Optics-aware super-resolution light-field microscopy for long-term volumetric imaging of dynamic intracellular processes at millisecond timescales

Lanxin Zhu[1,4], Jiahao Sun[1,4], Chengqiang Yi[1,4], Meng Zhang[2,4], Guangda Niu[3], Jiang Tang[3], Yuhui Zhang[2], Dongyu Li[1] and Peng Fei[1*]*

[1] *School of Optical and Electronic Information-Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, 430074, China.*

[2] *Britton Chance Center and MoE Key Laboratory for Biomedical Photonics, Wuhan National Laboratory for Optoelectronic, Huazhong University of Science and Technology, Wuhan, China.*

[3] *Wuhan National Laboratory for Optoelectronics-School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China.*

[4] *These authors contributed equally: Lanxin Zhu, Jiahao Sun, Chengqiang Yi, Meng Zhang*

[*] *Correspondence：feipeng@hust.edu.cn; li_dongyu@hust.edu.cn*

Abstract

Long-term and high-speed 3D imaging of live cells beyond diffraction limit remains a big challenge for current super-resolution microscopy implementations, owing to their severe phototoxicity and limited volumetric imaging rate. While emerging light-field microscopy (LFM) has mitigated this issue through rapid and mild 3D imaging of dynamic biological processes with single 2D snapshots, it suffers from a suboptimal spatial resolution close to diffraction limit, which greatly compromises its applications for live-cell imaging. Here, we propose a super-resolution light-field reconstruction strategy based on an optics-aware view-channel-depth (VCD 2.0) deep-learning strategy. Through a stepwise inversion of the physical degradation process during light-filed imaging, VCD 2.0 strongly pushes the light-field microscopy beyond diffraction limit, and achieves instant 3D reconstructions from raw 2D light-field snapshots with an ~14-fold resolution improvement. After being combined with a simple-and-robust light-field add-on design, VCD 2.0 further enables long-term 3D super-resolution imaging (over 30000 volumes) of various intracellular dynamics (FOV of ~220×220×10 μm^3) even on an ordinary 2D fluorescence microscope. Using this novel and readily-accessible approach, we successfully capture the fast morphology changes of diverse organelles, such as the extension and constriction in lysosome, at a near isotropic resolution of ~120 nm, and a high volume rate up to 50 Hz.

Introduction

Elucidating the volumetric architecture of organelles and molecules inside cells and their interactions requires microscopy methods with a sufficiently high spatiotemporal resolution in all three dimensions[1-4]. However, observing millisecond processes of 3D subcellular structures in living cells poses great challenge for current 3D microscopy owing to the constraints of the maximal photon budget a sample permits, resulting in an inevitable trades-off among imaging speed, spatial resolution, and photon efficiency[5-11]. A series of scanning-based 3D microscopy technique, including confocal microscopy[11], 3D structured illumination microscopy (3D SIM)[5, 7]

and light sheet microscopy (LSM)[9, 12] , have been developed for subcellular intravital imaging with high spatial resolution (~120 nm at lateral and 280-360nm at axial) through multi-frames acquisition but at a cost of reduced temporal resolution (limited to seconds per volume) and photon efficiency, making it difficult to observe life activities occurred in milliseconds.

Light field microscopy (LFM) is another attractive choice among 3D microscopy providing photon-efficient and high-speed volumetric imaging by encoding both position and angular information of 3D signals on single 2D camera snapshots without time-consuming axial scanning [13-18]. Benefiting from its high-speed scanning-free 3D imaging, LFM has facilitated various studies in intracellular imaging of live cells[16, 17, 19]. However, the spatial resolution is sacrificed for the simultaneous acquisition of multiple views in limited camera sensor. Though various efforts have been made to pushes the sacrificed resolution to the diffraction limit, current deconvolution-based light-field reconstruction algorithm without any prior information of fine structures can only yield limited resolution restricted by the optical system, compromising its potential for observing the nanoscale structures in live cells, such as the membrane of the organelles. In a word, there always exists an “impossible performance triangle” of current microscopy that constraint by imaging speed, spatial resolution and photon efficiency.

The incorporating of deep neural networks enlarges this design space of microscopy by introducing prior knowledge of high-resolution data[20, 21]. Previous view-channel-depth light-field microscopy (VCD-LFM) is capable of directly reconstructing high-resolution 3D volume from 2D light-field (LF) raw data by splitting various views from LF and incorporating successive extracted features in network into multiple channels to yield 3D image stacks, successfully pushing the spatial resolution of LFM to diffraction limit and showing outstanding performance in imaging cellular structures[22-24]. However, learning the mapping function to reconstruct from 2D under-sampling LF to 3D super-resolution (SR) volumes is challenging for usual one-stage model, since the degradation model of light field imaging is an extremely intricate process coupling multiple resolution degradation and noise and compressing the space bandwidth product by ~500 times, resulting in the limited performance of reaching sub-diffraction-limited resolution with high fidelity.

Herein, to further transcend the resolution limit of LFM, we developed a deep-learning based super-resolution LFM reconstruction strategy (VCD 2.0) to directly reconstruct a super resolution 3D volume from an under-sampling low-resolution 2D light field snapshot. According to the physical prior of light field degradation model, we specifically designed a multi-progressive light-field-aware network to decouple the complex inverse problem of LFM progressively and a light-field-specific pre-processing pipeline to generate synthetic multi-stage priors for progressively guiding the network. Combined with a simple and robust light-field add-on device that is designed for the rapid implementation of light-field imaging on commercial microscope, our VCD 2.0 is demonstrated to enables long-term visualization (over 30000 volumes) of the transient dynamics of intracellular structures in multiple live cells (at a FOV of 220×220×10 μm^3), at an isotropic spatial resolution of roughly 120 nm in three dimensions, yielding dozens of Giga voxels per second.

Results

Principle and Implementation of VCD 2.0

The VCD 2.0 was derived from our previous deep-learning-based view-channel-depth strategy (VCD 1.0) which allows for direct light-field reconstruction and enhancement[22]. However, the inference of SR images from low resolution (LR) images is an ill-posed problem since there exist infinite SR images that can be down sampled to LR images, the space of possible functions that map LR images to SR images can be very large depending on the complexity of degradation model[25, 26]. Hence, the space of candidate mapping function from 2D under-sampling LF to 3D SR volumes is significantly large, as the degradation encodes multidimensional information which compresses the space bandwidth product by ~500 times. This multidimensional degradation includes the degradation of axial resolution restricted by the NA of objectives, the noise caused by camera sensor and the under-sampling and dimension compression during light field encoding by a 5D light-field point spread function (PSF) which is far more complicated than common 2D or 3D PSF. Learning a good solution to decode these complex degradations at various dimensions is challenging for usual one-stage model, resulting in the limited performance in resolution and fidelity when performing subcellular structures[22].

To minimize the hypothesis space of the final solution and further transcend the spatiotemporal resolution limit, VCD 2.0 disentangles the complex light field inverse problem by designing four task-specific progressive modules to solve the mapping function of LF denoise, LF de-pixelization, 3D reconstruction and isotropic restoration step by step and incorporating multi-stage data priors to progressively guide these inverse tasks (Fig.1a). Furthermore, each function module is task-specific customized to further constrain the candidate function space by decoupling multi-dimensional LF feature, incorporating LF-specific loss based on the physical principle of light field and jointly optimized by weighting each module to regularize the network output. Firstly, to denoise the LF views with various signal-to-noise ratio (SNR), a view-attention denoise module was designed to adaptively capture discriminative noise features across different views, resulting in a substantial improvement of the SNR compared to the raw noisy LF. Next, the denoised LF were then input in a spatial-angular light field super-resolution module (SA), which extracted spatial and angular feature of LF by two convolutional kernels with different size. An EPI loss was designed to maintain the parallax of views while deblurring, which was crucial for the post reconstruction of axial information. The views of SR LF showed a substantially enhancement in the resolution by the SR LF module while remaining the high accuracy of the parallax. Then, a light field reconstruction module is revised based on our VCD 1.0 to reconstruct the SR LF images, which replaces the upsampling convolutional layers of U-Net block with multires blocks to reconcile the features at different scales. Finally, a 2D U-Net is used to enhance the axial resolution by incorporating the PSF prior of the optical system. These four modules are optimized jointly through inputting the output of each module into next module during each iteration and optimizing the loss functions of each module together, which can restrain the output of each module mutually and promotes the final output to be optimal. After being optimized to its optimal state, the network was capable of directly inferring 3D SR movies from the sequential light-field measurements (Fig. 1b). These multi-progressive reconstruction with intermediate

priors and joint optimizing strategy facilitated the convergence of the network and enhanced the accuracy of the inference results. As verified by the result of Airy scan microscopy of the same volume, the fine structure of lysosome was correctly reconstructed throughout the 3D volume and the fidelity was quantified to be significantly higher than other 3D imaging methods (Fig.2).

It is worth mentioning that, to generate training data without a complex system to in situ acquire SR and LF data, a data preprocessing pipeline was developed to synthesize multi-stage data priors (including the raw LF, denoise LF, SR LF and 3D SR) based on light-field propagation only from the SR label data (e.g. data acquired from Airy Scan microscopy) (Fig.1a), progressively guiding each module. It's notes worthy that we developed a strategy to generate SR LF, providing prior knowledge for the light field super resolution module to reduce the space of the mapping function of network. Unlike other light field super resolution network that generate training pairs by degrading the raw LF data as LR and using raw LF as GT, we synthesized the SR LF images without under-sampling views as label data. Since the low spatial resolution of LF images majorly caused by the under-sampling of microlens, we shifted the 3D SR volumes by the distance of the sub-microlens along the horizontal (both x and y respectively) and projected the shifted 3D volumes based on wave optics model, yielding multiple light field projections that include more spatial information through multiple sampling. The shifted light field projections are then realigned to generate the SR LF images according to the arrangement of the light. The resolution of synthetic SR LF images exceeded raw LF images by the improved sampling rate, bringing in a more accurate optimizing direction of mapping function. Another key point lied in the simulation of LF with accurate SNR matched with the experimental LF. The LFs without noise were first projected from 3D SR by convolving with the 5D PSF of LFM, which were also used for the intermediate label data for LF SR module. The intensity of clean LF images was normalized to the same range of experimental LFs and various noise in the range of the highest and lowest noise in the experimental LFs during long-term observation was applied on the clean LFs. The SNR of synthetic noisy LF was validated to be at the same level with experimental LF.

Furthermore, for wide application of our 3D super-resolution imaging strategy, a simple and cost-effective light-field add-on device (Add-on LFM) based on commercial microscopy was designed (Fig.1b). The implementation of this device only requires the assembling of off-the-shelf lenses and devices at low cost and mounting on the camera port of the commercial microscope. There is no requirement for any collimation of the light path and only a compact mechanic needed to slightly adjust for precisely projecting the back focal plane of the microlens array (MLA) on the camera sensor plane, which can be easily applied even for the biological labs without any optics background.

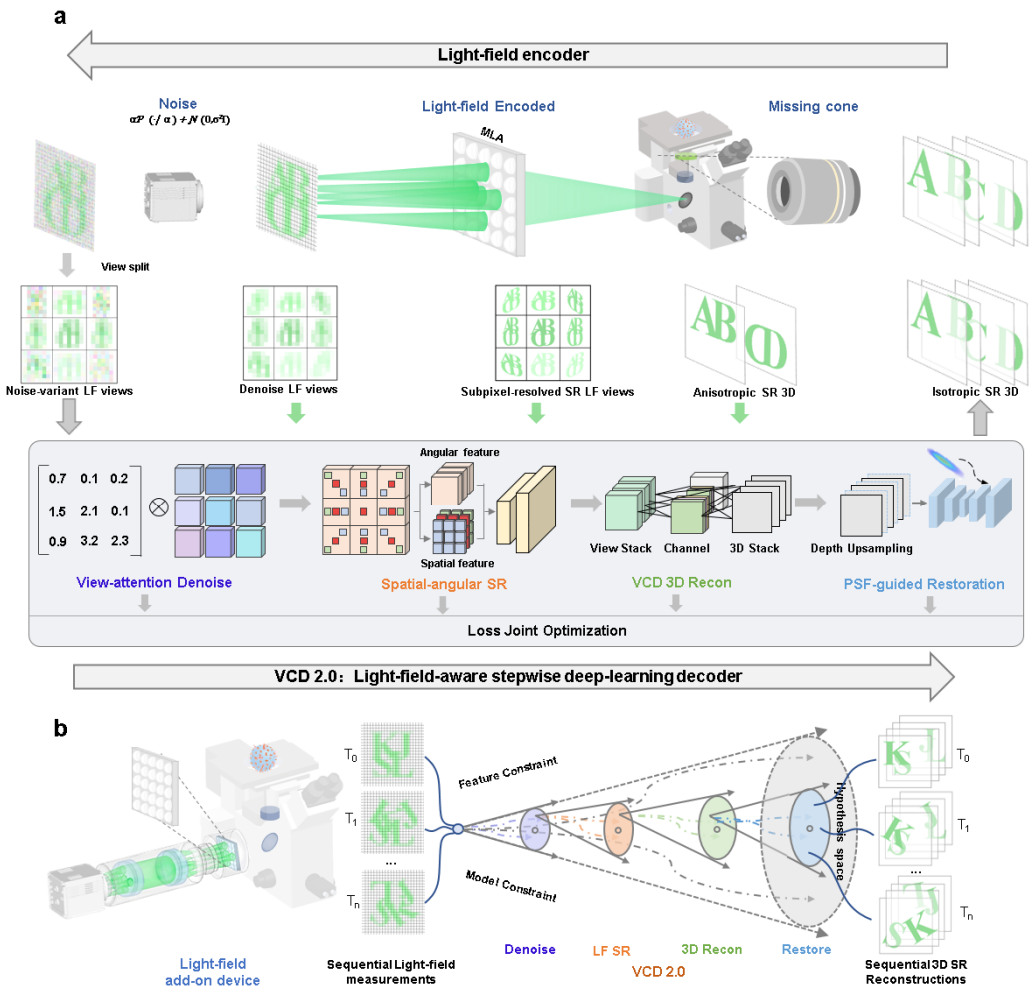

Fig.1 Principle of VCD 2.0. a, The multi-stage architecture of VCD 2.0 and the schematic of its training pipeline. The design of multi-stage network is an inverse process based on the physical model of the light-field encoding process, including denoise, super resolution, 3D reconstruction and restoration. **b,** The schematic of the inference process of VCD 2.0, including the acquisition of sequential light-field measurements from the light-field add-on device and the inference of VCD 2.0. The final hypothesis space of our VCD 2.0 indicated by the blue area could be substantially smaller than one-stage network (indicated by the gray area.

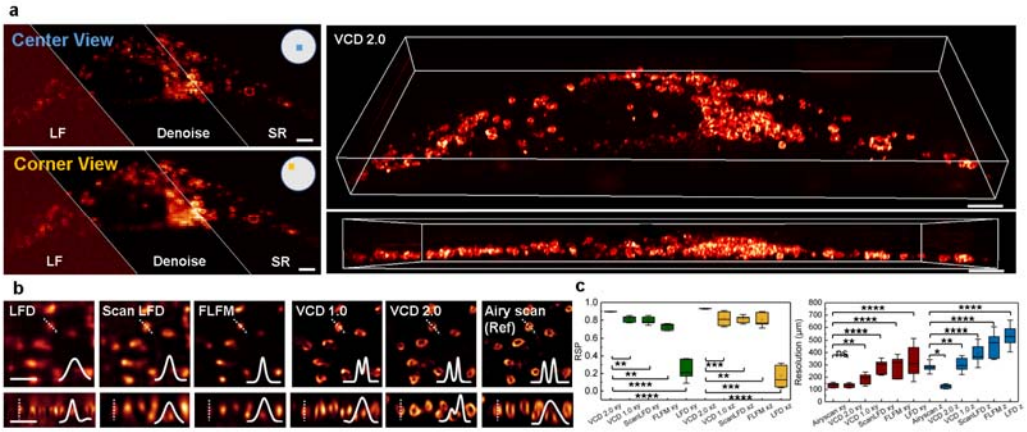

Fig. 2 The fidelity demonstration of VCD 2.0 on lysosome outer membranes. a, The denoise, SR and reconstructed results of VCD 2.0. Scale bar, 5 μm. **b,** The comparisons of the reconstructed results between our VCD 2.0 and current leading light field microscopy. White lines, intensity profiles of the areas indicated by dashed lines. Scale bar, 2 μm. **c,** Statistical comparisons of spatial resolution and fidelity between our VCD 2.0 and current leading light-field microscopy (n=20).

Characterization of Performance:

Using a inverted commercial microscope equipped with the add-on LFM module with a size of only ~220 mm×140 mm, light field imaging at a volumetric imaging rate up to 200 Hz and the maximum FOV of ~ 200 × 200 × 10 μm^3 containing dozen of cells were achieved under a × 60/1.3NA objective (Fig.3a). Lysosomes within an entire live cell were imaged and reconstructed via LFD, SA LFD, VCD 1.0 and VCD 2.0, respectively (Fig. 3b, c). The reconstructed results via VCD 2.0 yielded an isotropic resolution of ~120 nm (Fig. 3d) while other algorithm yielded an anisotropic resolution (LFD, SA LFD, VCD 1.0 yielded an average lateral resolution of 400 nm, 250 nm, 180 nm and an average axial resolution of 450 nm, 350 nm, 350 nm, respectively). The subcellular isotropic resolution enabled VCD 2.0 to clearly resolve the outer membrane of the lysosomes while the lysosomes remained unresolvable in LFD, SA LFD and VCD 1.0 and suffered from suboptimal resolution and artifact. The hundred-hertz volumetric imaging rate and subcellular resolution thus exclusively allowed the visualization of the fast deformation of the lysosome in 3D occurred in milliseconds (Fig. 3e). In addition, the inclusion of the denoise module resulted in a stable capability of VCD 2.0 to restore noisy data. As the LFM only needed one exposure per volume, the phototoxicity was intrinsically lower than the scanning-based 3D microscopy. The combination of the low phototoxicity with the strong denoise capability permitted live-cell imaging at 30000 volumes, yielding 4 million SR images with less than 50 % photobleaching while the signals have been photobleached after scanning imaging of 10 and 100 volumes by wide field and Airy scan imaging, respectively (Fig. 3f, g). In conclusion, current 3D microscopy is restricted by the “impossible triangle” which is the inevitable trade-off among imaging speed, spatial resolution, and photon efficiency owing to the constraints of maximal photon budget a sample permits (Fig. 2h). The point-scanning microscopy such as Airy scan achieves high spatial resolution but low volumetric speed and photo-efficiency, which lies in the

vertex of spatial resolution. The wide-field microscopy with the volume illumination and planes scanning has high photon efficiency while the spatiotemporal resolution is sacrificed. The LFM modality including Scan LFD and FLFM obviously attains fast volumetric speed by sacrificing the spatial resolution. The previous VCD-LFM breaks through the trade-off between volumetric speed and photo-efficiency, but still suffers low spatial resolution. Herein, our current VCD 2.0 thoroughly exceeds the constraints of this “impossible triangle” and shows much superior performance in both spatiotemporal resolution and photon efficiency.

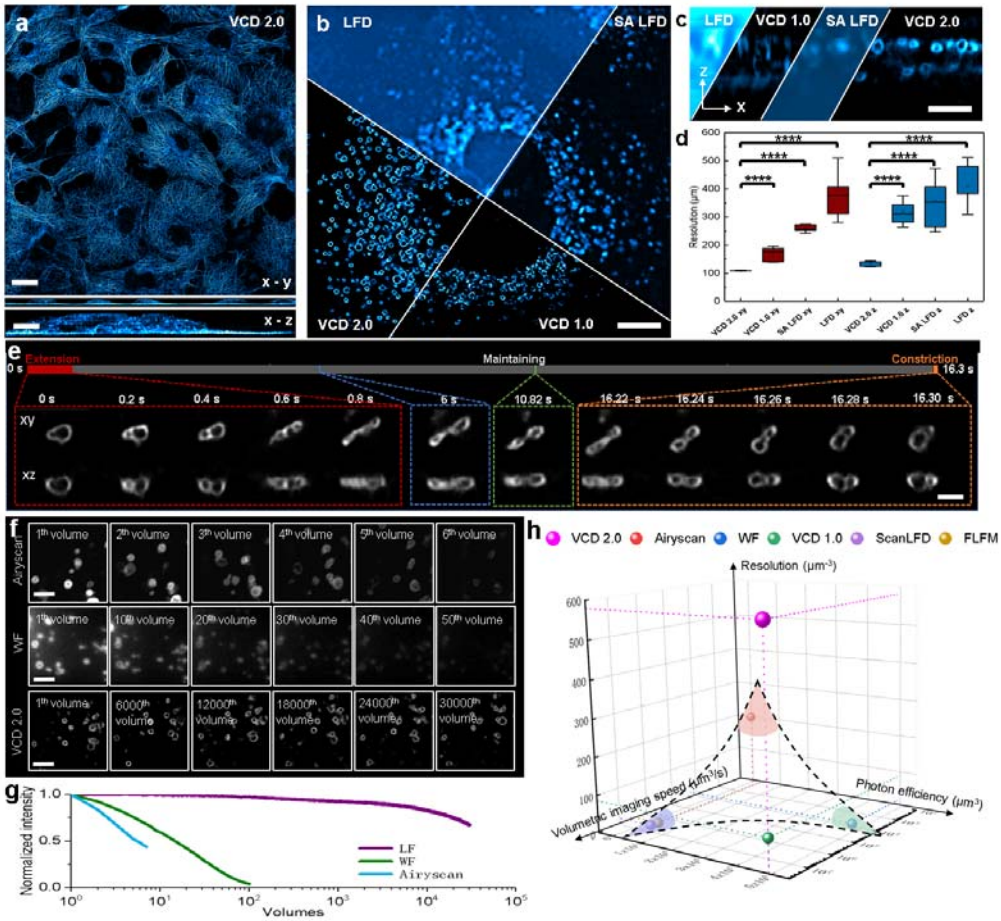

Fig. 3 The performance of VCD 2.0. a, The xy and xz MIPs of the VCD 2.0 result of microtubules in 13 fixed cos-7 cells within a large FOV of 220×220×10 μm3. Scale bar, 20 μm. **b, c,** The xy and xz MIPs of the reconstructed results of lysosome outer membranes in U2OS cells by LFD, SA LFD, VCD 1.0 and VCD 2.0. Scale bar, 10 μm. **d,** Decorrelation analysis quantifying spatial resolution of LFD, SA LFD, VCD 1.0 and VCD 2.0 (n=20). **e,** Time-lapse 3D visualization of the fast-morphological changes of lysosome outer membrane. Scale bar, 2 μm. **f,** The performance of the VCD 2.0, WF and Airyscan results in live U2OS cells along long timescale observation. Scale bar, 5 μm. **g,** Comparisons of the photobleaching rates between VCD 2.0, WF and Airyscan. h, The comparisons of the performance in both spatiotemporal resolution and photon efficiency between VCD 2.0, Airyscan, WF, VCD 1.0, ScanLFD and FLFM.

Methods

Add-on LFM device setup

An add-on device is designed to provide a fast and easy means for building a compact optical assembly for live-cell light-field imaging based on the commercial epifluorescence microscope. The device only requires the assembling of off-the-shelf lenses and few simple customizations without any complex alignment of the light path, which enables wider application of researchers that without optics background. To position MLA at native image plane of the microscope, a customized lens tube with a slot is designed to fix the MLA (MLA-S100-f28) at the Flange focal distance away from the camera port and one end of the lens tube can be easily mounted on the microscope by matching the thread. Then a pair of relay lens (TTL100-A) is utilized to relay the back focal plane of MLA on the camera sensor. To make sure the back focal plane is relayed on the camera sensor precisely, a high-precision zoom housing is used to connect the customized lens tube and the relay lens and tune the distance between them. Finally, a customized lens tube with a fixed length decided by the focal length of relay lens and Flange focal distance of camera is designed to mount the relay lens and camera. Optionally, this add-on light-field device can be extended to dual-channel detection by dividing after MLA via a dichromatic mirror and adding an extra camera. In dual-channel setup, the relay lens should be replaced by the tube lens with longer focal length (TTL180-A) and two mirrors are used to minimize the size of the assemble device. See Supplementary for more details about our add-on light-field device.

VCD 2.0 network

Previous deep-learning-based light-field reconstruction network enable resolution enhancement with data prior but simplify the multistage process of image degradation in LF. This one-step network produces hallucination or artifacts when solving hard inverse problem (ref multi stage super resolution). Similar with the physical process of light-field image generation, our VCD2.0 network disentangles this inverse problem into four sub-problems: denoise, LF de-pixelization, 3D reconstruction and isotropic restoration. The four sub-problems are divided and conquered by jointly optimizing task-specific networks with weighted objective function.

The architectures of the LF denoise net are based on channel-attention mechanism to reweight the extracted feature from input views with different SNR. This denoise module can be formulated as

$$x_1 = Fusion\left(Attention^{(2)}\left(Conv_{dilated}(x)\right) + Attention^{(1)}\left(Conv_{dilated}(x)\right)\right)$$

where $Conv_{dilated}$ means dilated convolution that enlarge receptive field. $Fusion$ is feature element-wise adding. In order to suppress the noise fluctuation, L1 loss is used as loss function during denoising stage:

$$Loss_{denoise} = MAE(x_1, y_1)$$

Due to the angular and spatial information encoded in LFs, directly applying traditional 2D super resolution network like ResNet, RCAN on extracted views induces the under-exploitation of the rich angular features. For LF de-pixelization, we adopt two types of dilated convolution operations [ref. SA] on 4D light-filed images to separate 2D spatial information and 2D angular information. These two convolution and series of activation function and feature fusion layers form Spatial-Angular (SA) block. According to sequential connection of four SABlocks and

Upsampling convolution (SI Note), the missing spatial information caused by sparse sampling rate of microlens array can be recovered:

$$x_2 = Upsampling(Fusion(Conv_{dilated}(x_1) + \underbrace{SABlock\big(\ldots SABlock(x_1)\big)}_{n\times SABlock}))$$

where $Fusion$ is channel concatenate operation. Considering the potential parallax information of various views, epipolar constraint ($EPI_{consistency}$) on SR prediction is proposed to keep geometry relationship among cross-views. The objective function of LFSR can be formulated as:

$$Loss_{LFSR} = MSE(x_2, y_2) + EPI_{consistency}(x_2, y_2)$$

where MSE means mean squared error between network output and digital scanning LF.

For 3D reconstruction model, we extended original VCD to handle super-resolution reconstruction. According to adding MultiRes blocks and modifying activation function (SI Note), the capability of VCD network is enhanced which achieves lower fitting errors when solving SR reconstruction:

$$x_3 = Unet_{MultiRes}(Conv(SubpixelConv(x_2)))$$

Correspondingly, for 3D reconstruction task, we adopt MSE loss and lateral gradients loss to preserve the structure information of ground truth.

$$Loss_{Reconstruction} = MSE(x_3, y_3) + Gradient_{xy}(x_3, y_3)$$

At isotropic restoration stage, similar with CARE[], a pre-measured PSF is used to blur the lateral slices of 3D stacks to mimic the image degradation of axial ones. Here, we establish the mapping between blurred slices from previous 3D prediction and high-resolution ones from ground truth by 2D Unet, which avoids the influence brought by errors of 3D prediction.

$$x_4 = Unet_{Isotropic}(x_3)$$

The loss function as below

$$Loss_{Isotropic} = MAE(Lateral(y_3), Lateral(Degraded(x_3)))$$

These four networks are optimized synchronously based on designed weighted objective function:

$$Loss = \lambda_1 Loss_{denoise} + \lambda_2 Loss_{LFSR} + \lambda_3 Loss_{Reconstruction} + \lambda_4 Loss_{Isotropic}$$

where λ_1, λ_2, λ_3, λ_4, are weight coefficients for different network respectively which are set as 0.05, 0.2, 0.8, 1.0. With this multi-loss optimization strategy and multi-GT supervision, the candidate solutions for LF inverse problem are reduced, which is important for yielding high-resolution 3D stacks without artifacts.

Training data acquisition and data processing pipeline

A data processing pipeline is developed to generate all training data (including the raw LF, denoise LF, SR LF and 3D SR) from the 3D SR data. In this paper, 3D SR data were acquired from AiryScan microscopy by imaging fixed cells. Firstly, a processing pipeline is designed to create SR LF images to guide the LF super resolution task. The 3D SR volumes are shifted by the distance of the sub-microlens along the horizontal (both x and y respectively) and then are projected by convolving with 5D PSF of LFM to yield multiple light field projections that include more spatial information through multiple sampling. The shifted times depend on the

under-sampling rate of the microlens, which is 3x3 times for the setup of Nnum 7 in this paper. The shifted light field projections are then realigned to generate the SR LF images according to the arrangement of the light. Secondly, the clean LF images is derived from the center image of SR LF images to guide the LF denoise task. Finally, to synthesize noisy LF with accurate SNR matched with the experimental LF, the clean LF images was normalized to the same range of experimental LFs and various noise in the range of the highest and lowest noise in the experimental LFs during long-term observation was applied on the clean LFs to generate noisy LF. The SNR was calculated by:$\mathrm{SNR} = \frac{\mu}{\sigma} = \frac{\mathrm{S} - I_{background}}{\sqrt{S - I_{background + \sigma^2_{background}}}}$, where S is the average signal intensity value in the image, $I_{background}$ and $\sigma_{background}$ are the mean and the standard deviation of the background, respectively.

Assessment of the resolution improvement and fidelity of VCD 2.0 network

We used the resolution-scaled Pearson coefficient (RSP) to assess the fidelity of our VCD 2.0 reconstructions and other 3D microscopy techniques including LFD ScanLFD, FLFM, VCD 1.0 using the AiryScan data as reference in both the xy and yz planes by using the NanoJ-SQUIRREL. We applied decorrelation analysis to quantify the resolution of VCD 2.0 results and other 3D microscopy. The analysis was conducted using the MATLAB. The axial resolution is measured by the sectorial resolution mode in the decorrelation analysis that only calculate the resolution in a narrow sectorial region along z direction.

Live cell imaging

Long-term light field imaging of the lysosome outer membrane in live U2OS cells is implemented using our add-on light-field device with a ×60/1.3NA objective (OlympusXXX) at an volumetric imaging rate of 200 Hz. The light-field imaging of live cells sustains 30000 volumes with less than 50 % photobleaching, yielding 4 million SR images by VCD 2.0 reconstructions. The fast deformation of lysosome outer membrane occurred in milliseconds could be observed by our VCD 2.0. To compare the photon bleaching rate with other 3D microscopy, we also imaged the lysosome outer membrane at the same condition by Airyscan and WF microscope. The laser intensity was carefully adjusted to make sure the SNR of the images of these microscopes is at the same level. The bleaching rates were then calculated using Matlab. Multiple areas are cropped and a threshold is used to discriminate the signal areas and background areas automatically. The photon bleaching rate over time was finally calculated using the following equation:$\mathrm{I(t)} = I_{signal}(t) - I_{background}(t))/MAX(I_{signal}(t) - I_{background}(t)$, where I_{signal} and $I_{background}$ are the mean intensity value of the multiple regions of signal and background images, respectively.

Discussion

In summary, our super-resolution VCD 2.0 model in conjunction with a compact light-field add-on design together readily enables sustained volumetric imaging of mitochondria-vesicle interactions at ~120-nm isotropic resolution, and high volume rate up to 50 Hz, yielding a spatiotemporal performance superior to the alternative 3D microscopy techniques, such as confocal microscopes, light-sheet microscopes, and other previously reported light-field microscopes. As compared to the end-to-end architecture in our previous VCD 1.0 LFM implementation, VCD 2.0 for

super-resolution LFM is a new optics-aware network that can progressively decouple the complex inverse problem in the optics of LFM and thereby achieve much improved light-field reconstructions with ~1.5 fold resolution enhancement and over 90% accuracy rate.

It should be noted that our approach is based on a simple-and-robust add-on device together with a generalizable computation framework, making it especially valuable for advancing cell biology research with conventional lab microscopes. For an instance, our light-field add-on only requires the assembling of a few off-the-shelf lenses and devices, and then can be easily attached to an inverted microscope. Meanwhile, VCD 2.0 only requires a small amount of label datasets, which could be obtained by any types of SR microscope, to generate all the priors for model training, and thereby eliminates the need of building complicated system for acquiring paired SR and LF datasets *in situ*. In addition to standard LFM, VCD 2.0 is also compatible with other LFM modalities, such as Fourier LFM, to improve the imaging performance without the need for complex hardware retrofit. We note that this network still requires abundant training blocks for model training and can only be generalized well on similar types of samples. We anticipate that this will be circumvented by incorporating weak training supervision or self-supervision strategies, which we are also developing now. We expect that our super-resolution VCD 2.0 could be a new technique-of-choice that helps the cell biologists to reveal numerous mysteries in the cells.

1. Valm, A.M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162-167 (2017).
2. Sahl, S.J., Hell, S.W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nat Rev Mol Cell Biol* **18**, 685-701 (2017).
3. Choquet, D., Sainlos, M. & Sibarita, J.B. Advanced imaging and labelling methods to decipher brain cell organization and function. *Nat Rev Neurosci* **22**, 237-255 (2021).
4. Schermelleh, L. et al. Super-resolution microscopy demystified. *Nat Cell Biol* **21**, 72-84 (2019).
5. Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., ... & Sedat, J. W. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320.5881**, 1332-1336 (2008).
6. Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**, 810-813

(2008).

7. Shao, L., Kner, P., Rego, E.H. & Gustafsson, M.G. Super-resolution 3D microscopy of live whole cells using structured illumination. *Nat Methods* **8**, 1044-1046 (2011).

8. York, A.G. et al. Instant super-resolution imaging in live cells and embryos via analog image processing. *Nat Methods* **10**, 1122-1126 (2013).

9. Chen, B.C. et al. Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).

10. Li, D. et al. ADVANCED IMAGING. Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science* **349**, aab3500 (2015).

11. Wu, Y. et al. Multiview confocal super-resolution microscopy. *Nature* **600**, 279-284 (2021).

12. Gao, L. et al. Noninvasive imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens. *Cell* **151**, 1370-1385 (2012).

13. Broxton, M. et al. Wave optics theory and 3-D deconvolution for the light field microscope. *Opt Express* **21**, 25418-25439 (2013).

14. Cohen, N. et al. Enhancing the performance of the light field microscope using wavefront coding. **22**, 24817-24839 (2014).

15. Prevedel, R. et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nat Methods* **11**, 727-730 (2014).

16. Guo, C., Liu, W., Hua, X., Li, H. & Jia, S. Fourier light-field microscopy. *Opt Express* **27**, 25573-25594 (2019).

17. Wu, J. et al. Iterative tomography with digital adaptive optics permits hour-long

intravital observation of 3D subcellular dynamics at millisecond scale. *Cell* **184**, 3318-3332 e3317 (2021).

18. Zhang, Z. et al. Imaging volumetric dynamics at high speed in mouse and zebrafish brain with confocal light field microscopy. *Nat Biotechnol* **39**, 74-83 (2021).

19. Hua, X., Liu, W. & Jia, S. High-resolution Fourier light-field microscopy for volumetric multi-color live-cell imaging. *Optica* **8**, 614-620 (2021).

20. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat Methods* **15**, 1090-1097 (2018).

21. Qiao, C. et al. Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes. *Nat Biotechnol* (2022).

22. Wang, Z. et al. Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning. *Nat Methods* **18**, 551-556 (2021).

23. Zhu, L., Yi, C. & Fei, P. A practical guide to deep-learning light-field microscopy for 3D imaging of biological dynamics. *STAR Protoc* **4**, 102078 (2023).

24. Zhu, T. et al. High-speed large-scale 4D activities mapping of moving C. elegans by deep-learning-enabled light-field microscopy on a chip. *Sensors and Actuators B: Chemical* **348** (2021).

25. Lempitsky, V., Vedaldi, A. & Ulyanov, D. in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition 9446-9454 (IEEE, 2018).

26. Guo, Y. et al. in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 5407-5416 (2020).