}{\partial \rho^2}, \qquad \Re_\theta\{\frac{\partial^2 f}{\partial y^2}\} = \sin^2\theta \frac{\partial^2 \overset{\vee}{f}(\rho,\theta)}{\partial \rho^2}. \tag{16}$$

Proof. The two sub-equations of Eq. (15) and Eq. (16) can be proved in the same manner, thus only the first sub-equations of Eq. (15) and Eq. (16) are proved. From the definition of derivative

$$\frac{\partial f}{\partial x} = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon/\cos\theta\ , y) - f(x,y)}{\varepsilon/\cos\theta} \tag{17}$$

Taking Radon transform to both sides of Eq. (17), we have

$$\begin{aligned}\Re_\theta \frac{\partial f}{\partial x} &= \lim_{\varepsilon \to 0} \frac{\Re_\theta f(x+\varepsilon/\cos\theta\ , y) - \Re_\theta f(x,y)}{\varepsilon/\cos\theta} \\ &= \cos\theta \cdot \lim_{\varepsilon \to 0} \frac{\Re_\theta f(x+\varepsilon/\cos\theta\ , y) - \overset{\vee}{f}(\rho,\theta)}{\varepsilon}.\end{aligned} \tag{18}$$

From lemma 1, when $\delta$=$(-\varepsilon/\cos\theta\ , 0)$, Eq. (18) can be simplified as

$$\begin{aligned}\Re_\theta \frac{\partial f}{\partial x} &= \cos\theta \cdot \lim_{\varepsilon \to 0} \frac{\overset{\vee}{f}(\rho,\theta) - \overset{\vee}{f}(\rho,\theta)}{\varepsilon} \\ &= \cos\theta \frac{\partial \overset{\vee}{f}(\rho+\varepsilon,\theta)}{\partial \rho}.\end{aligned} \tag{19}$$

Obviously, if $\partial f/\partial x$ is treated as the original function $f$ in Eq. (15), Eq. (16) can be derived by using the result of Eq. (19). □

**Theorem 1.** *Local Maximum Constraint Proposition. The direct illumination point on the pixel transport image is a local maximum point on the projection functions, if the corresponding projection line does not pass through any speckles caused by global illumination.*

Proof. Suppose $(x_0, y_0)$ is the direct illumination point, which is a local maximum point on the pixel transport image, and the small circular region $\Omega_R$ with radius $R$ centered at $(x_0, y_0)$ is the effective region of the direct illumination [Fig.8 (a)]. This effective region is caused by blurring effect of the lens, thus the value of pixel transport image inside $\Omega_R$ can be considered as axial symmetric and concave function, thus a radius square $r^2{=}(x-x_0)^2+(y-y_0)^2$ is enough to parameterize the function inside $\Omega_R$

$$\begin{aligned}
&f(x,y){=}f(r^2),\\
&\frac{\partial f(x,y)}{\partial x}=2{f'}_{r^2}x,\\
&\frac{\partial f(x,y)}{\partial y}=2{f'}_{r^2}y,\\
&\frac{\partial^2 f(x,y)}{\partial x^2}=4{f''}_{r^2}x^2+2{f'}_{r^2},\\
&\frac{\partial^2 f(x,y)}{\partial y^2}=4{f''}_{r^2}y^2+2{f'}_{r^2},
\end{aligned}\tag{20}$$

where $f(r^2)$ is a concave and monotone decreasing function with respect to $r^2$, and we have ${f'}_{r^2}<0, {f''}_{r^2}<0$, which are the derivative and second derivative of function $f(\boldsymbol{.})$ with respect to $r^2$. From Eq. (20), we have the following conclusions

$$\begin{aligned}
&\iint_{\Omega_R}\frac{\partial f(x,y)}{\partial x}\cdot\delta(\rho_0-x\cos\theta-y\sin\theta)dxdy=0,\\
&\iint_{\Omega_R}\frac{\partial f(x,y)}{\partial y}\cdot\delta(\rho_0-x\cos\theta-y\sin\theta)dxdy=0,
\end{aligned}\tag{21}$$

and

$$\begin{aligned}
&\iint_{\Omega_R}\frac{\partial^2 f(x,y)}{\partial x^2}\cdot\delta(\rho_0-x\cos\theta-y\sin\theta)dxdy<0,\\
&\iint_{\Omega_R}\frac{\partial^2 f(x,y)}{\partial y^2}\cdot\delta(\rho_0-x\cos\theta-y\sin\theta)dxdy<0,
\end{aligned}\tag{22}$$

where $\rho_0{=}x_0\cos\theta+y_0\sin\theta$. Eq. (21) and Eq. (22) indicates that the integral is only applied along the line $L(\rho_0,\theta): x\cos\theta+y\sin\theta-\rho_0{=}0$, since the delt function not equals to zero only along this line.

Eq. (21) holds because $\frac{\partial f}{\partial x}|_{x=x_0+\delta}=-\frac{\partial f}{\partial x}|_{x=x_0-\delta}$ and $\frac{\partial f}{\partial y}|_{y=y_0+\delta}=-\frac{\partial f}{\partial y}|_{y=y_0-\delta}$ can be concluded from Eq. (20). Eq. (22) holds because $\frac{\partial^2 f(x,y)}{\partial x^2}<0$and $\frac{\partial^2 f(x,y)}{\partial y^2}<0$ can be concluded according to Eq. (20).

Eq. (21) and Eq. (22) can be understood as the integral of the corresponding partial derivative functions along line $x\cos\theta+y\sin\theta-\rho_0{=}0$, where $\rho_0{=}x_0\cos\theta+y_0\sin\theta$. This is a line passing through point $(x_0, y_0)$, and the angle between $x$ axial is $\theta$. $\rho_0$ is the projection point of $(x_0, y_0)$.

From the definition of Radon transform, the Radon transform of the derivative of the original function can be written as

$$\Re_\theta \frac{\partial f}{\partial x} = \iint [\frac{\partial f}{\partial x} \cdot \delta(\rho - x\cos\theta - y\sin\theta)]dxdy$$
$$\Re_\theta \frac{\partial f}{\partial y} = \iint [\frac{\partial f}{\partial y} \cdot \delta(\rho - x\cos\theta - y\sin\theta)]dxdy. \tag{23}$$

Because the projection line does not pass through any speckles caused by global illumination for a certain $\theta$, the values of function $f(x, y)$ (also function $\partial f(x, y)/\partial x$ ) on line $L(\rho_0, \theta)$ : $x\cos\theta + y\sin\theta - \rho_0{=}0$ are zeros, except the partial line in region $\Omega_R$. For the integral inside $\Omega_R$, the integral along a fixed line $\rho_0{=}x\cos\theta + y\sin\theta$ in Eq. (23) equals to zeros, according to Eq. (21). Then, from Eq. (15), we have

$$\Re_\theta \frac{\partial f}{\partial x}|_{(x,y)\in L(\rho_0,\theta)} = \cos\theta \frac{\partial \overset{\vee}{f}(\rho,\theta)}{\partial \rho}|_{\rho=x_0\cos\theta+y_0\sin\theta}{=}0,$$
$$\Re_\theta \frac{\partial f}{\partial y}|_{(x,y)\in L(\rho_0,\theta)} = \sin\theta \frac{\partial \overset{\vee}{f}(\rho,\theta)}{\partial \rho}|_{\rho=x_0\cos\theta+y_0\sin\theta}{=}0. \tag{24}$$

For any value of $\theta$, $\sin\theta$ and $\cos\theta$ are not equal to zero at the same time. Thus, we can conclude that

$$\frac{\partial \overset{\vee}{f}(\rho,\theta)}{\partial \rho}|_{\rho=x_0\cos\theta+y_0\sin\theta}{=}0. \tag{25}$$

From the definition of Radon transform, the second derivative of original function can be written as

$$\Re_\theta \frac{\partial^2 f}{\partial x^2} = \iint [\frac{\partial^2 f}{\partial x^2} \cdot \delta(\rho - x\cos\theta - y\sin\theta)]dxdy$$
$$\Re_\theta \frac{\partial^2 f}{\partial y^2} = \iint [\frac{\partial^2 f}{\partial y^2} \cdot \delta(\rho - x\cos\theta - y\sin\theta)]dxdy. \tag{26}$$

By a similar derivation between Eq. (23) and Eq. (24), and from Eq. (22), Eq. (26) can be written as

$$\begin{aligned}
\Re_\theta \frac{\partial^2 f}{\partial x^2}|_{(x,y)\in L(\rho_0,\theta)} &= \cos^2\theta \frac{\partial^2 \overset{\vee}{f}(\rho,\theta)}{\partial \rho^2}|_{\rho=x_0\cos\theta+y_0\sin\theta} \\
&= \iint_{\Omega_R} [\frac{\partial^2 f}{\partial x^2} \cdot \delta(\rho_0 - x\cos\theta - y\sin\theta)]dxdy < 0 \\
\Re_\theta \frac{\partial^2 f}{\partial y^2}|_{(x,y)\in L(\rho_0,\theta)} &= \sin^2\theta \frac{\partial^2 \overset{\vee}{f}(\rho,\theta)}{\partial \rho^2}|_{\rho=x_0\cos\theta+y_0\sin\theta} \\
&= \iint_{\Omega_R} \frac{\partial^2 f(x,y)}{\partial y^2} \cdot \delta(\rho_0 - x\cos\theta - y\sin\theta)dxdy < 0.
\end{aligned} \tag{27}$$

Because $\sin^2\theta \geq 0$and $\cos^2\theta \geq 0$, thus, we can conclude that

$$\frac{\partial^2 \overset{\vee}{f}(\rho,\theta)}{\partial \rho^2}\Big|_{\rho=x_0\cos\theta+y_0\sin\theta} < 0. \tag{28}$$

Thus, from Eq. (25) and (28), the projection point $\rho_0$ is a local maximum on the projection function $\overset{\vee}{f}(\rho,\theta)$. □

# B Perfect Reconstruction Property of Local Slice Extension Method

In the main text, the projection functions are captured by the proposed local slice extension method, which is implemented by a "coarse to fine" localization procedure. In this section, we will introduce local slice extension method from a theoretical aspect, and prove that the local slice extension method can perfectly reconstruct the projection function. First, we provide Lemma 3 which states that the reconstructed projection function by inverse discrete Fourier transform (IDFT) corresponds to a periodic extension version of the projection function $f_\theta(\rho; u, v)$ when the patterns generated by Eq. (9) are projected.

**Lemma 3.** *Assume $f_\theta(\rho; u, v)$ is the projection function with direction $\theta$ for camera pixel $(u, v)$, by projecting patterns in the form of Eq. (9), the reconstructed function of camera pixel $(u, v)$ by IDFT becomes a periodic extension version of the original projection function*

$$\tilde{f}^r_\theta(\rho_r; u, v) = \sum_{r_1=-\infty}^{+\infty} f(\rho_r - r_1 M_\theta; u, v), \tag{29}$$

*where $\rho_r$ is a pixel on the reconstructed function, $M_\theta$ is given in Eq. (9), which is the size of the maximum of $\theta$ projected reception field for each camera pixel, and $r_1$ is integer.*

Proof. Similar to Eqs. (7) and (8), when each sample in the frequency domain is obtained by using the patterns generated by Eq. (9), the reconstructed function by applying IDFT on the captured intensity is calculated as

$$\begin{aligned}
\tilde{f}^r_\theta(\rho_r; u, v) &= IDFT\{\frac{Sb}{2} \cdot \sum_{v'=0}^{N-1}\sum_{u'=0}^{M-1} \cdot h(u', v'; u, v) \cdot \exp[-\frac{2\pi k}{M_\theta}(u'\cos\theta + v'\sin\theta)]\} \\
&= \frac{Sb}{2} \cdot \sum_{r_1=-\infty}^{+\infty}\sum_{v'=0}^{N-1}\sum_{u'=0}^{M-1} \cdot h(u', v'; u, v) \cdot \delta(\rho_r - u'\cos\theta - v'\sin\theta - r_1 M_\theta) \\
&= \frac{Sb}{2} \cdot \sum_{r_1=-\infty}^{+\infty} f_\theta^{Radon}(\rho_r - r_1 M_\theta; u, v) \\
&= \sum_{r_1=-\infty}^{+\infty} f_\theta(\rho_r - r_1 M_\theta; u, v),
\end{aligned} \tag{30}$$

where $f_\theta^{Radon}(\rho; u, v)$ is the discrete Radon transform of LTCs along direction $\theta$, as defined in Eq. (2). □

Figure 7 shows the periodic extension version of projection function. Provided with Lemma 3, we can prove the local slice extension theorem.

**Theorem 2.** *Local Slice Extension Theorem. If the size of the maximum of $\theta$ projected reception field $M_\theta$ covers the non-zero regions of $f_\theta(\rho; u, v)$, projection function can be perfectly reconstructed by adopting local slice extension method, that is, the projection function obtained by the local slice extension method implemented by the "coarse to fine" procedure is exactly equal to the projection function captured and reconstructed according to Eq. (4) - Eq. (8).*

Proof. If the size of the maximum of $\theta$ projected reception field $M_\theta$ covers the non-zero regions of $f_\theta(\rho; u, v)$, then from Lemma 3, the reconstructed projection function can be regarded as a periodic extension version of the original projection function, with step size of $M_\theta$. Aliasing does not occur in this situation (Fig. 7), and all information of $f_\theta(\rho; u, v)$ is preserved in $\tilde{f}_\theta^r(\rho; u, v)$. Thus, by adopting local slice extension method, projection function can be exactly reconstructed, which equals to projection function captured and reconstructed according to Eq. (4) - Eq. (8).

Provided that the nonzero region $C_\theta(\rho; u, v)$ of $f_\theta(\rho; u, v)$ is known, the projection function $f_\theta^r(\rho; u, v)$ can be exactly obtained by taking the values inside the visible region $C_\theta(\rho; u, v)$ of $\tilde{f}_\theta^r(\rho; u, v)$ and setting zeros outside $C_\theta(\rho; u, v)$ of $\tilde{f}_\theta^r(\rho; u, v)$ as given by Eq. (12). $M_\theta$ is obtained by coarse localization step, which can avoid the occurence of alizaing as shown in the second row of Fig. 7.□

# C Localization by Truncation

## C.1 Derivation of the Coarse Localization Accuracy in Terms of Frequency Number

Suppose the projection function $f(\rho)$ has a length of $L$. The process of truncation the higher frequencies of the projection function $f(\rho)$ can be described as applying a window function $P(k)$ on $F(k)$, which is the discrete Fourier transform (DFT) of $f(\rho)$

$$F^C(k) = P(k)F(k), \tag{31}$$

where $F^C(k)$ is the captured frequencies for coarse localization, and $P(k)$ has following form

$$P(k)=\begin{cases} 1 & 0 \le k \le K-1 \text{ or } L-K-1 \le k \le L-1 \\ 0 & \text{otherwise} \end{cases}, \tag{32}$$

where $K$ is the captured number of low frequencies. Conjugate symmetry property is considered, which resulting the two banded nonzero values in Eq. (32).

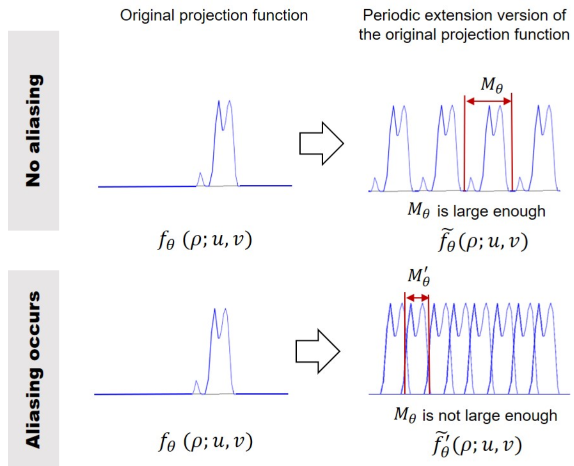



**Fig. 7.** Perfect reconstruction property of local slice extension method. Periodic extension version of projection function is shown. In the first row, $M_\theta$ is large enough such that aliasing does not occur. While in the second row, $M_\theta$ is not large enough and aliasing occurs.

From circular convolution property, Eq. (31) is equivalent to applying circular convolution in the spatial domain

$$F^C(k) = P(k)F(k) \overset{DFT}{\leftrightarrow} f^C(\rho) = p(\rho) \otimes_L f(\rho), \tag{33}$$

where $\otimes_L$ denotes L-point circular convolution. $\overset{DFT}{\leftrightarrow}$ denotes that the two functions are a DFT pair. If $p(\rho)$ is a periodic impulse train with period L, $f^C(\rho)$ equals exactly to $f(\rho)$. This corresponds to capture the whole frequencies $k = 0, \cdots, L-1$.

When truncation is applied, the reconstructed coarse projection function $f^C(\rho)$ is the L-point circular convolution of $p(\rho)$ and $f(\rho)$, shown in Fig. 8 (a)–(c). The approximated size of the projected reception field $M_s$is determined by the width of the main lobe of $p(\rho)$, which is depicted in Fig. 8 (b). The length of $M_s$ can be approximated by adding the actual size of the projected reception field and the size of the main lobe.

$p(\rho)$ can be obtained by applying IDFT to $P(k)$, as expressed by

$$
\begin{aligned}
p(\rho) &= \frac{1}{L}\sum_{k=0}^{K-1}\exp(j\frac{2\pi}{L}k\rho) \\
&+ \frac{1}{L}\sum_{k=L-K-1}^{L-1}\exp(j\frac{2\pi}{L}k\rho) \\
&= \frac{1}{L}\sum_{k=-(K-1)}^{K-1}\exp(j\frac{2\pi}{L}k\rho) \\
&= \frac{1}{L}\frac{\sin[\frac{2\pi}{L}(K-\frac{1}{2})\rho]}{\sin(\frac{\pi}{L}\rho)}.
\end{aligned}
\tag{34}
$$

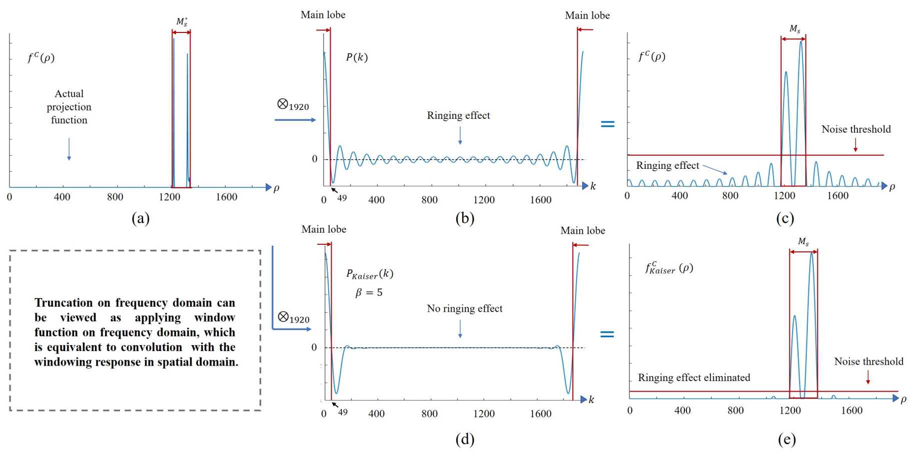



**Fig. 8.** Truncation in frequency domain. (a) Actual projection function. The projected reception field is indicated by the red lines. (b) Convolution kernel in spatial domain when direct truncation is applied in frequency domain. The size of the main lobe is important for localization accuracy. Ringing effect is obvious if the high frequency is removed directly. (c) The reconstructed coarse projection function when direct truncation is applied. The location of the projected reception field can be determined by setting the region where the values are larger than a noise threshold. (d) Convolution kernel in spatial domain when Kaiser window, with the shape parameter set as 5, is applied in frequency domain. Ringing effect is not obvious. (e) The reconstructed coarse projection function when Kaiser window is applied in frequency domain.

The size of the main lobe can be determined by setting Eq. (34) equals to zero. The first zero-crossing point $\rho = L/(2K-1)$ is half of the size of the main lobe. For a fixed $L$, with more low frequency information captured, the more concentrated the main lobe area, thus the more accurate the coarse localization process.

In summary, for a projection function $f(\rho)$ with a length of $L$, when $K$ low frequencies are captured, The approximated size of the projected reception field $M_s$ is $2L/(2K-1)$ wider than the actual size.

There is a trade-off between the number of low frequencies captured and the localization accuracy. To solve this problem, suppose $M_s^*$ be the actual size of the projected reception field, and

$$R = \frac{2L}{(2K-1)M_s^*}, \tag{35}$$

be the localization uncertain ratio. In our experiments, we believe $R$=1 , which indicates that the approximated size of the projected reception field $M_s$ is only one time wider than the actual size, is enough. $L$ takes value of 1920, and $M_s^*$ is assumed as 150 pixels. Putting these values into Eq. (35), we get $K \approx 13$. We found that the number of low frequency samples of 10 is enough in most situations.

However, removing high frequency information directly results in severely ringing effect, shown in Fig. 8 (b)–(c). Alternatively, Kaiser window [2], with the shape parameter $\beta$ set to 5, is applied to the sampled low frequency samples. The ringing effect can be largely eliminated, as shown in Fig. 8 (d)–(e). Kaiser window is originally applied on spatial/time domain, however, in the context of pPSI, Kaiser window should by applied on frequency domain.

### C.2 Partial Scan for Fine Localization Step

This section provides experimental results to show the excellent compressive property of Fourier spectrum, which provides opportunity to obtain satisfactory results by using patrial observations. Although in section B we prove that perfect reconstruction can be achieved by projecting patterns generated by Eq. (9) with all frequencies $k_\theta$=$0, 1 \dots M_\theta - 1$, we can only project partial frequency to obtain satisfactory results in most real applications. In fine localization step, only the patterns with low frequencies are projected (We used a scan ratio of 25% in the main text). The unscanned frequencies are filled with zeros. We conducted an experiment to analyze the influence of the scan ratio on the subpixel matching error (SME), and explain the reason for using scan ratio of 25%. SME is defined as followed.

Performing both PSI and pPSI for a scene. For a camera pixel $(u, v)$, suppose $\boldsymbol{u'}_R = (u'_R, v'_R)$ is the subpixel matched point found by PSI, and $\boldsymbol{u'_P} = (u'_P, v'_P)$ is the subpixel matched point found by pPSI. We will conduct pPSI with various scan ratio. The SME is defined as

$$SME(\boldsymbol{u'_R}, \boldsymbol{u'_P}) = \frac{1}{2}[(u'_R - u'_P)^2 + (v'_R - v'_P)^2]. \tag{36}$$

In this section, we study the influence of scan ratio on SME for the composite scene (the scene displayed in Fig.3) in the main text, the result is shown Fig. 9(a). The vertical axis corresponds to the averaged SME vale. We averaged the SME value of every point in the composite scene. The number of points participating in the average calculation was 289,145. The horizontal axis corresponds to the scan ratio. The calculated scan ratios were 5%, 8%, 10%, 12%, 14%,16%, 18%, 20%, 22% 25%,30%, 35%, 40%, 50%, 60%, 70%, 80%, and 90%, 100%, and each ratio corresponds to one data point on the curve. As can be seen in the figure, the tenth data point or the 25% scan ratio point (represented by a red star) provides the best trade-off between accuracy and efficiency, since the gradient of the curve changes greatly near this point. The curve drops rapidly before the point, and descents slowly after the point. The averaged SME of 25% scan is 0.047 pixel, which is not very worse than the SME value 0.042 of 100% scan. Thus, we believe that using the scan ratio of 25% provides satisfactory results.

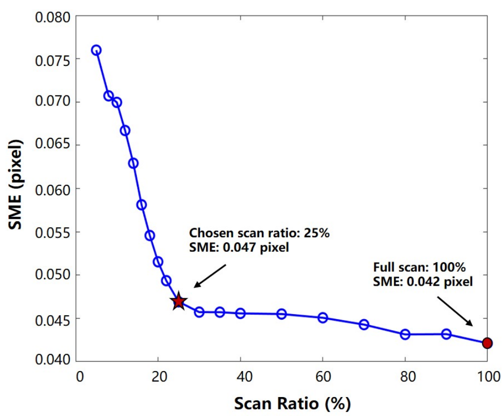
To answer why 25% scan ratio is possible to provide satisfactory results, we show the normalized energy distribution in Fig. 9(b). This energy distribution is calculated by the $\theta = 0°$ projection function. We captured the 100% Fourier spectrum by pPSI, and calculated the norm of each frequency position in the Fourier spectrum (The value on each frequency position is a complex value). The norm corresponds to energy for the frequency index. We repeated the calculation for every camera pixel, and summed the energy according to its frequency. Taking account of the conjugate symmetric property of Fourier spectrum of real-valued signal, we show half of the normalized energy distribution in Fig. 9(b). As can be seen, the energy of the first 25% frequencies account 88.24% of the total energy. This explains the reason why 25% scan ratio is possible to provide satisfactory results. For the scenes in this study, 25% scan corresponds to 19 frequencies in the fine localization step. In addition, the patterns with zero frequency has been captured in the coarse localization step. Thus, the number of frequencies required by fine localization step is 18.

## D Choosing the Number of Slices

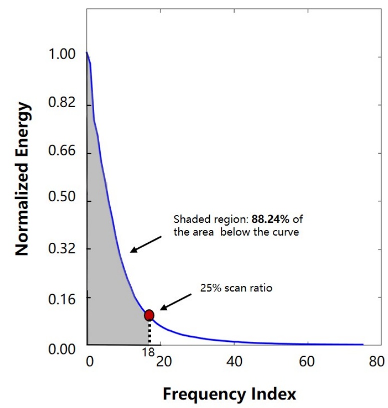
In this section, we explain why slices with four directions is used in this study. The fundamental reason that we use four directions is that we want to keep the simplicity of the calculation of candidate matching points, and also ensure the robustness.

As an illustration, a pixel transport image is shown in Fig. 10(a), with its four projection functions along angles $\theta = 0°, 45°, 90°, 135°$. In this example, the 45° projection line that passes through the direct speckle also passes through another speckle, as shown in Fig. 10(a). The local maximum constraint proposition is violated in this situation, which is the fundamental reason for using four slices if robust correspondence points are required.

When only one slice is used, epipolar constraint has to be applied for calculating the candidate points. However, these candidate points cannot be eliminated further to obtain the final correspondence matching point, if no more constraint is provided, as shown in Fig. 10(b).




(a)




(b)

**Fig. 9.** Analysis of partial scan in fine localization step. (a) The influence of scan ratio on SME for the composite scene. (b) Normalized energy distribution with respect to frequency.

When two slices are used, more virtual points are generated, as shown in Fig. 10(c). Some of these virtual points may near to the epipolar line, and still cannot be eliminated.

Three slices can handle most situations. However, there may exist situations where the local maximum constraint proposition is violated. The 45° projection line that passes through the direct speckle also passes through another speckle, thus only 2 peaks, instead of 3 peaks, can be found on the 45° projection function, shown in Fig. 10(a). We refer to this situation as degenerate, and refer to the peak generated by a projection line that passes more than one speckles as degenerate peak. The position of the degenerate peak is influenced by several speckles, and thus its position is not exactly equal to the projected position of the maximum point in any speckle. In this situation, if we back project the degenerate peak, the resulted line is not intersected exactly with other projection lines, and large error can be incurred, refer to Fig. 10(a) and (d).

Using four slices can better deal with the degenerate situation. Although the 45° projection line leads to degenerate situation, the 135° projection line does not. We can calculate the direct illumination point by projection functions with $\theta = 0°, 90°, 135°$.To be more general, if we find a combination of peaks in four slices does not determine a point, we can check whether by eliminating one peak, a point can be determined. If the remaining three peaks does intersect as one point, we store the intersected point as candidate, and check whether it passes epipolar constraint.

When four slices are used, the angles of $\theta = 0°, 45°, 90°, 135°$ are chosen because they evenly divide 180°.

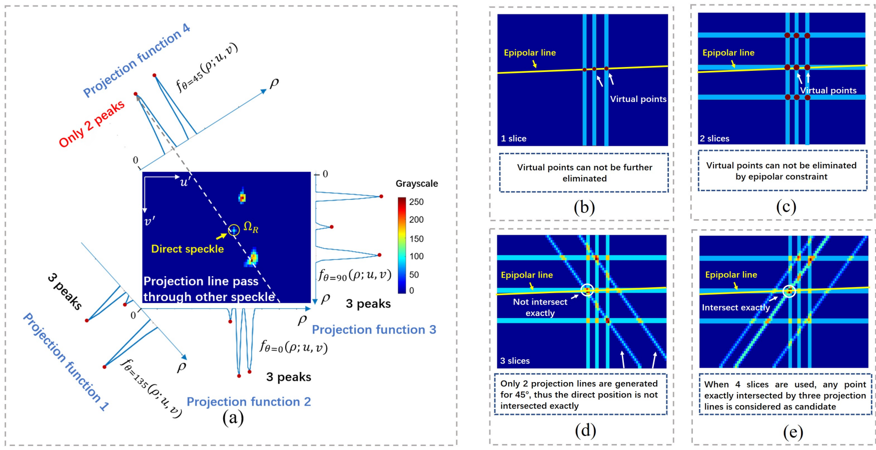



**Fig. 10.** pPSI with different number of slices. (a) A PTI where the 45° projection line passes through another speckle. Only 2 peaks can be observed on the 45° projection function. (b) pPSI by one slice. Epipolar constraint has to be applied for calculating the candidate point. However, virtual points cannot be further eliminated. (c) pPSI by two slices. More virtual points are incurred. The virtual points near to the epipolar line cannot be eliminated. (d) pPSI by three slices. If one of the projection lines which passes through the direct speckle also passes through other speckles, the local maximum constraint proposition is violated. The direct illumination point may not intersect exactly. (e) pPSI by four slices. Any point exactly intersected by three projection lines is considered as candidate points.

# E Calculation of the Number of Patterns

This section introduces the calculation of the number of patterns required by pPSI. For a slice with certain direction, suppose that $N_c$ is the number of frequencies for coarse location, and $N_f$ is the size of the projection reception field in the fine location step. $\eta$ is the scan ratio in fine localization step. $S$ is the step number. Then, taking account of the conjugate symmetric property of Fourier spectrum of real-valued signal, the required number of patterns $N_t$ is calculated by

$$N_t = \begin{cases} SN_c + \eta \frac{S}{2} N_f - S & S \text{ is even} \\ SN_c + \eta \frac{S}{2}(N_f+1) - S & S \text{ is odd, } N_f \text{ is odd} \\ SN_c + \eta S(\frac{N_f}{2}+1) - S & S \text{ is odd, } N_f \text{ is even.} \end{cases} \quad (37)$$

The minus $S$ means that the zero frequency patterns are projected in the coarse localization step, which should be eliminated from the fine localization step.

## F Compare with Micro-phase Shifting Method

In this section, several additional experiments are made to compare pPSI with micro-phase shifting method [1]. Scenes with strong interreflections and subsurface scattering are investigated.

For interreflections, three scenes are investigated. 1. The scene of gypsum bear with a mirror nearby; 2. The V-groove scene in the main text; 3. A workpiece made by aluminum alloy, shown in Fig. 11(a). In Fig. 11(b), we showed the overlapped patterns on the surfaces when high-frequency patterns projected. As can be seen, glossy interreflections dominated in this scene.

For subsurface scattering, we also tested three scenes. 1. The jade horse; 2. The polyamide sphere; 3. The candle appeared in the main text, also shown in Fig. 11 (c). In Fig. 11 (d), we showed the degraded patterns when low frequency patterns projected. As can be seen, strong subsurface scattering occurred in this scene.

We conduced micro-phase shifting method and pPSI to the above scenes, and the 3D reconstruction results are shown in Fig. 12. For the three interreflections scenes, the projected patterns for micro-phase shifting were chosen as the 16-15 patterns set, which contains patterns for a frequency-band around 16 pixels, and 15 frequencies. For the three subsurface scattering scenes, the projected patterns for micro-phase shifting were chosen as the 64-15 patterns set, which contains patterns for a frequency-band around 64 pixels, and 15 frequencies. For pPSI, all the scenes are illuminated by the same pattern set, i.e., the patterns used in section 5.2 and 5.3 in the main manuscript.

As can be seen from the results, micro-phase shifting method failed to reconstruct complete 3D data for the following scenes: the scene of gypsum bear with a mirror, the aluminum alloy workpiece scene. pPSI achieves complete 3D reconstruction for these scenes, and the quality of the 3D data seems good. Although micro-phase shifting succeeded in reconstructing complete 3D data for the V-groove scene and the three subsurface scattering scenes, the quality of the 3D data is not as good as that of pPSI. There are obvious ripple errors in the point clouds reconstructed by micro-phase shifting method.

To test the quality of the 3D point cloud data reconstructed by micro-phase shifting and pPSI, we calculated the root-mean-square (RMS) error for the V-groove scene and the candle scene. We separately fitted planes for each surface of the V-groove scene and the candle scene such that the RMS error can be obtained by calculating the average error between the reconstructed 3D data points and the fitted planes. The RMS errors are shown for these two scenes in Fig. 13. For micro-phase shifting method, the RMS value of the upper plane and lower plane of the V-groove was 0.058 (mm) and 0.056 (mm), respectively. For pPSI, the RMS value of the upper plane and lower plane of the V-groove was 0.021 (mm) and 0.015 (mm), respectively. The accuracy of pPSI can is better than micro-phase shifting. Similar conclusions can be drawn for the candle scene. For micro-phase shifting method, the RMS value of the candle was 0.228 (mm). For pPSI, the RMS value of the candle was 0.073 (mm).

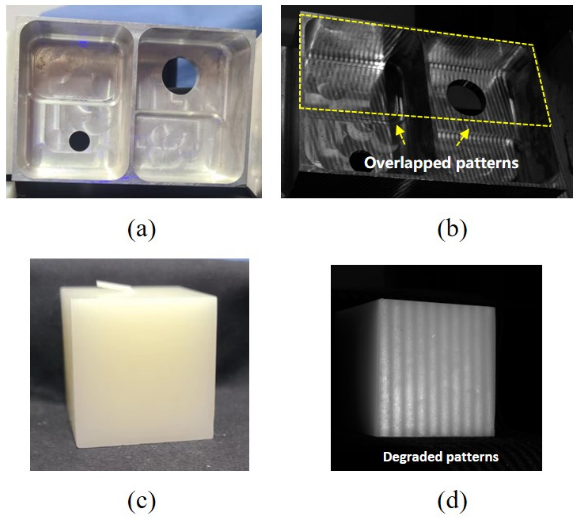



**Fig. 11.** The workpiece made by aluminum alloy and the candle scene. (a) Workpiece made by aluminum alloy. (b) The overlapped patterns occurred in the workpiece. (c) Candle. (d) The degraded patterns on the candle.

The reason that explains the better ability of pPSI is that pPSI explicitly separates the influences of global illumination and direct illumination, this property enables pPSI to reconstruct 3D data under more complex global illumination. Also, pPSI projects both lower and higher frequency patterns, and all the response information is synthesized by Fourier transform. There is no need for pattern set selection. Thus, pPSI is able to reconstruct interreflections and strong subsurface scattering simultaneously (with a same pattern set to overcome both interreflections and subsurface scattering).

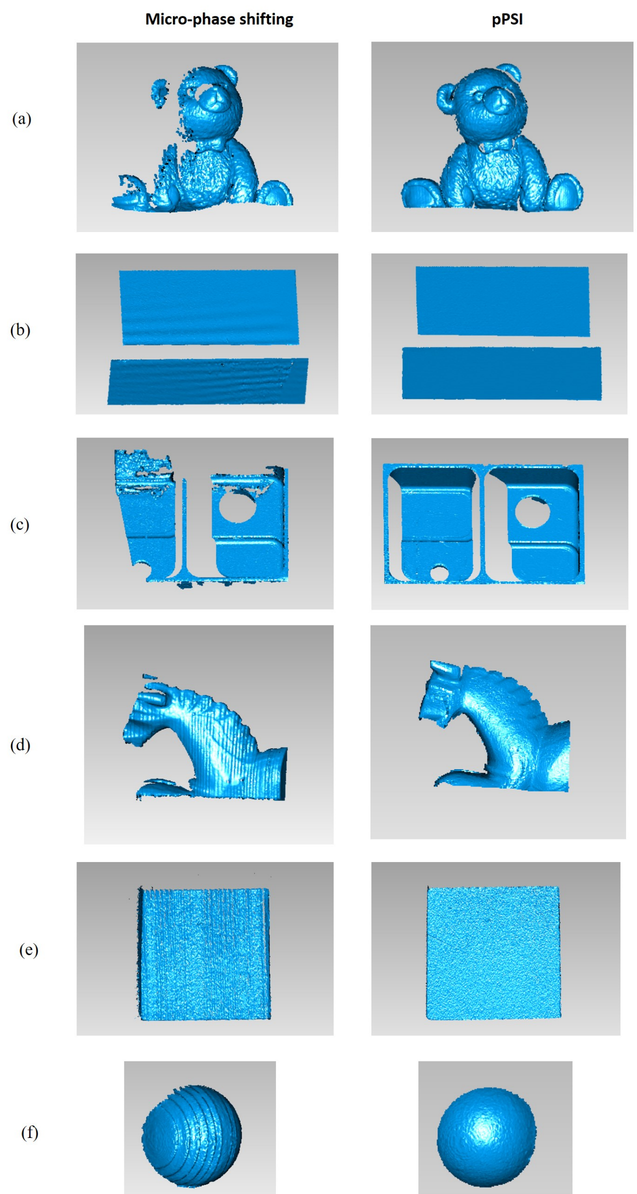



**Fig. 12.** The reconstructed 3D point clouds. (a) The 3D data of gypsum bear with a mirror nearby. (b) The 3D data of V-groove. (c) The 3D data of the workpiece. (d) The 3D data of the jade horse. (e) The 3D data of the candle. (f) The 3D data of the polyamide sphere.

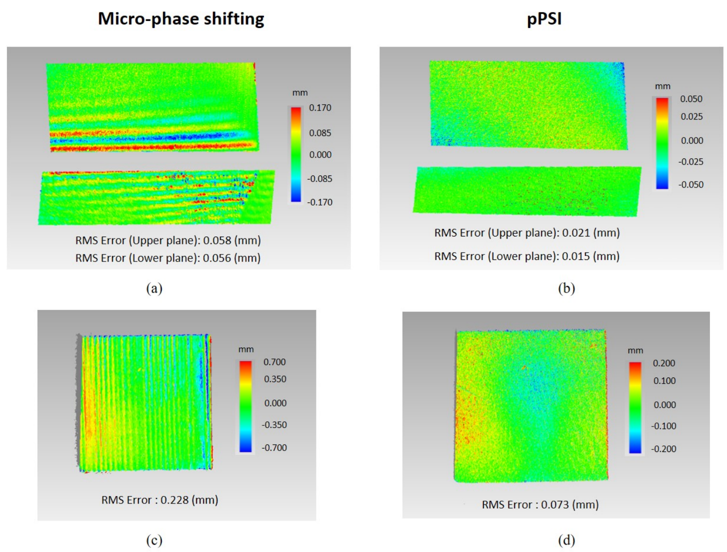



**Fig. 13.** Accuracy comparision between micro-phase shifting and pPSI. (a) The RMS error for the V-groove scene reconstructed by micro-phase shifting method. (b) The RMS error for the V-groove scene reconstructed by pPSI. (c) The RMS error for the candle scene reconstructed by micro-phase shifting method. (d) The RMS error for the candle scene reconstructed by pPSI.

## References

1. Gupta, M., Nayar, S.K.: Micro phase shifting. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. pp. 813–820. IEEE (2012)
2. Kaiser, J.F.: Nonrecursive digital filter design using the i_0-sinh window function. In: Proc. 1974 IEEE International Symposium on Circuits & Systems, San Francisco DA, April. pp. 20–23 (1974)