



Published in final edited form as:

Nat Methods. 2021 May ; 18(5): 551–556. doi:10.1038/s41592-021-01058-x.

Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning

Zhaoqiang Wang^{1,2,3,+}, Lanxin Zhu^{1,+}, Hao Zhang^{1,+}, Guo Li¹, Chengqiang Yi¹, Yi Li^{1,4}, Yicong Yang¹, Yichen Ding³, Mei Zhen⁵, Shangbang Gao^{4,*}, Tzung K. Hsiai^{2,3,6,*}, Peng Fei^{1,*}

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

²Department of Bioengineering, Henry Samueli School of Engineering and Applied Science, University of California, Los Angeles, Los Angeles, 90095, U.S.A.

³Division of Cardiology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, 90095, U.S.A

⁴College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China.

⁵Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, M5G 1X5, Canada.

⁶Division of Cardiology, Department of Medicine, Greater Los Angeles VA Healthcare System, Los Angeles, U.S.A

Abstract

Light-field microscopy has emerged as a technique of choice for high-speed volumetric imaging of fast biological processes. However, artefacts, non-uniform resolution, and a slow reconstruction speed have limited its full capabilities for *in toto* extraction of the dynamic spatiotemporal patterns

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence feipeng@hust.edu.cn (P.F.), THsiai@mednet.ucla.edu (T.K.H) and sgao@hust.edu.cn (S.G.).

+Equal contribution

Author contributions

P.F., Z.W. and H.Z. conceived the idea. P.F., S.G. and T.K.H. oversaw the project. Z.W., L.Z., G.L., Y.Y. developed the optical setups. H.Z., Z.W. and L.Z. developed the programs. Z.W., L.Z., G.L., Y.D. and Y.L. conducted the experiments, processed the images. Z.W., L.Z., H.Z., L.X., M.Z., S.G., T.K.H. and P.F. analyzed the data and wrote the paper.

Competing interests

The authors declare no conflicts of interest.

Data availability.

The datasets generated and analyzed in this study are available at <https://zenodo.org/record/4390067#.X-NsrNgzYuU>. The supplementary information are also available from the corresponding authors upon request. The example model checkpoints and data are packaged with the code. Source data for Fig. 1–3 are provided with this paper.

Code availability.

Codes for VCD-Net are available at <https://github.com/feilab-hust/VCD-Net>. Codes for data analysis and synthetic data generation used in current study are available from the corresponding authors upon request.

Editor summary: Reconstruction of light-field microscopy data with a deep learning network achieves high reconstruction speed and reduces artifacts, as illustrated for moving *C. elegans* and beating zebrafish hearts.

Editor recognition statement: Peer review information: Nina Vogt was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

in samples. Here, we combined a view-channel-depth (VCD) neural network with light-field microscopy to mitigate these limitations, yielding artefact-free three-dimensional image sequences with uniform spatial resolution and high video-rate reconstruction throughput. We imaged neuronal activities across moving *C. elegans* and blood flow in a beating zebrafish heart at single-cell resolution with volumetric imaging rates up to 200 Hz.

Introduction

A recurring challenge in biology is the extraction of ever more spatiotemporal information from targets, as many millisecond transient cellular processes occur in three-dimensional (3D) tissues and across long time scales. Several imaging techniques, including epifluorescence and plane illumination methods, can image live samples in three dimensions at high spatial resolution^{1–5}. However, they require recording a number of two-dimensional (2D) images to create a 3D volume, and the temporal resolution is compromised by the extended acquisition time of multiple frames by the camera.

Light-field microscopy (LFM) has become the technique of choice for instantaneous volumetric imaging^{6–14}. It permits the acquisition of transient 3D signals via post-processing of the light-field information recorded by single 2D camera snapshots. Because LFM provides high-speed volumetric imaging limited only by the camera frame rate, it has delivered promising results in various applications, such as the recording of neuronal activity^{7–10} and visualization of cardiac dynamics in model organisms¹¹. Despite these advances, the low and non-uniform spatial resolution and the presence of reconstruction artefacts prevents LFM's more widespread application for capturing millisecond time-scale biological processes at single-cell resolution^{7, 11}. While these problems can be mitigated by optimizing the way the light field is recorded^{9, 12, 14}, the extra system complexity could impede the wide dissemination of the LFM technique. Furthermore, current LFM's rely on a computationally demanding, iterative recovery process that limits the overall throughput of LFM reconstruction, compromising its potential for long time-scale applications.

Here, we describe a LFM strategy based on view-channel-depth neural network processing of the light-field data, termed VCD-LFM. Using a wave optics model¹³, we generated synthetic light-field images from high-resolution 3D images that we previously obtained experimentally, and paired them as input and ground-truth data for network training. We designed the VCD network (VCD-Net) to extract multiple views from these 2D light-fields and transform them back to 3D depth images, which would be compared with the high-resolution ground truth images to guide optimization of the network. Through iteratively minimizing the loss of spatial resolution by incorporating abundant structural information from training data, this deep-learning VCD-Net could be gradually strengthened until it is capable of deducing 3D high-fidelity signals at uniform resolution across the imaging depth. Once the VCD-Net is properly trained, it can deduce an image stack from a light-field measurement at a millisecond time scale, achieving a throughput suitable for time-lapse video processing. We demonstrated the ability of the VCD-LFM method by imaging motor neuron activity in L4-stage *C. elegans* rapidly moving inside a $300 \times 300 \times 50 \mu\text{m}$ microfluidic chamber, at an acquisition rate of 100 Hz and processing throughput of 13

volumes per second. This allowed us to extract the 4D spatiotemporal patterns of neuronal calcium signaling and track correlated worm behaviors at single-cell resolution, which was notably better than classic light-field deconvolution microscopy (LFD). Furthermore, we performed *in toto* imaging of blood flow in the beating heart of zebrafish larvae, enabling velocity tracking of blood cells and ejection fraction analysis of the heartbeat across a $250 \times 250 \times 150 \mu\text{m}$ volume at 200 Hz.

Results

Principle and performance of VCD-LFM

Our VCD-LFM involves training of a VCD convolutional neural network and its inferences based on the images obtained from LFM (Fig. 1a, Supplementary Fig. 1, 2). To create the data for network training, we first obtained high-resolution (HR) 3D images of stationary samples using synthetic or experimental methods (Fig. 1a). Using the wave optics model of LFM¹³, we projected these HR 3D images into 2D light-field images, which we then used as the input for network training (Fig. 1b, **step 1**, Supplementary Note 1). Each synthetic light-field image was first re-arranged into different views, from which features were extracted and incorporated into multiple channels in each convolutional layer. The final output channels were then assigned to a number of planes representing different depths to generate an image stack. Using cascaded convolution layers (U-Net architecture, Supplementary Fig. 3 and Supplementary Table 1) to repetitively extract features, this VCD procedure generated intermediate 3D reconstructions (Fig. 1b, **step 2**). The pixel-wise mean-square-error (MSE) was counted as the loss function to indicate how different these outputs were to the ground truth images. By iteratively minimizing the loss function (Fig. 1b, **step 3**), the network was gradually optimized until it could transform the synthetic light-fields into 3D images that were similar to the ground-truth images (Supplementary Fig. 4, Supplementary Note 2). After training on gigavoxels of data (Supplementary Table 2), the network was capable of implementing VCD transformations of sequential light-field measurements of dynamic processes, and inferring a sequence of 3D volumes at a rate up to 13 volumes s^{-1} (Fig. 1b, **step 4**).

We built a custom microscope to enable *in situ* light-field recording and 3D wide-field imaging (Supplementary Fig. 1). To demonstrate the capability of the VCD-LFM, we reconstructed sub-diffraction beads captured using a $40\times/0.8\text{w}$ objective, and quantified the resolution improvement resulting from the network by comparing the results with those from conventional light-field deconvolution (LFD, Fig. 1c–f). As verified by 3D wide-field imaging of the same volume, the fluorescence of individual beads was correctly localized throughout the volume (Fig. 1e, f, Supplementary Fig. 5). The VCD-LFM with a network trained on wide-field 3D image yielded an average resolution of $1.1 \mu\text{m}$ (x,y) and $3.0 \mu\text{m}$ (z) ($n = 2527$ beads), which was uniform across a $60\text{-}\mu\text{m}$ imaging depth ($1.0 \mu\text{m}$ (x,y) and $2.9 \mu\text{m}$ (z) at the best plane, $1.3 \mu\text{m}$ (x,y) and $3.1 \mu\text{m}$ (z) at the outer edge of the axial field of view) (Fig. 1g, Supplementary Fig. 6). This demonstrates notable improvements compared to LFD on the same image data with an average resolution of $2.6 \mu\text{m}$ (x,y) and $5.0 \mu\text{m}$ (z), and a broader distribution ($1.6 \mu\text{m}$ (x,y) and $3.8 \mu\text{m}$ (z) at the best plane, $4.0 \mu\text{m}$ (x,y) and $7.0 \mu\text{m}$ (z) at the outer edge of the axial field of view). We note that the performance of

VCD-LFM is dependent on training data; hence, the beads can be reconstructed isotropically (1.0 μm x,y,z) by including higher-resolution data in the training (Fig. 1f, g, Supplementary Fig. 7). Furthermore, VCD-LFM substantially removed the mosaic-like artefacts near the focal plane that are common in LFD (Fig. 1d, Supplementary Fig. 5, 6), and it performed well even when the signals were weak with high background noise or dense (Supplementary Fig. 8, 9). To further validate the accuracy of reconstructed signals, we applied VCD-Net to the reconstruction of synthetic firing neurons which are adjacent to each other. The improved image quality achieved by VCD-Net suppressed signal cross-talk from blurring and artefacts (Fig. 1h–j), and thus contributed to accurate recovery of signal fluctuations when recording the activity of densely labeled neurons. We further validated the reconstruction accuracy of VCD-Net on both static and moving neurons with varying signal magnitude and density (Supplementary Fig. 10, 11).

VCD-LFM reveals locomotion-associated neural activity in moving *C.elegans*

We show VCD-LFM to be suitable for capturing dynamic processes in live animals by demonstrating the imaging of neuronal activity in moving *C. elegans* and the beating heart of a zebrafish larva. We used a microfluidic chip to permit *C. elegans* (L4-stage) to rapidly move inside a micro-chamber (300 \times 300 \times 50 μm , Fig. 2a). We used a 40 \times /0.8w objective to image the activity of fluorescently-labeled motor neurons (*ZM9128 hpIs595[Pacr-2(s)::GCaMP6(f)::wCherry]*) at a 100-Hz acquisition rate, yielding 6000 light-fields in both green and red channels in a 1-minute observation (Supplementary Fig. 1). VCD-Net reconstructed the neuronal signals at single-cell resolution during fast body movement (Fig. 2b, Supplementary Fig. 12, Supplementary Video 1, 2). In contrast, LFD suffered from suboptimal cellular resolution and deteriorated image quality around the focal plane, as shown on the same image data (Fig. 2c). We identified A- and B- motor neurons that have been associated with motor-program selection^{15–17}, and mapped their calcium activity over time after ratiometrically correcting the calcium signals (GCaMP6(f)) using RFP baselines (wCherry) to remove motion noise¹⁸ (Fig. 2d, e, Supplementary Fig. 13, Supplementary Video 3). By applying an automatic segmentation of the worm body contours (Supplementary Note 3), we calculated the worm's velocity and curvatures related to its locomotion and behavior, thereby allowing classification of the worm motion into forward, backward, or irregular crawling (Fig. 2f–h). We found the patterns of transient calcium signaling to be relevant to the switches from forward to backward crawling of the worm, which is consistent with previous findings^{18–20}. Furthermore, the non-iterative VCD reconstruction could sequentially recover 3D images at a volume rate of 13.5 Hz, ~900-times faster than the iterative LFD method (Fig. 2i, Supplementary Table 3). Our VCD-LFM thus demonstrated advantages for visualizing sustained biological dynamics, which is computationally challenging using conventional deconvolution methods.

Volumetric imaging of fast dynamics in the beating zebrafish heart

We captured the cardiac hemodynamics in the beating zebrafish heart. To reduce the background from body tissue, we generated a rod-like laser beam to selectively illuminate the heart region²¹, and we recorded the light-field video at a 200-Hz volume rate using a 20 \times /0.5w objective (Fig. 3a, Supplementary Fig. 2). We reconstructed red blood cells (RBCs, labeled with *Tg(gata1a:dsRed)*) and beating cardiomyocytes (nuclei were labeled

with *Tg(my17:nls-gfp)* in four dimensions using VCD-Net with resolution, structural similarity, and processing throughput notably better than conventional LFD approaches (Fig. 3b–g, Supplementary Fig. 14–16, Supplementary Video 4, 5). We also show that direct 2D–3D VCD transformation is better than enhancing 3D LFD results by further processing with established deep-learning restoration methods²² in term of reconstruction accuracy and when dealing with densely labeled signals (Supplementary Note 4). We further demonstrated that our VCD-LFM could reconstruct a 3D beating heart with densely labeled trabecular myocardium (*Tg(cmlc2:gfp)*), Fig. 3h, i, Supplementary Fig. 17, Supplementary Video 6), and the limit of signal density that could be accurately recovered by VCD was relatively high, as compared to LFD (Supplementary Fig. 18). VCD-Net could also generalize well when trained on one type of cardiac sample, e.g. RBCs, and applied to another, e.g. cardiomyocyte nuclei, or trained on hybrid cardiac datasets including both and applied to each of them (Supplementary Note 5). VCD-LFM reconstruction with single-cell resolution (Supplementary Fig. 19) permitted quantitative investigation of transient cardiac hemodynamics. We tracked 19 individual RBCs throughout the entire cardiac cycle of 415 ms, during which the blood was pumped in and out of the ventricle at a speed of over 3000 $\mu\text{m s}^{-1}$ (Fig. 3d, e). By segmenting the heart boundary, we quantified the volume change of the myocardium during the diastole and systole, and calculated the ejection fraction of the heartbeat (Fig. 3j, k). Furthermore, through a combination of VCD-LFM and selective plane illumination microscopy (SPIM), we visualized in 3D and analyzed blood flow in the zebrafish circulation system (labeled with *Tg(fli1:gfp; gata1a:dsRed)* for blood cells in vessels, *Tg(gata1a:dsRed; cmlc2:gfp)* for blood cells in the heart), Supplementary Fig. 20, Supplementary Video 7–9).

Discussion

The VCD-LFM achieves real-time recording and video-rate reconstruction of instantaneous 3D processes in whole moving *C. elegans*. Combined with efficient network-based locomotion analysis, it offers an efficient pipeline for the study of sustained worm neural activity and related locomotion behavior at high throughput. The robust performance of VCD-LFM enables the suppression of signal artefacts, elimination of motion blurs, and accurate quantification of calcium signaling. For cardiovascular imaging, recent scanning-based approaches for volumetrically imaging in zebrafish larvae have required complicated optics, an ultra-fast camera, and high-intensity excitation^{23, 24}. In contrast, our method based on a relatively simple system and easily adoptable deep-learning framework offers a compelling solution for investigating the dynamic properties and functions of the cardiovascular system. Therefore, VCD-LFM could be a valuable tool for studying dynamics on fast timescales, potentially benefiting a variety of applications such as behaviorally relevant neuronal activity studies and dysfunctions of the heart and blood transport system in model organisms.

In summary, we introduced a VCD-LFM approach and demonstrated its ability to image transient biological dynamics with improved spatial resolution, minimal reconstruction artefacts, and increased reconstruction throughput compared to conventional LFD approaches. The network-based VCD computational model is robust, versatile, and ready for widespread application. While VCD-LFM improves the reconstruction quality from one

originally determined by the optical system to one can be optimized via the training procedure, it requires the pre-acquisition of a considerable number of training images. We expect this will improve with the continued development of deep learning technique, which aims for strong generalization ability and weak training supervision, thus allowing the model implementation with much fewer training data being required. Aside from the combination with a basic LFM setup^{7,13}, we note that VCD-Net is compatible with modified LFM modalities, such as a dual-objective setup¹¹ or a Fourier LFM setup¹⁴. Finally, we expect that VCD-LFM could potentially bring new insights for computational imaging techniques by raising the possibility of restoring image beyond the system optical limit rather than just approaching it, and showing the capability of increasing the image dimension while minimally compromising the image quality. Taken together, we can further push the spatiotemporal limits for *in toto* observation of dynamic biological processes.

Methods

Epi-illumination LFM setup.

An epi-fluorescence light-field setup was built on an upright microscope (BX51, Olympus). The light-field and wide-field detection paths were appended to the camera port of host microscope, with using a flip mirror to switch between two detection modes. A motorized z-stage (Z812B, Thorlabs) together with a water chamber were directly mounted onto the microscope stage (x-y), to three-dimensionally control the samples inside the chamber. A water immersion objective (LUMPlanFLN 40 \times /0.8w, Olympus) was used to collect the epifluorescence signals from samples. For recording the light field, a microlens array (MLA, APO-Q-P150-F3.5 (633), OKO Optics) was placed at the native image plane to collect the light-field signals. A 1:1 relay system (AF 60 mm 2.8D, Nikon) was used to conjugate the back focal plane of MLA with the camera sensor plane (Flash 4.0 V2, Hamamatsu). The light-field path was optionally extended to dual-channel detection by dividing after MLA and adding an extra camera sensor for the *C.elegans* experiments. See Supplementary Fig. 1 for more details of Epi-LFM setup.

Selective volume illumination LFM setup.

We also developed a LFM setup based on selective volume illumination. Two pairs of beam reducers combined with an adjustable iris were used to generate a scalable rod-like beam (473 or 532-nm), which was finally projected onto the sample through a 4 \times illumination objective (Plan Fluor 4 \times /0.13, Nikon) placed perpendicular to the detection path. It confined the fluorescence excitation within the heart region of zebrafish embryo, reducing the excessive emission from out of the volume of interest that could smear the desired signals. This selective volume illumination mode provided light-field image with less background noise and increased contrast²¹. For observing the dynamic process of blood flowing through vessels, we also integrated a standard SPIM channel (473-nm, 4- μ m-thickness laser-sheet) to implement the *in situ* 3D imaging of static vessels. The illumination paths were aligned, providing double excitation to the sample from its dual sides. The detection path used a water immersion objective (Fluor 20 \times /0.5w, Nikon) to collect the fluorescence signals. A dichroic mirror split the GFP (vessels) and DsRed (RBCs) signals for wide-field and light-field detection, respectively, when performing dual-channel imaging. The light-field

detection here followed the same design used in the epi-illumination LFM. See Supplementary Fig. 2 for more details of this hybrid setup.

VCD light-field reconstruction network.

In a general convolutional neural network (CNN), a certain M th convolutional layer receives feature maps from the previous $(N-1)$ th layer, and generates new feature maps using different convolution kernels. The network finally produces a multi-channel output, in which each channel is a non-linear combination of the original input. This concept shares similarities with the digital refocusing algorithms in the light-field photography, where each synthetic focal plane of the reconstructed volume can be interpreted as a superposition of different views extracted from the light-fields²⁵. Through cascaded layers, our model is expected to gradually transform the original angular information from the light-field raw image into depth features, eventually forming the conventional 3D image stack, and reconstructing the scene. In our implementation, the customized VCD-Net is based on a modified U-Net architecture that contains a downsampling path and a symmetric upsampling path²⁶. Along both paths, each layer has three parameters: \mathbf{n} , \mathbf{f} and \mathbf{s} , denoting the output channels number, the filter size of convolution kernel and the step size of the moving kernel, respectively, as specified in Supplementary Fig. 3 and Supplementary Table 1. The pixels from the input 2D light-field raw image (dimension: $\mathbf{a} \times \mathbf{b}$, height \times width) are first reformatted into a series of different views (dimension: $\mathbf{a}/\mathbf{d} \times \mathbf{b}/\mathbf{d} \times \mathbf{d}^2$, height \times width \times views) according to their relative positions to each lenslet. A subpixel up-scaling part further interpolates these views to dimension $\mathbf{a} \times \mathbf{b} \times \mathbf{d}^2$ (height \times width \times views). Then in the first VCD layer, the initial transformation from “view” to “channel” is done by convolutions with all these views using different kernels of each channel, generating an output with dimension $\mathbf{a} \times \mathbf{b} \times \mathbf{n}$, where \mathbf{n} is the channel numbers. The following convolution layers keep combining old channels from previous layers and generating new ones to excavate the hidden features from the input image. Local residual connections are integrated into the downsampling path to fully extract the hierarchical features. Finally, the last layer outputs a 3D image with the channel number \mathbf{n} equal to the desired number of synthetic focal planes \mathbf{c} , thereby finishing the transformation from “channel” to “depth” (dimension: $\mathbf{a} \times \mathbf{b} \times \mathbf{c}$, height \times width \times depth).

For VCD-Net training, HR 3D reference images were acquired from confocal microscopy of static samples or synthetic data. Then a light-field projection based on wave optics model¹³ was applied to these HR references to generate corresponding synthetic 2D light-field images as inputs (Supplementary Note 1). The VCD-Net was gradually trained via iteratively minimizing the difference between its intermediate outputs and the HR references. By setting appropriate loss functions, such as the MSE (Mean Square Error) of the pixel intensity, the VCD-Net obtained optimized kernel parameters for each layer and efficiently converged to a well-trained status, at which the network can transform the synthetic light-field inputs back to 3D images. At inference stage, the trained VCD-Net directly infers a sequence of 3D images from an input light-field video, which contains many light-field frames recording the dynamic biological processes. The time consumption for the VCD-Net procedure depends on the dataset size and computational resources. As a reference point, the VCD-Net converged after training on 4580 pairs of blood cell image

patches (size: $176 \times 176 \times 51$) pixels) with 110 epochs. The time cost was ~4 hours on a single GPU. Then the trained network spent ~15 seconds to reconstruct 450 consecutive volumes (size: $341 \times 341 \times 51$ pixels) from acquired light-field videos. This 4D reconstruction throughput was compared to ~11.8 hours (~42467 s) by running LFD (8 iterations) on the same workstation. The computation was performed on a workstation equipped with Intel(R) Core (TM) i9-7900X CPU @3.3GHz, 128G RAM, and Nvidia GeForce RTX 2080 Ti graphic cards. For more details about VCD-Net training and inference, see Supplementary Note 2 and our open-source code. The comparative performance of our direct 2D-3D VCD-Net recovery and LFD recovery (2D-3D) plus other deep-learning image restoration (3D-better 3D) is shown in Supplementary Note 4. The generalization ability of VCD-Net, reflected by the hybrid-sample and cross-sample learning application, is discussed in Supplementary Note 5.

PSF measurements.

We used both light-field ($40\times/0.8w$ objective) and 3D wide-field ($40\times/0.8w$ objective plus $0.5\text{-}\mu\text{m}$ z-step) modes to image the same volume of sub-diffraction fluorescent beads ($0.5\ \mu\text{m}$ Lumisphere, BaseLine) distributed in a piece of hydrogel (0.7% low melting agarose solution, BBI Life Sciences). To demonstrate the tunable performance of VCD-Net with applying different training data, we trained anisotropic and isotropic VCD-Nets using synthetic anisotropic and isotropic beads, respectively. These synthetic beads with random 3D distributions were generated using a 3D Gaussian kernel with controllable FWHMs (anisotropic: $1 \times 1 \times 3\ \mu\text{m}$; isotropic: $1 \times 1 \times 1\ \mu\text{m}$). Then 3D image stacks with $60\text{-}\mu\text{m}$ depth were recovered from the recorded light-field image using LFD, anisotropic and isotropic VCD-Nets. The reconstruction yielded 61 z planes with voxel size $0.34 \times 0.34 \times 1\ \mu\text{m}$ for anisotropic VCD-Net and LFD (8 iterations), and 177 z planes, with voxel size $0.34 \times 0.34 \times 0.34\ \mu\text{m}$ for isotropic VCD-Net. The reconstructed beads were then detected and fitted with 1D Gaussian function in each dimension to determine the full width at half maximum (FWHM) as spatial resolution metrics using a custom-written MATLAB script. The PSFs of LFD and VCD-Nets at certain depth were then compared via plotting the line profiles of the same resolved beads. The achieved lateral and axial resolutions at certain depth by each method were indicated through calculating the averaged FWHM values of the resolved beads, e.g., $z = -14\ \mu\text{m}$. The performance of anisotropic and isotropic VCD-Net *versus* LFD at different depths were further analyzed via measuring the FWHMs of resolved beads across a $60\text{-}\mu\text{m}$ depth, with their variation indicating the non-uniformity of light-field recovery.

C. elegans strain.

The strain ZM9128 *hpIs595[Pacr-2(s)::GCaMP6(f)::wCherry]*, expressing GCaMP6f in A- and B- class motor neurons, was used to detect neuronal activity in the moving worm (Fig. 2). The strain QW1217: *hpIs491[Prgef-1::GCaMP6(f)::wCherry]; hpIs467[Prab-3::NLS::RFP]*, with nuclei labelled in all neurons, was used to demonstrate the imaging performance of VCD-LFM (Supplementary Fig. 12). All *C. elegans* were cultured on standard Nematode Growth Medium (NGM) plates seeded with OP50 and maintained at $22\ ^\circ\text{C}$ incubators until L4 stage.

C. *elegans* imaging.

To obtain HR 3D data for network training, the L4-stage anesthetized worms (QW1217 *hpIs467*, ZM9128 *hpIs595*, by 2.5 mM levamisole in M9 buffer, Sigma-Aldrich) were first imaged using a 40×/0.95 objective on a confocal microscope (FV3000, Olympus). To demonstrate the performance of VCD-LFM, anesthetized worms (QW1217 *hpIs467*) embedded in agarose were *in situ* imaged by both light-field and wide-field detection modes in our Epi-illumination LFM. The acquired 3D wide-field image stacks (1- μ m step size) were deconvolved in Amira software (Thermo Scientific), for comparison with light-field reconstructions by LFD and VCD-Net. The 3D reconstructions encompassed 31 z planes, with voxel size of $0.34 \times 0.34 \times 1 \mu\text{m}$. In the behavior studies of *C. elegans*, the awake L4-stage worms (ZM9128 *hpIs595*) were loaded into a microfluidic chamber ($\sim 300 \times 300 \times 50 \mu\text{m}$), which allows the worms to move within the FOV of a 40× objective. The GCaMP and RFP signals of the moving worm were then recorded for one minute at 100 Hz frame rate (2-ms exposure time).

Quantitative Analysis of neuronal activity and behavior of moving *C. elegans*.

We performed semi-automated tracking of the movement and intensity fluctuation of each individual neuron using TrackMate Fiji Plugin²⁷. Neurons were detected automatically in each volume by applying a circular ROI through the Difference of Gaussian (DoG) detector and then tracked using Kalman filter. In cases when this automatic tracking failed due to fast movement of the neuron, the missing detections and tracking mistakes were manually corrected. For each neuron in each volume of GCaMP or RFP channel, all the pixels within the ROI of a certain neuron were averaged to generate a single value F_g or F_r representing the fluorescence intensity of this neuron in each channel. To extract calcium dynamics, we measured $(F - F_0)/F_0$, where $F = F_g / F_r$, is the ratio of GCaMP fluorescence F_g to RFP fluorescence F_r , and F_0 is the neuron-specific baseline being the average of the lowest 100 values. Worm behavior analysis was then implemented based on the same fluorescence images. We developed an efficient worm analysis pipeline which could 1. rapidly infer the worm outlines throughout the whole time period using a U-Net based image segmentation²⁶; 2. extract the center lines from the segmented outlines to calculate the changing curvatures and motion velocities²⁸. More details are given in Supplementary Note 3. The body curvature map shown in Fig. 2f indicates the time-varying worm postures. The velocity curve shown in Fig. 2g represents the movement along the worm body.

Fish husbandry and lines.

Transgenic zebrafish lines *Tg(gata1a:dsRed; cmlc2:gfp)* (Fig. 3, Supplementary Fig. 14, Video 4, 9), Transgenic zebrafish lines *Tg(cmlc2:gfp)* (Fig. 3, Supplementary Fig. 17, Video 6), *Tg(fli1:gfp; gata1a:dsRed)* (Supplementary Fig. 14, 20, Video 7, 8), *Tg(my17:nls-gfp)* (Supplementary Fig. 15, Video 5) were used in our experiments. Embryonic fishes were maintained until 3–4 dpf in standard E3 medium, which was supplemented with extra PTU (Sigma Aldrich, MO) to inhibit melanogenesis. Then, the larvae were anesthetized with tricaine (3-amino benzoic acid ethylester, Sigma Aldrich, MO) and immobilized in 1% low-melting-point agarose inside FEP (Fluorinated Ethylene Propylene) tube for further imaging. In the imaging of cardiac blood flow, the embryos were injected with *gata2a* morpholino

oligonucleotide at the single-cell stage, to slow down hematopoiesis, and thereby reduce the density of RBCs. All the experiments were performed in compliance with and with the approval of a UCLA IACUC protocol.

Zebrafish imaging.

The fish samples were moved by a custom stage (xyz + rotation) so that their signals of interest could be positioned within the optimal imaging region, determined by the size of selective volume illumination and the FOV of the detection objective (LUMPlanFLN 20×/0.5w, Olympus). During cardiac imaging, the center of the heart was moved to the focal plane so that the volumetric illumination (473 or 532-nm) could selectively excite the RBCs, myocytes nuclei or myocardium. The light-field movies were recorded at 200 Hz frame rate, with 768 × 768 frame size that corresponds to a lateral FOV of ~250 × 250 μm. The movies covered 4–5 cardiac cycles and contained 450 frames (5 ms exposure for each frame). For imaging the dynamic process of blood flow inside vessels, the cameras of dsRed and GFP channels recorded a light-field video of rapidly flowing RBCs and a SPIM image stack of static tail vessels, respectively. Each light-field movie contained 600 frames (10 or 5 ms exposure). The SPIM stack contained 50 planes that covered a 100-μm z depth of the fish tail (step size 2 μm).

To obtain HR 3D images of RBCs, myocyte nuclei and myocardium for VCD-Net training, deeply anesthetized fish larvae with immobilized hearts were embedded in 1% low-melting agarose for sustained confocal imaging (SP8-STED/FLIM/FCS, Leica) using a 20×/0.75 Objective (HC PL APO CS2). Then we trained three VCD-Nets based on these HR images of static blood cells (16 fishes), nuclei (23 fishes), and myocardium (11 fishes), and their paired LFPs. Besides the confocal imaging of immobilized heart, we also combined SPIM with a retrospective gating method^{23, 29, 30}, to obtain the light-sheet based HR 3D image of periodically beating myocardium for network training. A movie stack of light-sheet images of the dense myocardium was first acquired, containing 60–70 movies at different planes (z step = 2 μm) that captured the complete myocardial structures. Each movie was then temporally aligned using retrospective gating, to reconstruct a beating 3D heart. When the training using these prior data was finished, the network inferred 3D images from empirical light-fields with output voxel size 0.68 × 0.68 × 2 μm (blood cell, myocardium) or 0.326 × 0.326 × 3 μm (myocyte nucleus). The output image stacks contain 51 planes with depth spanning from –50 to 50 μm or –75 to 75 μm.

Resolution quantification of VCD-Net and LFD reconstructions.

We applied decorrelation analysis³¹ to quantify the resolutions of VCD-Net and LFD reconstructions. The analysis was conducted using the ImageJ plugin (image decorrelation analysis). We set recommended parameters (radius, the range of normalized frequencies, 0–1; Nr, the number of sampling points, 50; Ng, the number of intermediate high-pass filtering used to find the resolution, 10) to calculate the cut-off frequency k_c . Supplementary Figs. 15–17 show higher k_c obtained in VCD-Net results than LFD results, which quantitatively validate the better resolution achieved by VCD-Net. We also used the Fourier spectrograms (by ImageJ plugin), to further confirm the resolution advantage of VCD-Net.

Velocity map of RBCs and volume-based ejection fraction analysis of myocardium in zebrafish heart.

Tracking of the flowing RBCs was performed using Imaris (Bitplane). Fig. 3d shows the trajectories of 19 RBCs throughout one cardiac cycle (415 ms). Fig. 3e demonstrates the velocity map that has been extracted at one specific time point during systole, which contains 31 vectors from all trackable RBCs in that frame. The analysis was written in Matlab and the vector field was visualized by Mayavi³². Volumetric segmentation of beating myocardium was performed using Amira. A blow tool was used to allow semiautomatic definition of the inner boundary of the ventricle at each slice. Once all the slices in certain 3D image stack of one time point were correctly segmented, the software could calculate the volume of segmented ventricle based on the slice thickness and the defined area. After the volume of the beating heart at all time points was calculated, we obtained the volume change ratio of the ventricle during the diastole and systole in one cardiac cycle.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank J.-N. Chen and Y. Dong at UCLA for their help with zebrafish experiments. We thank Y. Huang at Peking University and J. Wang at Tsinghua University for their discussions and comments on the manuscript. We thank L. Xiao for his help on the code implementation. We acknowledge the David Geffen School of Medicine at UCLA, for providing the Zebrafish Core Facilities and laboratory space. This work was supported by the following grants: National Natural Science Foundation of China (grant nos. 21874052 and 21927802 to P.F., 31671052 and 32020103007 to S.G.), National Key R&D program of China (grant no. 2017YFA0700501 to P.F.), Innovation Fund of WNLO (P.F.), National Science Foundation of Hubei (grant no. 2018CFA039 to S.G.), National Institutes of Health (NIH) (grant nos. R01 HL149808, R01 HL118650, R01 HL129727 and R01 HL111437 to T.K.H. and K99 HL148493 to Y.D.), Department of Veterans Affairs (grant no. I01 BX004356 to T.K.H.), the Fundamental Research Funds for the Central Universities (grant no. 2020JYCXJJ076 to L.Z.), and the Junior Thousand Talents Program of China (P.F. and S.G.). The confocal experiment was done at the Advanced Light Microscopy/Spectroscopy Laboratory and the Leica Microsystems Center of Excellence at the California Nano Systems Institute at UCLA.

References

1. Keller PJ, Schmidt AD, Wittbrodt J & Stelzer EHK Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy. *Science* 322, 1065–1069 (2008). [PubMed: 18845710]
2. Planchon TA et al. Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. *Nat. Methods* 8, 417 (2011). [PubMed: 21378978]
3. Conchello J-A & Lichtman JW Optical sectioning microscopy. *Nat. Methods* 2, 920 (2005). [PubMed: 16299477]
4. Huisken J, Swoger J, Del BF, Wittbrodt J & Stelzer EH Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* 305, 1007–1009 (2004). [PubMed: 15310904]
5. Yang W & Yuste R. In vivo imaging of neural activity. *Nat. Methods* 14, 349 (2017). [PubMed: 28362436]

6. Levoy M, Ng R, Adams A, Footer M & Horowitz M. Light field microscopy. *ACM Trans. Graph* 25, 924–934 (2006).
7. Prevedel R. et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nat. Methods* 11, 727–730 (2014). [PubMed: 24836920]
8. Pégard NC et al. Compressive light-field microscopy for 3D neural activity recording. *Optica* 3, 517 (2016).
9. Cong L. et al. Rapid whole brain imaging of neural activity in freely behaving larval zebrafish (*Danio rerio*). *Elife* 6, e28158 (2017).
10. Nobauer T. et al. Video rate volumetric Ca(2+) imaging across cortex using seeded iterative demixing (SID) microscopy. *Nat. Methods* 14, 811–818 (2017). [PubMed: 28650477]
11. Wagner N. et al. Instantaneous isotropic volumetric imaging of fast biological processes. *Nat. Methods* 16, 497–500 (2019). [PubMed: 31036959]
12. Cohen N. et al. Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express* 22, 24817–24839 (2014).
13. Broxton M. et al. Wave optics theory and 3-D deconvolution for the light field microscope. *Opt. Express* 21, 25418–25439 (2013).
14. Guo C, Liu W, Hua X, Li H & Jia S. Fourier light-field microscopy. *Opt. Express* 27, 25573–25594 (2019).
15. Gao S. et al. Excitatory motor neurons are local oscillators for backward locomotion. *Elife* 7, e29915 (2018).
16. Wen Q, Gao S & Zhen M. *Caenorhabditis elegans* excitatory ventral cord motor neurons derive rhythm for body undulation. *Philos. Trans. R. Soc. B* 373, 20170370 (2018).
17. Xu T. et al. Descending pathway facilitates undulatory wave propagation in *Caenorhabditis elegans* through gap junctions. *Proc. Natl Acad. Sci. USA* 115, E4493-E4502 (2018).
18. Nguyen JP et al. Whole-brain calcium imaging with cellular resolution in freely behaving *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* 113, E1074-E1081 (2016).
19. Venkatachalam V. et al. Pan-neuronal imaging in roaming *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* 113, E1082-E1088 (2016).
20. Kawano T. et al. An imbalancing act: gap junctions reduce the backward motor circuit activity to bias *C. elegans* for forward locomotion. *Neuron* 72, 572–586 (2011). [PubMed: 22099460]
21. Truong TV et al. High-contrast, synchronous volumetric imaging with selective volume illumination microscopy. *Commun. Biol* 3, 1–8 (2020). [PubMed: 31925316]
22. Weigert M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* 15, 1090–1097 (2018). [PubMed: 30478326]
23. Mickoleit M. et al. High-resolution reconstruction of the beating zebrafish heart. *Nat. Methods* 11, 919 (2014). [PubMed: 25042787]
24. Voleti V. et al. Real-time volumetric microscopy of in vivo dynamics and large-scale samples with SCAPE 2.0. *Nat. Methods* 16, 1054–1062 (2019). [PubMed: 31562489]
25. Ng Ren, et al. Light field photography with a hand-held plenoptic camera. *Diss. Stanford University* (2005).
26. Falk T. et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* 16, 67 (2019). [PubMed: 30559429]
27. Tinevez JY et al. TrackMate: An open and extensible platform for single-particle tracking. *Methods* 115, 80–90 (2017). [PubMed: 27713081]
28. Restif C. et al. CeleST: computer vision software for quantitative analysis of *C. elegans* swim behavior reveals novel features of locomotion. *PLoS Comput. Biol* 10, e1003702 (2014).
29. Lee J. et al. 4-Dimensional light-sheet microscopy to elucidate shear stress modulation of cardiac trabeculation. *J.Clin. Investig* 126, 1679–1690 (2016). [PubMed: 27018592]
30. Liebling M, Forouhar AS, Gharib M, Fraser SE & Dickinson ME Four-dimensional cardiac imaging in living embryos via postacquisition synchronization of nongated slice sequences. *J. Biomed. Opt* 10, 054001 (2005).
31. Descloux AC, Grussmayer KS & Radenovic A. Parameter-free image resolution estimation based on decorrelation analysis. *Nat. Methods* 16, 918–924 (2019). [PubMed: 31451766]

32. Ramachandran P & Varoquaux G. Mayavi: 3D visualization of scientific data. *Comput. Sci. Eng* 13, 40–51 (2011).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

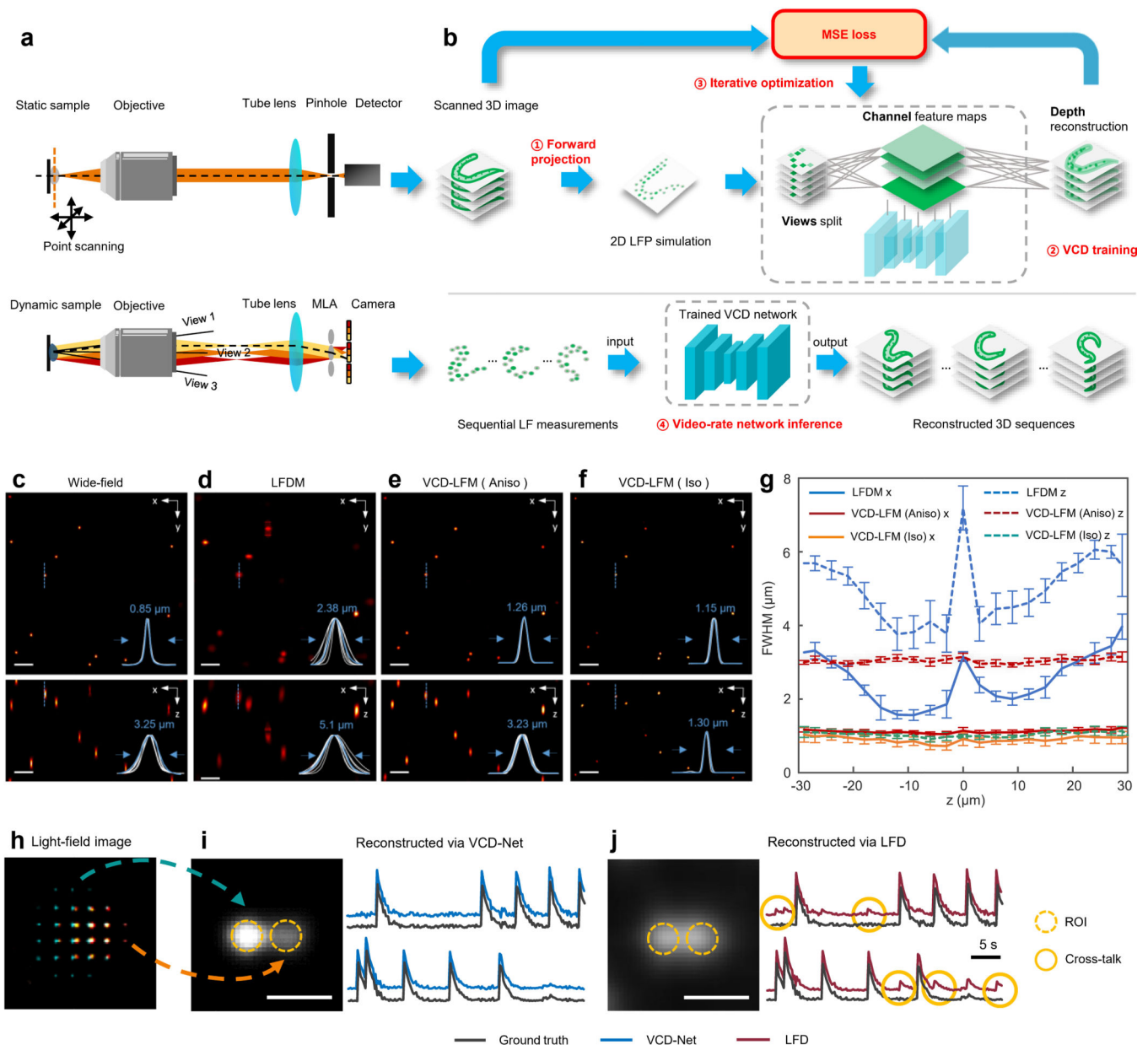


Figure 1. VCD-LFM and its performance.

(a) HR 3D imaging of stationary samples by a confocal microscope (upper row) and instantaneous recording of dynamic samples by a light-field microscope (bottom row). (b) The VCD-Net reconstruction pipeline, containing: 1. Forward light-field projection (LFP) from the HR image stacks; 2. VCD transformation of synthetic light-field inputs into intermediate 3D image stacks; 3. Network training via iteratively minimizing the difference between VCD inferences and confocal ground truths; 4. Inference of 3D images from the recorded light-field images by a trained VCD-Net. (c)–(f) Maximum intensity projections (MIPs) of the same fluorescent beads and achieved resolution (FWHM) by wide-field microscopy (c), LFDM (d), and VCD-LFM trained with anisotropic (e) and isotropic (f) HR

data, respectively. White lines, intensity profiles of all the resolved beads shown in the MIPs. Blue lines, intensity profiles across a selected bead (indicated by a vertical line) at 20 μm off the focal plane. Scale bars, 10 μm . **(g)** Average axial (dashed lines) and lateral (solid lines) FWHM of the beads across the volumes reconstructed by LFD (n = 2,039 beads), anisotropic (n = 2,527 beads) and isotropic VCD-LFM (n = 2,731 beads), respectively. Aniso-VCD achieves uniform resolution of 1.1 ± 0.08 (mean \pm standard deviation) μm and 3.0 ± 0.1 μm in x/y and z, respectively, and Iso-VCD achieves near isotropic resolution of 1.0 ± 0.15 μm . Center lines represent means and error bars denote standard deviations. **(h)** One frame of a light-field video recording the activity of two synthetic firing neurons adjacent to each other (indicated by blue and red colors). **(i),(j)** Reconstructions of the light-field video by VCD-Net (i) and LFD (j), with signal traces extracted in the indicated ROIs (dashed circles), and compared to the ground truth. Circles denote the signal cross-talks between adjacent neurons due to the blurring. Scale bar, 5 μm .

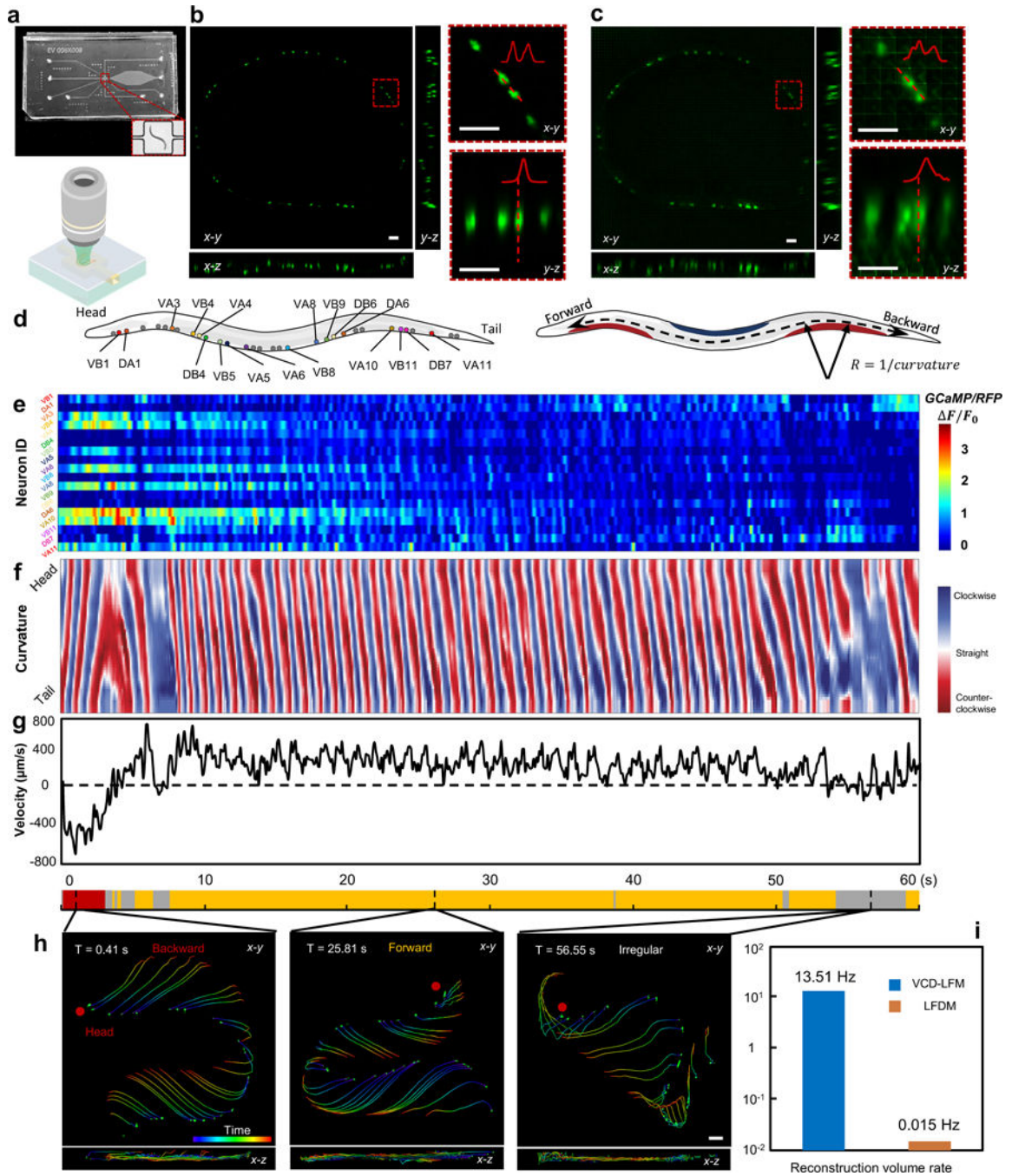


Figure 2.

Whole-animal Ca^{2+} imaging of moving *C. elegans* using VCD-LFM. (a) Configuration combining light field microscopy with a microfluidic technique for imaging motor neuron activity in L4-stage *C. elegans* (strain ZM9128 *hplIs.595[Pacr-2(s)::GCaMP6(f)::wCherry]*) moving in a microfluidic chip ($300 \times 300 \times 50 \mu\text{m}$, top panel) at 100-Hz recording rate. (b), (c) MIPs as well as magnified views of the indicated regions of one instantaneous volume reconstructed by VCD (b) and LFD (c), respectively. The data shown are representative of $n = 10$ independent *C. elegans*. Scale bars, $10 \mu\text{m}$. (d) Schematics of the worm with identified

motor-neurons labeled (left) and body curvature annotated (right). **(e)** Heatmap visualizing the neuronal activity of 18 identified motor neurons during a 1-minute observation of the moving worm. Each row shows the Ca^{2+} signal fluctuation of an individual neuron with color indicating the percent fluorescence changes ($\Delta F/F_0$), where F is ratiometrically corrected by the ratio of GCaMP6(f) fluorescence to wCherry fluorescence. **(f)** Curvature kymograms along the body of the moving worm. **(g)** Velocity plot shows the displacement in the direction of body. An ethogram describing the worm behavior over time (lower panel) is obtained by analyzing the curvature and velocity change. **(h)** Selected volumes with time-coded traces (duration: left and middle, 150 ms; right, 500 ms) in accordance with the ethogram visualizing the backward (left), forward (middle), and irregular (right) crawling tendency of the worm. Scale bars, 20 μm . **(i)** The reconstruction throughput of VCD-LFM and LFDM for processing the same *C.elegans* light-field video.

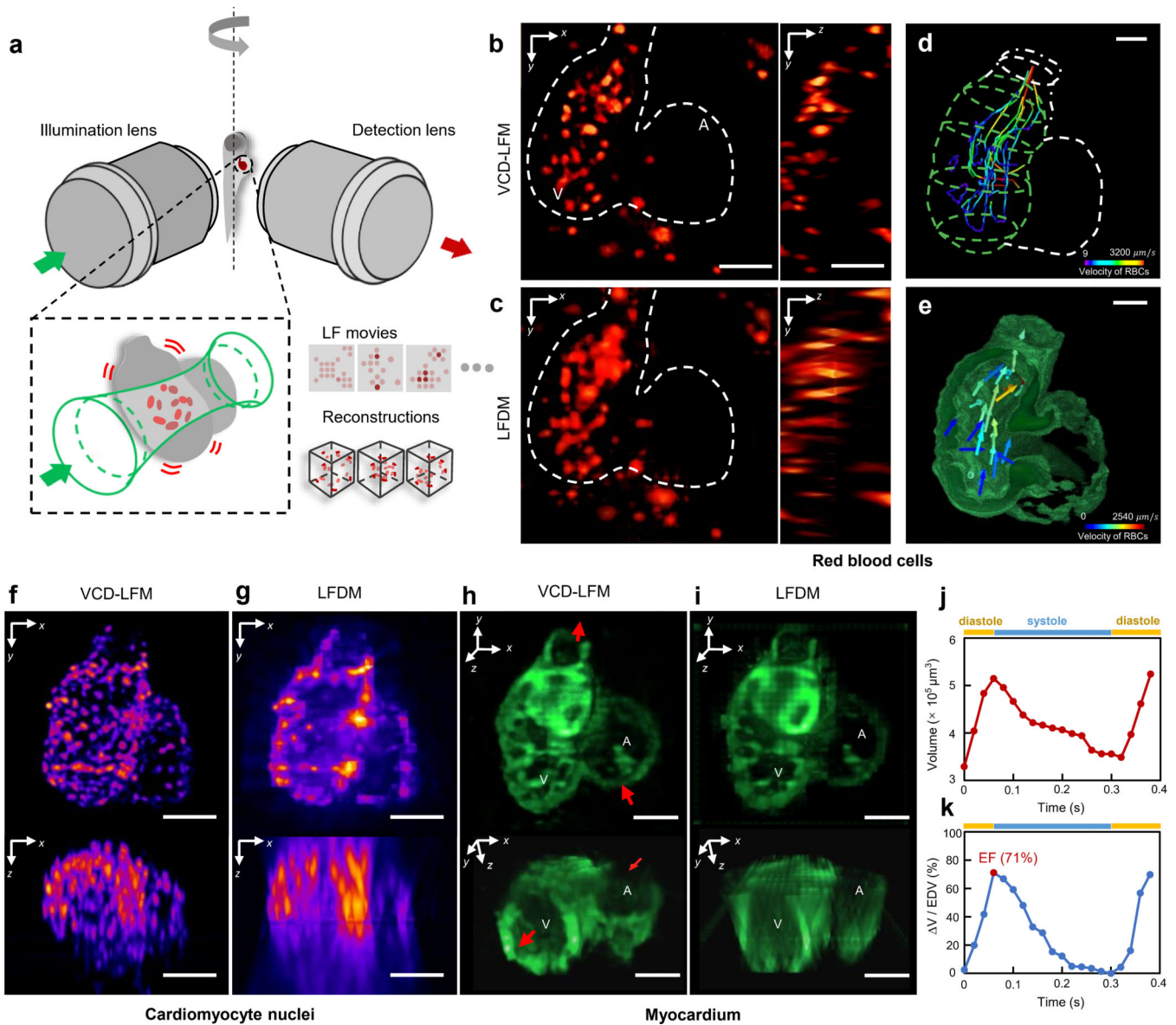


Figure 3. Imaging of various cardiac dynamics in beating zebrafish heart using VCD-LFM. (a) Schematic of selective volume illumination based light-field imaging for zebrafish experiments. High-contrast light-field sequences of RBCs in the beating heart are recorded under a rod-like beam illumination setup that selectively excites the fluorescence signals within the volumes of interest. (b),(c) MIPs in x - y (left) and y - z (right) planes of one instantaneous volume of RBCs by VCD-LFM (b) and LFDM (c), respectively. The dash lines indicate the heart. (d) Tracks of 19 single RBCs throughout the cardiac cycle. A static heart is outlined for reference. (e) Velocity map computed from two consecutive volumes of RBCs during systole. (f),(g) MIPs in x - y (top) and x - z (bottom) planes of one instantaneous volume of nuclei of beating cardiomyocytes by VCD-LFM (f) and LFDM (g), respectively. (h),(i) 3D visualization of beating myocardium at one time point by VCD-LFM (h) and LFDM (i), respectively. The myocardium was labeled by GFP. Arrows indicate the inlet and outlet of cardiac pumping. A: atrium; V: ventricle. (j) Volume of the myocardium during the

diastole and systole in one cardiac cycle. **(k)** Volume change ratio of the ventricle during the diastole and systole in one cardiac cycle. The rate is calculated by $(V_{time} - ESV) / EDV$, where V_{time} is the time-varying volume of the ventricle during heartbeat and ESV and EDV represent the volumes at the end of systole and diastole, respectively. The ejection fraction of the heartbeat (EF), given by $(EDV - ESV) / EDV$, also shown as the peak in the curve, is calculated to be ~71%. Twenty out of a total of 120 time points (1 of each 6) are selected for the analysis during a ~400 ms cardiac cycle. Scale bars, 50 μm . The data shown are representative of $n = 8, 4, 3$ independent fish for blood cells, cardiomyocyte nuclei and myocardium imaging, respectively.