Supplementary Materials for

High-fidelity CUDA-based SIM parallel reconstruction method

Hongyu Wang,[a] Ruijie Cao,[a] Wenyi Wang,[b] Yaning Li,[c] Peng Xi[a,b,*]
Peking University, College of Future Technology, Beijing 100871, China
[b]Airy Technologies Co., Ltd., Beijing 100871, China
[c]China Academy of Space Technology, Beijng Institute of Space Mechanics and electricity, 100094, China

S1. GPU kernel structure

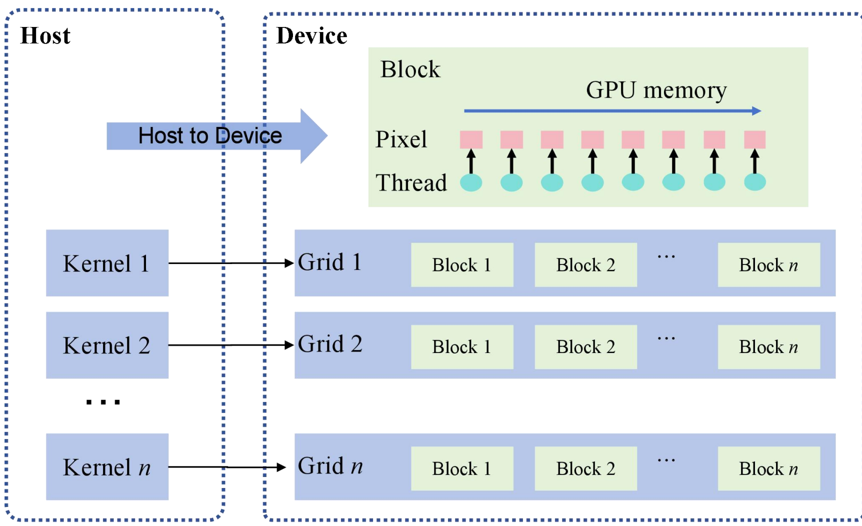

Fig. S1 GPU kernel structure.

The GPU kernel structure is shown in Fig.S1. It is defined as different threads, blocks, grids, and kernels with pixels as the basic units. In the proposed parallel reconstruction architecture, multi-layer images form a spatial three-dimensional matrix data. A kernel computation can be regarded as one basic GPU operation. This speed depends more on the performance of the GPU, including graphics memory, frequency, and so on. The complexity of these computations is the main reason for the time consuming of 3DSIM reconstruction. In the GPU space, the maximum space of a single high-performance mainstream graphics card is 24GB. Therefore, it is necessary to make the algorithm execution process and variable data structure lightweight. Due to the large amount of spatial transformation of three-dimensional data and the mixed computation of multi-layer data in the 3DSIM reconstruction process, we employ matrix and parallelization methods to enhance the computing speed.

S2. 3DSIM structured illumination principle

3DSIM can achieve twice the resolution enhancement in three-dimensional space by filling OTF through three-dimensional spectrum transfer and concatenation. When three coherent light beams intersect and interfere on the sample, the fluorescence distribution at the illumination angle θ and phase φ can be expressed as:

$$C_{\theta,\varphi}(x,y,z)=\left[S(x,y,z)\cdot I_{\theta,\varphi}(x,y,z)\right]\otimes PSF(x,y,z) \tag{S1}$$

Where $I_{\theta,\varphi}(x,y,z)$ is the three-dimensional illumination structure light generated by the interference of three beams of light, $PSF(x,y,z)$ is the three-dimensional point spread function of the system, and its Fourier transform is $OTF(k_x,k_y,k_z)$. Unlike the 2DSIM illumination structured light that only contains 0th and ±1st harmonics, the illumination structured light of 3DSIM contains 0th, ±1st, and ±2nd harmonics, which can be expressed in spatial domain as:

$$\begin{aligned} I_{\theta,\varphi}(x,y,z)=I_0\big[&1+2\mathrm{m}^2+4m\cos(2\pi p_z z)\cos(2\pi p_x x+2\pi p_y y+\varphi) \\ &+2\mathrm{m}^2\cos(4\pi p_x x+4\pi p_y y+2\varphi)\big] \end{aligned} \tag{S2}$$

Where m is the modulation depth, p_x, p_x and p_z are the periodic components of the illumination structured light in the x, y, and z directions respectively. It can be seen that only the 1st harmonic $4m\cos(2\pi p_z z)$ is related to the z direction, while the other harmonics are only related to the x and y directions.

Here we simplify the variables m_0, m_1, and m_2 in equation (2):

$$\begin{aligned} m_0 &= 1+2m^2 \\ m_1 &= 4m \\ m_{1z}(z) &= 4m\cos(2\pi p_z z) \\ m_2 &= 2m^2 \end{aligned} \tag{S3}$$

Therefore, when the illumination pattern sequence $I_{\theta,\varphi}(x,y,z)$ excites the sample, the fluorescence image $C_{\theta,\varphi}(x,y,z)$ obtained by the camera in the frequency domain $C_{\theta,\varphi}(k)$ is represented as:

$$C_{\theta,\varphi}(k)=\left[S(k)\otimes I_{\theta,\varphi}(k)\right]OTF(k) \tag{S4}$$

Then the fluorescence image obtained by the camera in the frequency domain can be further expressed as:

$$C_{\theta,\varphi}(k)=I_0\begin{bmatrix} m_0 S(k) \\ +m_1 S(k_{x,y}-p_{x,y},k_z-p_z)e^{j\varphi} \\ +m_1 S(k_{x,y}-p_{x,y},k_z+p_z)e^{j\varphi} \\ +m_1 S(k_{x,y}+p_{x,y},k_z-p_z)e^{-j\varphi} \\ +m_1 S(k_{x,y}+p_{x,y},k_z+p_z)e^{-j\varphi} \\ +m_2 S(k_{x,y}-2p_{x,y},k_z)e^{j2\varphi} \\ +m_2 S(k_{x,y}+2p_{x,y},k_z)e^{-j2\varphi} \end{bmatrix}OTF(k) \tag{S5}$$

In each illumination direction, five phase (φ_1=0, φ_2=2π/5, φ_3=4π/5, φ_4=6π/5, φ_5=8π/5)can be set using a five step phase-shift matrix separation method to solve for the 0 level, ±1 level, and ±2 level spectral components:

$$\begin{aligned} S_0(k)&=S(k) \\ S_{-1}(k)&=S(k_{x,y}-p_{x,y},k_z-p_z)+S(k_{x,y}-p_{x,y},k_z+p_z) \\ S_{+1}(k)&=S(k_{x,y}+p_{x,y},k_z-p_z)+S(k_{x,y}+p_{x,y},k_z+p_z) \\ S_{-2}(k)&=S(k_{x,y}-2p_{x,y},k_z) \\ S_{+2}(k)&=S(k_{x,y}+2p_{x,y},k_z) \end{aligned} \tag{S6}$$

Therefore, equation (5) can be rewritten as:

$$\begin{bmatrix} C_{\theta,\varphi_1}(k) \\ C_{\theta,\varphi_2}(k) \\ C_{\theta,\varphi_3}(k) \\ C_{\theta,\varphi_4}(k) \\ C_{\theta,\varphi_5}(k) \end{bmatrix}=I_0 M\begin{bmatrix} S_0(k) \\ S_{-1}(k) \\ S_{+1}(k) \\ S_{-2}(k) \\ S_{+2}(k) \end{bmatrix}OTF(k) \tag{S7}$$

Where M is seperation matrix expressed as:

$$M=\begin{bmatrix} m_0 & m_1e^{j\varphi_1} & m_1e^{-j\varphi_1} & m_1e^{j2\varphi_1} & m_1e^{-j2\varphi_1} \\ m_0 & m_1e^{j\varphi_2} & m_1e^{-j\varphi_2} & m_1e^{j2\varphi_2} & m_1e^{-j2\varphi_2} \\ m_0 & m_1e^{j\varphi_3} & m_1e^{-j\varphi_3} & m_1e^{j2\varphi_3} & m_1e^{-j2\varphi_3} \\ m_0 & m_1e^{j\varphi_4} & m_1e^{-j\varphi_4} & m_1e^{j2\varphi_4} & m_1e^{-j2\varphi_4} \\ m_0 & m_1e^{j\varphi_5} & m_1e^{-j\varphi_5} & m_1e^{j2\varphi_5} & m_1e^{-j2\varphi_5} \end{bmatrix} \tag{S8}$$

Based on separation matrix, five frequency components can be given by:

$$\begin{bmatrix} C_0(k) \\ C_{-1}(k) \\ C_{+1}(k) \\ C_{-2}(k) \\ C_{+2}(k) \end{bmatrix} = \begin{bmatrix} S_0(k)\cdot OTF(k) \\ S_{-1}(k)\cdot OTF(k) \\ S_{+1}(k)\cdot OTF(k) \\ S_{-2}(k)\cdot OTF(k) \\ S_{+2}(k)\cdot OTF(k) \end{bmatrix} = \frac{1}{I_0}M^{-1}\begin{bmatrix} C_{\theta,\varphi_1}(k) \\ C_{\theta,\varphi_2}(k) \\ C_{\theta,\varphi_3}(k) \\ C_{\theta,\varphi_4}(k) \\ C_{\theta,\varphi_5}(k) \end{bmatrix} \tag{S9}$$

S3. Automatic determination of illumination pattern parameters

In order to accurately determine the parameters of the structured illumination pattern, we use COR methods to estimate the frequency, angle, phase, and modulation depth of the structured illumination pattern during the reconstruction process. Because only estimation methods based on COR can solve frequency vectors with sub-pixel accuracy. However, the iterative nature of cross-correlation inevitably leads to longer computation times.

Under the GPU parallel architecture, we utilize both +1 and +2 frequency components simultaneously. If the result of the +2 order frequency peak is reliable, we first choose to use the +2 order frequency peak estimation result. On the contrary, we choose the result estimated by the +1 order spectral frequency vector. This can ensure the effectiveness of frequency vector estimation and improve the efficiency of parameter estimation. After obtaining the frequency vector estimation results, the corresponding parameters such as phase and angle can be analyzed. Generally, the peak value of+second frequency is more accurate than+first frequency. However, in low SNR samples, the peak value of +first frequency has a higher contrast than +first frequency. Therefore, this method can improve the accuracy of parameter estimation for low SNR images.

As shown in equation 8, C_0, C_1, and C_2 can be expressed as follows:

$$\begin{aligned} &C_0 = OTF(k)S(k) \\ &C_1 = OTF(k)m_1\left[S\left(k_{x,y}+p_{x,y},k_z-p_z\right)+S\left(k_{x,y}+p_{x,y},k_z+p_z\right)\right]e^{-j\varphi} \\ &C_2 = OTF(k)m_2S\left(k_{x,y}+2p_{x,y},k_z\right)e^{-j2\varphi} \end{aligned} \tag{S10}$$

where, C_0 represents wide-field information, C_1 and C_2 represents super-resolution signals modulated into OTF. The purpose of super-resolution is to demodulate high-frequency information in C_1 and C_2. For this, it is necessary to estimate reliable $p_{x,y}$, φ, and m, based on which the 0, ±1 order and ±2 order spectra can be accurately separated. If we remove the OTF terms in C_0, C_1 and C_2, then they can be represented as:

If OTF terms in C_0, C_1 and C_2 is removed and shifted by $p'_{x,y}$, they can be expressed as:

$$\begin{aligned} &C'_0 = C_0\frac{OTF^*(k)}{|OTF(k)|^2} = S(k) \\ &C'_1 = C_1\frac{OTF^*\left(k-p'_{x,y}\right)}{\left|OTF\left(k-p'_{x,y}\right)\right|^2} = m_1S\left(k+p_{x,y}-p'_{x,y}\right)e^{-j\varphi} \\ &C'_2 = C_1\frac{OTF^*\left(k-2p'_{x,y}\right)}{\left|OTF\left(k-2p'_{x,y}\right)\right|^2} = m_2S\left(k+2p_{x,y}-2p'_{x,y}\right)e^{-j2\varphi} \end{aligned} \tag{S11}$$

where * represents the result of complex conjugation, and OTF can be obtained through calibration or simulation. Therefore, C'_0 and C'_1, or C'_0 and C'_2 have a high correlation. When C'_1 is shifted by $p_{x,y}$ or C'_2 is shifted by $2p_{x,y}$, the correlation reaches its maximum value. Based on the above characteristics, we can globally move C'_1 or C'_2 by sub-pixel $p'_{x,y}$ and gradually find the maximum cross-correlation value between C'_0 and C'_1, or C'_0 and C'_2 :

$$\begin{aligned} &CC_1\left(p'_{x,y}\right) = \frac{\sum_k m_1S^*(k)S\left(k+p_{x,y}-p'_{x,y}\right)e^{-j\varphi}}{\sum_k |S(k)|^2} \\ &CC_2\left(p'_{x,y}\right) = \frac{\sum_k m_2S^*(k)S\left(k+2p_{x,y}-2p'_{x,y}\right)e^{-j2\varphi}}{\sum_k |S(k)|^2} \end{aligned} \tag{S12}$$

The COR coefficient determines the similarity between the two results after translation. Therefore, the cross-correlation coefficient can be used to characterize the accuracy of parameter estimation. Therefore, parameter estimation results $p_{x,y}$ can be obtained based on the +1 and +2 frequency components separately, shown as following:

$$\begin{aligned} p_{x,y}(CC_1) &= \underset{p'_{x,y}\in p}{\arg\max}\, CC_1(p'_{x,y}) \\ p_{x,y}(CC_2) &= \underset{p'_{x,y}\in p}{\arg\max}\, CC_2(p'_{x,y}) \end{aligned} \tag{S13}$$

In order to obtain isotropic results, it is necessary to calculate the data of three directional fringe with an angle difference of 120 degrees, so the parameters of the three directions $p_{1x,y}$, $p_{2x,y}$ and $p_{3x,y}$ need to be estimated separately. In our method, parallel GPU architecture is used to calculate the parameter estimation results of +2 and +1 order frequency, as well as the corresponding results in three directions simultaneously. Then, the frequency and angle values are calculated to evaluate the quality of the parameters in the three directions :

$$\begin{aligned} Freq_{m,n} &= \left| p_{nx}(CC_m), p_{ny}(CC_m) \right| \\ Ang_{m,n} &= a\tan\frac{p_{ny}(CC_m)}{p_{nx}(CC_m)} \end{aligned} \tag{S14}$$

where the range of m is 1-3, representing different fringe directions, while the range of n is 1 and 2, representing the results of the 1st and 2nd order frequency peak estimation, respectively.Therefore, the estimated parameters can be expressed as:

$$p_{nx,y} = \begin{cases} p_{nx,y}(CC_2) & \begin{bmatrix} std(Freq_{2,1}, Freq_{2,2}, Freq_{2,3}) < \theta_{freq} \\ std(Ang_{2,1}, Ang_{2,2}, Ang_{2,3}) < \theta_{ang} \end{bmatrix} \\ p_{nx,y}(CC_1) & else \end{cases} \tag{S15}$$

The +2 order frequency peak estimation result is confirmed to use firstly. When the standard deviation of frequency and angle obtained based on the +2 frequency components is greater than the threshold θ_{freq} and θ_{ang}, this result is considered to be of lower quality, and the result estimated with +1 frequency components is used to to reconstruct the super-resolution information.

S4. 3DSIM reconstruction principle

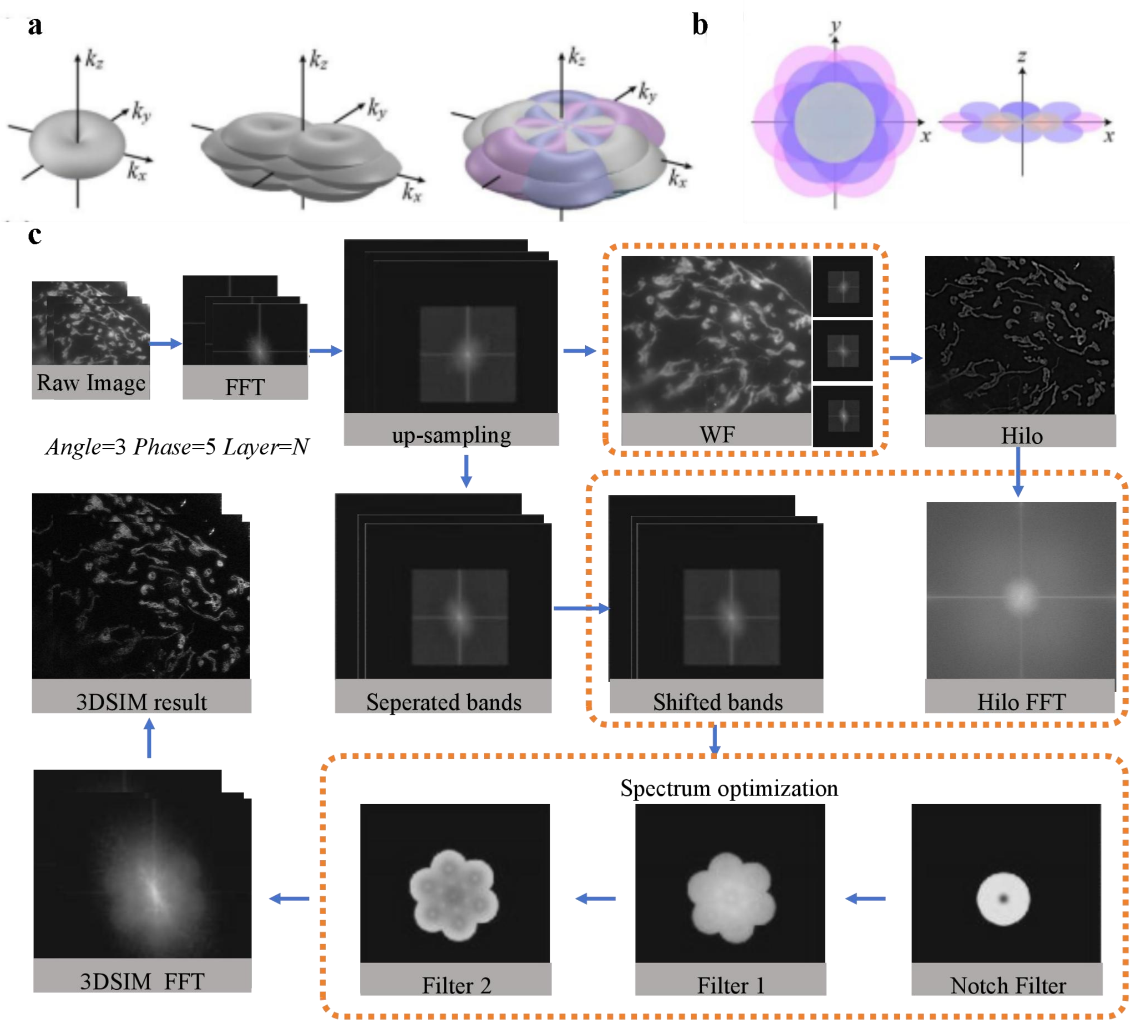

Fig. S2 The principle of the proposed method. (a) is the three-dimensional spectrum of 3DSIM in the xoy plane and the yoz plane. (b) is the two-dimensional spectrum of unidirectional 3DSIM and isotropic 3DSIM. (c) is the 3DSIM algorithm flow.

The reconstruction process is shown as Fig.S2. Five frequency components $C_n(k)$ are shifted to the correct position $C_{nr}(k)$ based on parameter estimation results respectively. Where n=0,±1 and ±2 donates the index of five frequency components. Therefore, the shifted spectra $C_{nr}(k)$ can be represented as:

$$C_{nr}(k) = F\left\{F^{-1}\left[C_n(k)e^{-jn2\pi(p_x,p_y)}\right]\right\} \tag{S16}$$

Here, F and F^{-1} represent the three-dimensional Fourier transform and the inverse three-dimensional Fourier transform respectively.

The first spectral optimization is the applied of notch filters to suppress the reconstruction artifacts caused by high-frequency spectral peak points. In general reconstruction results, the background information mainly comes from 0 frequency, which is reflected in the reconstruction result when 0 frequency is added to the result. This background can be suppressed by using a notch filter. But the filter creates artifacts while removing the background. In order to solve the problem of background interference in the reconstruction results, Hilo result with lower background are applied to represent zero frequency. In this way, the background interference of the reconstruction results is weaker and the reconstructed details are more prominent. The combined spectra $C_{SR\_0}(k)$ can be represented as:

$$C_{SR\_0}(k) = \sum_{n=[-2,2],\ n\neq 0} C_{nr}(k) north \cdot OTF^{att} + Hilo(k) \tag{S17}$$

Where north(x,y,z,0) is the notch-filter designed with the estimated frequency vector on the *xoy* and *yoz* plane.

$$north(x, y, z, 0) = 1 - d \cdot e^{\frac{\frac{x^2+y^2}{|px, py|^2} + \frac{z^2}{|pz|^2}}{2w^2}} \tag{S18}$$

In equation (10) $north(x,y,z,n)$ is the shifted notch-filter in frequency-domain position based on $north(x,y,z,0)$. Similarly, $OTF^{att}(x,y,z,n)$ is the shifted result with attenuation of the optical transfer function to the n-th position in the frequency domain. d and w are the notch depth and width respectively.

In order to suppress plaque artifacts after reconstruction and further improve the weak signal retention ability of the image, a spectral filter is constructed based on apodization function (Apo), optical transfer function and notch-filter notch width w:

$$Filter(w)=\frac{Apo}{\sum_{n=-2}^{2} m(n)\cdot OTF(k,n)\cdot north(k,n)+w^2} \quad (S19)$$

The optimized $C_{SR\_0}(k)$ with the spectrum filter can be expressed as $C_{SR\_1}(k)$:

$$C_{SR\_1}(k)=C_{SR\_0}(k)Filter(w_1)Filter(w_2) \quad (S20)$$

So, the final 3DSIM super-resolution result is S_{3DSIM}:

$$S_{3DSIM}(x,y,z)=F^{-1}\left[C_{SR\_1}(k)\right] \quad (S21)$$

Due to the isotropic nature of super-resolution, it is necessary to rotate the direction of the other two sets of fringe by 120 degrees separately, and then perform the same operation on the original image.