Network-based instantaneous recording and video-rate reconstruction of 4D biological dynamics

Zhaoqiang Wang[1,4+], Hao Zhang[1+], Lanxin Zhu[1+], Guo Li[1], Yi Li[1,3], Yicong Yang[1], Mehrdad Roustaei[4], Shangbang Gao[3], Tzung K. Hsiai[4,5*] and Peng Fei[1,2*]*

[1] *School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, 430074, China.*

[2] *Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, 430074, China.*

[3] *College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China.*

[4] *Department of Bioengineering, University of California, Los Angeles, Los Angeles, 90095, U.S.A.*

[5] *School of Medicine, University of California, Los Angeles, Los Angeles, 90095, U.S.A*

[+] *These authors contribute equally to this work*

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

Abstract-The artefacts and non-uniform resolution accompanied with slow reconstruction speed in light-field microscopy compromises its full capability for *intoto* observing fast biological dynamics. Here we demonstrate that combining a view-channel-depth (VCD) neural network with light-field microscopy can mitigate these limitations, yielding artefact-free 3D image sequences with uniform spatial resolution and three-orders higher, video-rate reconstruction throughput. We image neuronal activities across moving *C. elegans* and pumping blood flow in beating zebrafish heart at single-cell resolution and volume rate up to 200 Hz.

A recurring challenge in biology is the quest to extract ever more spatiotemporal information from the targets since many millisecond- transient cellular processes occur in three-dimensional tissues and across long time scale. Several imaging techniques, including epifluorescence and plane illumination methods, can image live samples in three dimensions at high spatial resolution[1-4]. Meanwhile, they need to record a number of 2D images to comprise a 3D volume, in which the transient temporal profiles are compromised due to extended acquisition time.

The recent advent of Light field microscopy (LFM) has become a unique technique of choice for rapid and instantaneous volumetric imaging[5-9]. It particularly permits the retrieval of transient 3D signal distribution through post-processing of 4D light field recorded by single camera snapshot. As LFM provides high-speed volumetric imaging only limited by the camera frame rate, it has delivered promising results for various applications, such as recording of neuronal activities and visualization of cardiac dynamics in model organisms[9-11]. Despite this advancement, its generally low spatial resolution, especially non-uniform axial resolution, and presence of reconstruction artefacts by traditional light-field recovery methods yet prevents its more widespread applications to capture millisecond time-scale biological processes at cellular resolution. Though these problems can be mitigated through optimizing the way light field being recorded[9,12], the extra system complexity could impede the wide dissemination of LFM technique. Furthermore, current LFMs heavily rely on computationally demanding, iterative recovery process that intrinsically limits the overall throughput of LFM reconstruction, compromising its potentials for long time-scale applications.

Here we propose a novel LFM strategy based on a View-Channel-Depth neural network processing of light field data[13], termed VCD-LFM. By developing a light-field projection based on wave optics model, we generated plenty of synthetic light-field images from high-resolution 3D images experimentally obtained beforehand, and readily paired them as target and ground-truth data, respectively, for network training. The VCD network (VCD-Net) procedure was designed to enable the extraction of multiple views from these 2D light fields, and the involvement of multi-channel architecture in convolutional network to transform the views back to 3D depth images, which would be compared with the high-resolution ground truths to guide the optimization of 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 capable of deducing 3D, high-fidelity signals at uniform resolution across depth. In addition to higher resolution and less artefact provided, once the VCD-Net being properly trained, it could instantly deduce an image stack from a light-field measurement in milliseconds time scale, showing ultra-high throughput suited for time-lapsed video processing. We demonstrated the ability of VCD-LFM method via imaging the motor neurons activities of L4-stage *C. elegan*s rapidly moving inside a 300 μm × 300 μm × 50 μm microfluidic chamber, at acquisition rate of 100 Hz and processing throughput of 16 volumes per second. This allowed us to extract the spatiotemporal patterns of neuronal calcium signaling and track correlated worm behaviors at single-cell resolution, notably better than conventional light-field deconvolution microscopy (LFDM). Furthermore, we combined VCD-LFM with light-sheet microscopy to *intoto* image blood flow in beating heart of zebrafish larvae at 4 days post fertilization (dpf) with 3D spatial, temporal and spectral (termed 5D) information resolved, thereby enabling the velocity tracking of single blood cells and synchronized displacement tracking of beating cardiomyocytes across a 250 μm × 250 μm × 150 μm space with volume rate 200 Hz.

Our VCD-LFM contains the training of VCD convolutional neural network and its inference based on the images obtained by LFMs (**Fig. 1a, Methods, Supplementary Fig. 1,2**). To create the data for network training, we first obtained a number of high- resolution (HR) 3D images of stationary samples using synthetic or experimental method (**Fig. 1a, Methods**). Referring to wave optic model of LFM[8], we projected these HR 3D images into 2D light-field images, which were used as the input of the network training (**Fig. 1b, step 1, Supplementary Fig. 3**). Each synthetic light-field image was first in principle re-arranged into different views, from which features were extracted and incorporated into multiple channels in each convolutional layer. The output features of each channel were then assigned into a number of planes representing different depths to generate an image stack. With repetitively extracting the features by using several convolution and deconvolution layers (U-net architecture, Supplementary Fig. 4), this view-channel-depth procedure generated intermediate 3D reconstructions (**step 2**). Pixel-wise mean-square-error (MSE) between them and HR ground truths were counted as the loss function that indicated how different these outputs were as compared to the ground truths. By iteratively minimizing the loss function (**step 3**), the network was gradually optimized till it could transform the synthetic light fields into 3D images similar enough to the ground truths **(Supplementary Fig. 5**). Based on the learning from gigavoxels prior data, the trained network was capable of instantly implementing VCD transformation of sequential light-field measurements that recorded the dynamic process and inferred a sequence of 3D volumes at a rate up to 16 volumes s^{-1} (**step 4**)

Experimentally, we built a microscope to enable *insitu* light-field recording and 3D wide-field imaging (**Methods, Supplementary Fig. 1**). To demonstrate the capability of VCD-LFM procedure, we reconstructed

sub-diffraction beads captured using a 40×/0.8w objective and quantified the resolution improvement owing to the network capability via comparing the results with those by conventional light-field deconvolution (LFD, **Fig. 1c**). As verified by 3D wide-field imaging of the same volume, the fluorescence of individual beads was correctly localized throughout the volume by VCD reconstruction (**Fig. 1c, Supplementary Fig. 6**). The averaged lateral and axial extents of beads at 11 μm off the focal plane resolved by each method are plotted in Fig. 1d. VCD-LFM yields averaged x,y~1.3± 0.1 and z~3.5± 0.4 μm resolutions (n = 15 beads) which are also more uniform across 90-μm imaging depth (**Fig. 1e, Supplementary Fig. 7**), showing significant improvement as compared to x,y~2.1 and z~5.1 μm of the LFDM results (~2.7 and 7.1 μm at the outer edge of the axial FOV). Furthermore, VCD-LFM substantially wipes off the mosaic-like artefacts near the focal plane that are common in LFDM method (**Supplementary Fig. 6,7**). It's also noted that the performance of VCD-Net is highly training data-oriented and less coupled with the optical limit of LFM system, indicating its reconstruction quality can be further improved by including higher-resolution prior data (**Supplementary Fig. 8**).

We demonstrate VCD-LFM suited for capturing highly dynamic cellular processes of live animals via imaging neuronal activities of moving *C. elegans* and beating heart of zebrafish larva. A microfluidic chip was used to permit *C.elegans* (L4-stage) rapidly moving inside a micro-chamber (300 × 300× 50 μm, **Fig. 2a**). We used 40×/0.8w objective to image the activities of motor neurons labelled with GFP (*Pacr-2(s)*::GCaMP6(f)) at 100 Hz acquisition rate, yielding 12000 light fields in 2 minutes- observation. The VCD-Net provided high-quality and quantitatively accurate reconstruction to visualize the neuron calcium signaling during the fast body movement at single-cell resolution (**Fig. 2b, Supplementary Fig. 9,10**). In contrast, LFD suffered from ambiguous cellular resolution and deteriorated image quality around focal plane, a known limitation of LFD (**Fig. 2c**). Furthermore, the non-iterative VCD reconstruction could sequentially recover 3D images at a volume rate of 16.5 Hz, ~1000-times faster than iterative LFD method (**Fig. 2d**). Our VCD-LFM thus shows significant advantage for visualizing sustained biological dynamics which is computationally challenging to conventional methods (**Supplementary Table 1, Video 1**). With applying an automatic segmentation of the worm body contours based on the location and morphology of neuron signals (**Methods, Supplementary Fig. 11**), we calculated the worm's velocity and curvatures related with the locomotion and animal behavior, thereby classifying the worm motion into forward, backward and irregular crawling (**Fig. 2e, f**). Meanwhile, we identified specific A- and B- motor neurons that have been associated with motor-program selection, and mapped their calcium activities over time (**Fig. 2g, Methods, Supplementary Video 2**). The patterns of transient Ca^{2+} signaling was found relevant to the switches from forward to backward crawling of the worm, consistent with the previous findings[14, 15]. (**Fig. 2g-i**).

We also captured the red fluorescence protein (RFP)-tagged red blood cells (RBCs) flowing inside a beating zebrafish heart expressing highly dense GFP at cardiomyocytes. This dual-color instantaneous 3D imaging was realized by a selective volume illumination-based LFM setup combined with a time-stamp post-synchronization (**Methods, Supplementary Fig. 2**). To reduce the background, a ~250-μm diameter laser-rod (532-nm) was generated to selectively illuminate blood cells in the heart region, recording light-field video at 200 Hz volume rate by a 20×/0.5w objective. A 488-nm SPIM[3] channel simultaneously imaged heartbeat (4-5 cycles) at different depths (2-μm z-step), followed by a retrospective gating algorithm to volumetrically reconstruct the periodically beating cardiomyocytes (**Fig. 3a**). By using a pair of light-sheet/light-field images at certain plane as time-stamp (**Methods**), we spatiotemporally aligned VCD-

121 reconstructed blood flow with retrospectively-gated heartbeat, finally forming a 5D image capable of *intoto*
122 visualizing the rapid cardiac hemodynamics at single-cell resolution (**Fig. 3b-d, Supplementary Fig. 12,**
123 **Video 3**). Single-color beating cardiomyocytes nuclei and dual-color hemodynamics at dorsal vein were also
124 visualized using our method (**Supplementary Fig.13, Video 4,5**). We manually tracked 14 individual RBCs
125 within the three dimensions of ventricle geometry throughout the entire cardiac cycle of 445 milliseconds,
126 owing to the 5D VCD-LFM reconstruction (**Fig. 3e**). We also computed the 3D velocity distribution of the
127 entire blood flow and the simultaneous movement of ventricle wall between two consecutive volumes (10-
128 ms interval) (**Fig. 3f,g**). These permitted the quantitative investigation of a transient cardiac hemodynamics,
129 in which the blood flow was pumped in-and-out the ventricle at high speed over 3000 $\mu m\ s^{-1}$ near the AV
130 valve during the diastole and systole of the heart. We note that visualization of such 5D dynamic process
131 remains unmet for earlier LFM demonstrations, and conventional LFD approaches would be unable to resolve
132 these dynamics at such resolution (**Supplementary Fig. 12,13**) and throughput.
133 To record the neural activities of acting *C. elegans* at single-cell resolution, previous studies image either
134 the whole worm in 2D at 50 Hz frame rate[16], or alternatively immobilized worm in 3D at up to 50 Hz volume
135 rate[6]. An additional infrared channel is also required to provide bright-field images of worm behaviors as
136 reference. VCD-LFM, capable of 100 Hz recording and 16.5 Hz reconstruction of fast 3D processes,
137 combined with network-based all-fluorescence analysis, offer a new paradigm shift to the investigation of
138 neural activities and relevant worm behaviors in whole moving worms without motion blur or computation
139 traffic. As for cardiovascular imaging, previous approaches also image beating heart/ blood flow in zebrafish
140 larvae[17], at the cost of non-uniform axial sampling, limited imaging depth and high excitation intensity. A
141 mild volumetric imaging method based on simple system and universal computation is still a crucial demand
142 for studying the dynamic properties and functions of the cardiovascular system at single-cell resolution.
143 Therefore, VCD-LFM could be a valuable tool for studying dynamics on these fast timescales, potentially
144 benefiting a variety of applications, such as neural signaling regulated phenotype studies, and medically
145 relevant dysfunctions of the heart and blood transport system in model organisms.
146 In summary, we introduce a VCD network LFM approach and demonstrate its ability to image biological
147 dynamics with increased spatial resolution, minimal reconstruction artefacts and 3-order higher
148 reconstruction throughput. These advances are crucial for sustained volumetric imaging of blood flow in
149 zebrafish larva and neural activities in acting *C.elegans* at single-cell resolution. While VCD-LFM greatly
150 improves the reconstruction quality from one originally coupled with the system optics to one can be
151 optimized by the prior data training, it yet requires the pre-acquisition of the training images from the same
152 type of specimens. We expect this could be improved by the unceasing development of deep learning, to
153 achieve more versatile computation. We also note that VCD models the LFM using scalar diffraction theory
154 in isotropic sample with free optic access. Non-negligible scattering in deep regions and aberration from
155 unmatched diffraction index will therefore impair the reconstruction quality. We anticipate future
156 combination with other enhancement methods, e.g., multi-view deconvolution or adaptive optics, to alleviate
157 the influence from non-ideal samples. It's noteworthy that VCD procedure is also compatible with modified
158 LFM modalities, such as dual-objective setup[11] or scanning light-sheet illumination[18]. Hence, higher-quality
159 light-field acquisitions will also contribute to better VCD reconstruction. Finally, we expect VCD-LFM could
160 potentially bring new insight for computational imaging techniques so that we can further push the
161 spatiotemporal resolution limits in observing biological dynamics.

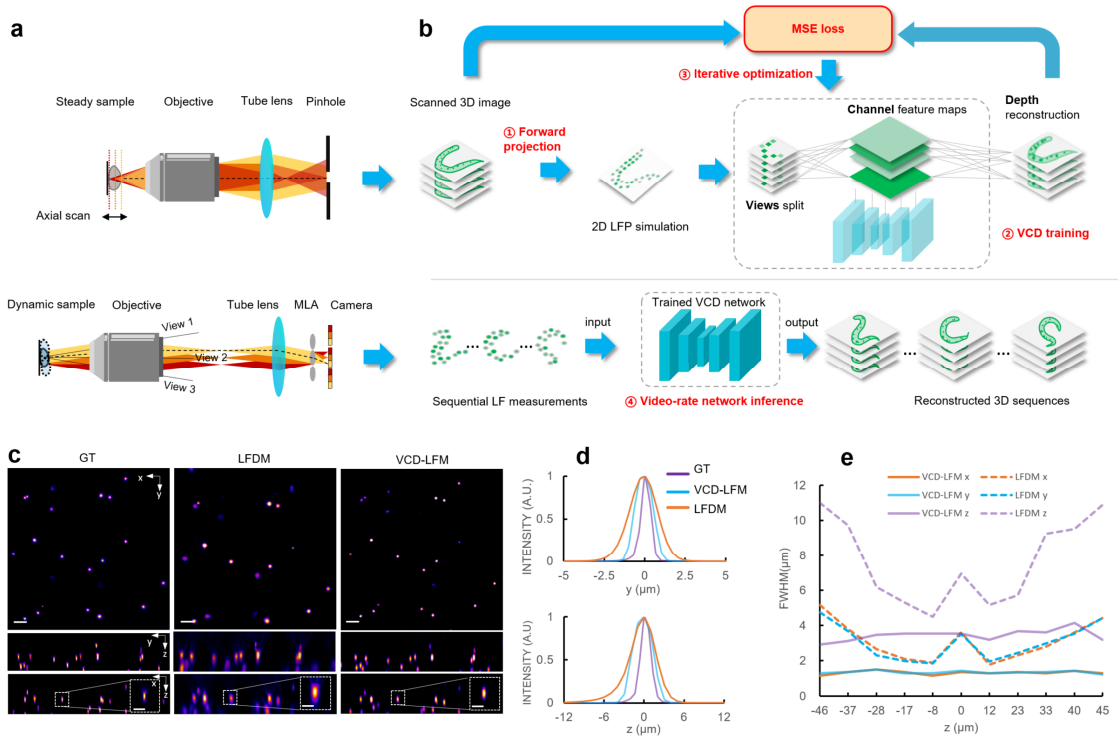

Figure 1. VCD-LFM and its performance. (a) Instantaneous recording of dynamic samples by a light-field microscope and 3D imaging of steady samples by a confocal microscope. **(b)** The VCD-Net reconstruction pipeline. First, a number of 2D synthetic light-field images are generated from the 3D confocal stacks (step1). VCD procedure then transforms these light-field inputs into intermediate 3D image stacks (step 2), which will be compared with confocal ground truths through pixel-wise MSE and be iteratively updated. The network is finally well trained with obtaining optimal parameters that can generate 3D images similar enough to the ground truths (step 3). The trained network is then capable of directly deducing 3D images from the recorded light-field images (step 4). **(c)** Maximum intensity projections (MIPs) of the same fluorescent beads and their resolution (FWHM) by 3D wide-field microscopy, LFDM and VCD-LFM, respectively, demonstrating an improvement in lateral and axial resolution for VCD-LFM. **(d)** PSF measurements at 11 μm off the focal plane with lateral (top) and axial (bottom) FWHM indicated. Scale bars, 10 μm (xy planes) and 5 μm (vignettes in x-z planes). **(e)** Average axial and lateral FWHM of the beads across the volumes reconstructed by LFDM (dashed lines) and VCD-LFM (solid lines), respectively. VCD-LFM improves lateral/axial resolution across the entire 90-μm overlapping volume and achieves uniform resolution of ~1.3 ± 0.1 μm and ~3.5 ± 0.4 μm in x/y and z, respectively.

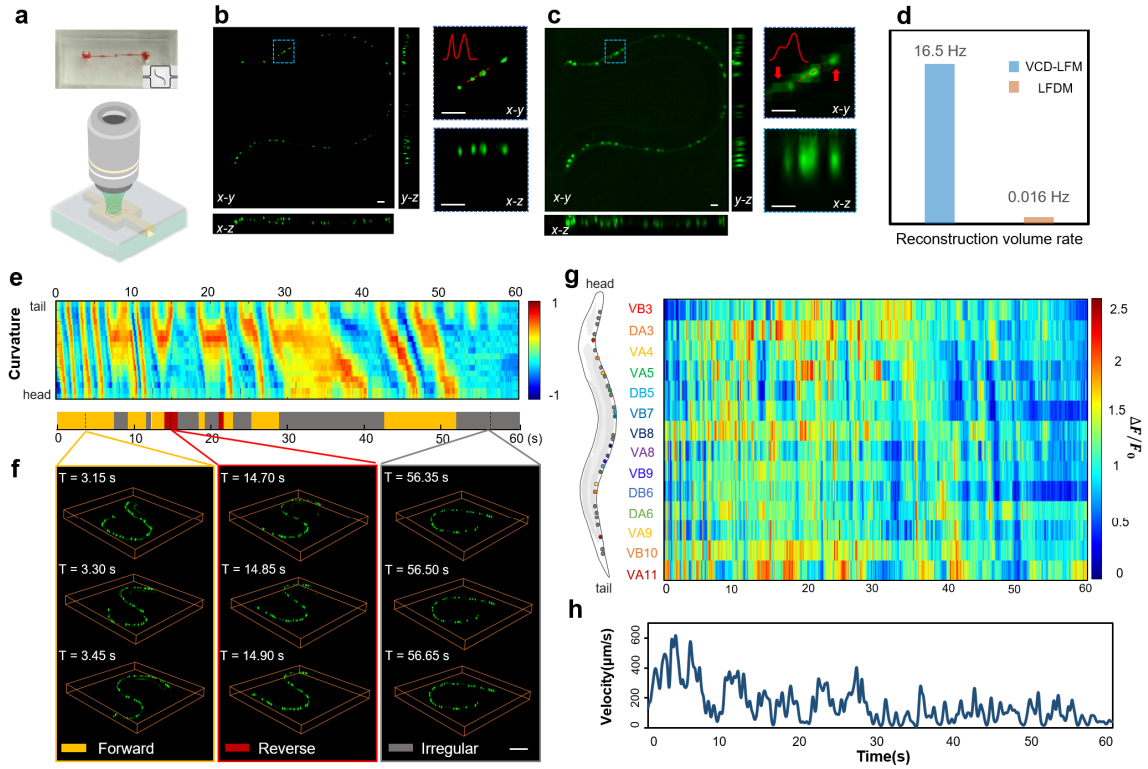

Figure 2. Whole-animal Ca^{2+} imaging of moving *C.elegans* using VCD-LFM. (a) Configuration combining light field with microfluidic technique for imaging motor neuron activities (*hpls595* line) of L4-stage *C.elegans* acting in a microfluidic chip (300 × 300 × 50 μm, top panel) at 100 Hz recording rate. (**b**,**c**) MIPs of one instantaneous volume reconstructed by VCD and LFD, respectively. The magnified views of selected region show that VCD restores single-cell resolution and removes artefacts (red arrow). **(d)** The reconstruction throughput of VCD-LFM and LFDM for processing light-field video. **(e)** Curvature kymograms along the body of moving worm based on deep learning (U-net) segmentation of reconstructed worm body contours. An ethogram describing the worm behavior over the time (lower panel) is obtained by analyzing the curvature change. **(f)** Selected consecutive volumes (9 of 6000) in accordance to the ethogram visualizing the forward (left), backward (middle) and irregular (right) crawling tendency of the worm. **(g)** Neural activities map of the acting worm (right) by quantifying the Ca^{2+} signal fluctuation of identified motor-neurons (left)**.** Each row shows the fluctuation of an individual neuron with color indicating percent fluorescence changes ($\Delta F/F_0$). (**h**) The average velocity plot of the worm during 1-minute observation. Scale bars, 10 μm **(b)**,**(c)** , 50 μm in **(f)**.

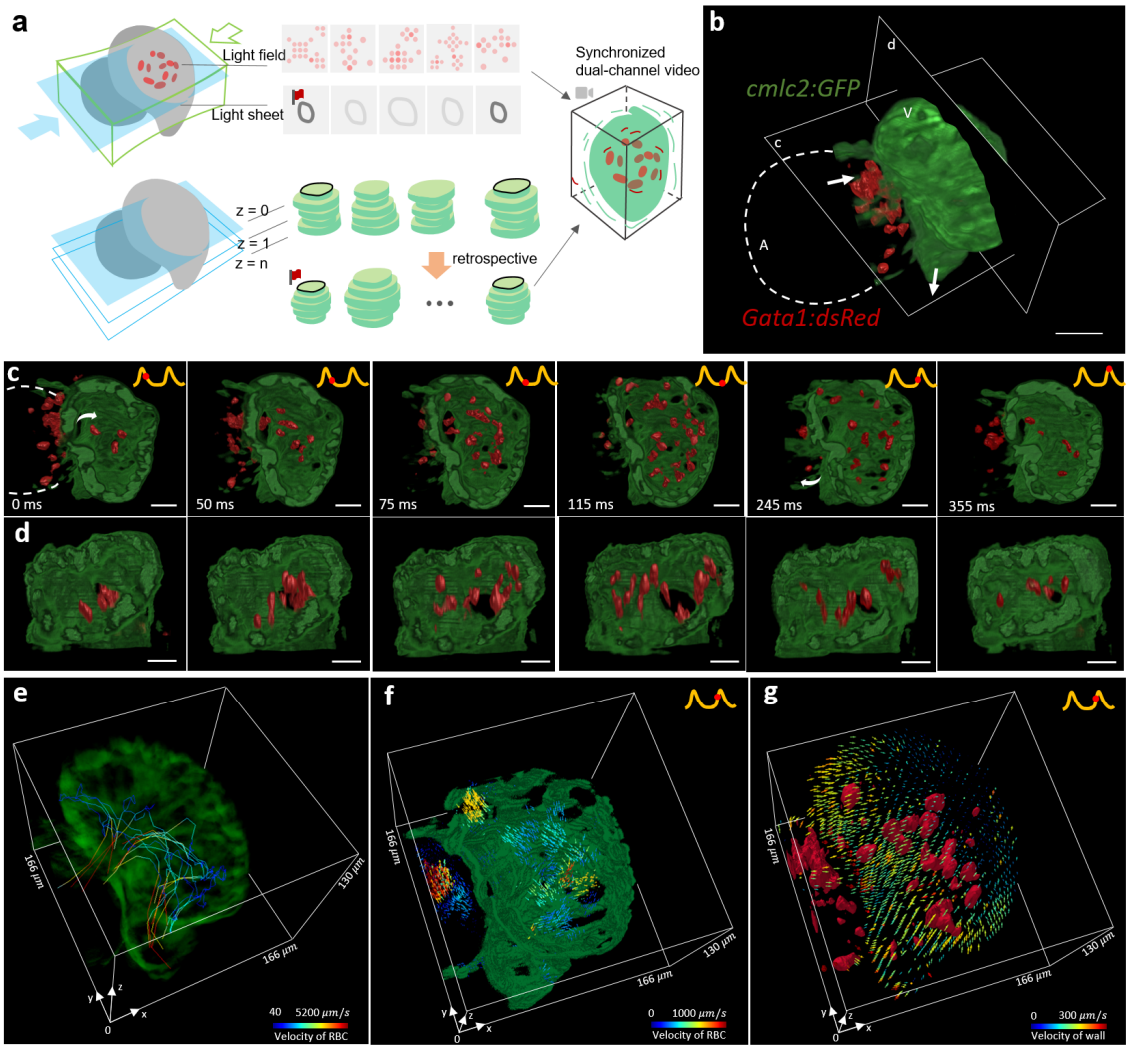

Figure 3. 5D imaging of blood flow and beating heart in zebrafish using a selective plane/volume illumination LFM setup. **(a)** Schematic of 5D (space + time + spectrum) images acquisition and reconstruction process. 2D time sequences of flowing RBCs (gata1a::dsRed) and beating cardiomyocytes (cmlc2::GFP) were recorded simultaneously with LFM and SPIM detections, respectively, under selective volume (532-nm) and plane (488-nm) illumination (upper panel). This images provides a time stamp for spatiotemporally aligning the VCD-reconstructed 4D blood flow and retrospectively-synchronized 4D heartbeat, forming the final 5D cardiac hemodynamics (lower panel). **(b)** 4D visualization (space + spectrum) of the heart in a transient moment. Arrows indicate the direction of blood flow, from the AV canal to outflow tract. Dotted line marks the region of atrium. A: atrium; V: Ventricle. Scale bar, 40 µm. **(c)**, **(d)** 5D bisected heartbeat along sectioning planes indicated in **(b)**. Myocytes were synchronized with blood blow using the periodic movement of heart wall. Dotted line indicates the region of atrium. Scale bar, 30 µm. **(e)** Tracks of 14 single RBCs throughout the cardiac cycle. A static heart is shown for reference of cardiac geometry. **(f),** Particle image velocimetry (PIV) of two temporally adjacent volumes of RBCs at the end of systole. A static bisected heart is shown for reference of cardiac geometry. **(g)** Displacement analysis of the cardiac wall (cardiomyocytes) movement between the same 2 adjacent volumes used in **f**. Each arrow demonstrates the direction and velocity of the corresponding part of tissue. Static RBCs of one volume are shown for reference.

Methods

Epi−illumination LFM setup

A light-field setup was built on an upright microscope (BX51, Olympus). The light-field/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×/0.8W, Olympus) was used to collect the epifluorescence signals from samples. For recording the light field, a microlens array (OKO Optics, 40K elements) was placed at the native image plane, projecting the light-field signals $(x, y, \theta_x, \theta_y)$ onto the camera (Hamamatsu Flash 4.0). See **Supplementary Fig. 1** for more details of Epi-LFM setup

Selective plane/volume illumination based LFM setup

In addition to the traditional LFM geometry, we also developed a light-field setup based on selective volume illumination. Two pairs of beam reducers combined with an adjustable iris were used to generate scalable rod-like beam (532-nm) perpendicular to the detection path. It confined the fluorescence excitation within the heart region of zebrafish embryo, reducing the excessive emission of RBCs that could smear the desired signals. This selective volume illumination provided light-field image with less background and increased contrast[19]. For dual-color imaging of RBCs/cardiomyocytes, we also introduced a 4-μm laser-sheet illumination (488-nm) in conjunction with z-scan of sample so that selective plane illumination imaging of periodically beating cardiomyocytes as well as selective volume illumination light-field imaging of RBCs could be implemented simultaneously. The light-rod/light-sheet illumination paths were precisely aligned, with providing double excitation to the sample from its dual sides. A water immersion objective (Fluor 20×/0.50W, Nikon) were used to collect the dual-color fluorescence signals from the RBCs/cardiomyocytes. A dichroic mirror split the GFP (cardiomyocytes)/DsRed (RBCs) signals for wide-field and light-field detection, respectively. The light-field detection here followed the same design used in the aforementioned 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), the output channels of certain *N*th convolutional layer receives output feature maps from the last *(N-1)*th layer, and generate new channel feature maps using different convolution kernels of each channel. This *N*th feature map then flows to the next *(N+1)*th layer as the input to further distill the features and generate the *(N+1)*th map. The network finally generates a multi-channel output where each channel is a non-linear combination of the original input channels. This mechanism intrinsically agrees with the LFD, in which each axial slice of the reconstructed volume was generated by a superposition of all angular views of light field projection convolved with the point spread function (PSF). Therefore, the VCD transformation is fundamentally suited for reconstructing 3D volume with depth information from a bunch of angular views of light field images. Our VCD network is based on an U−Net architecture containing 6 convolution layers and 5 deconvolution layers. Each convolution layer and deconvolution layer has three parameters: **n, f** and **s**, denoting the output channels number of this layer,

the filter size of convolution kernel and the down/up scaling rate when kernel moves, respectively. With applying a light-field projection based on wave optics mode [8], we first convolved high-resolution 3D label images (dimension: height **a** × width **b** × depth **c**) with a calculated light field PSF, generating synthetic 2D light-field images (dimension: **a** × **b**) for network training. After the VCD network took these light-fields as input, the initial convolutional layer dealt with a re-formatted 3D ([**a/d**, **b/d**, $\mathbf{d^2}$]) data, in which the 3rd dimension $\mathbf{d^2}$ was the extracted and successively arranged angular views from the raw light-fields. A subvoxel up-scaling part of this initial layer interpolates the lateral dimension of these 3D view images with dimension **a** (width) × **b** (height) × $\mathbf{d^2}$ (view). Then in the first VCD convolution layer, all the **n** channels with its own convolution kernel size **f** extracted feature maps from all these views, generating an output with dimensional **a** × **b** × **n**, where **n** is the channel numbers, thus finished the initial transformation from "view" to "channel". To fully excavate the hidden features from input images, the following convolution and deconvolution layers kept combining old channels from previous layer and generating new ones according to their (**n**, **f**, **s**) parameters. The last deconvolution layer finally outputted a 3D image with the channel number **n** equal to the desired number of plane **c**, thereby generating a network output with the same dimension **a** × **b** × **c** as those HR references. This step fulfilled the transformation from "channel" to "depth". Such a 2D-to-3D VCD transformation at beginning stage unsurprisingly generated results far from those high-resolution references so this VCD neural network need to be gradually optimized via iteratively minimizing the difference between its intermediate outputs and the high-resolution references. With setting appropriate penalty functions, such as the pixel intensity MSE of outputs and references, the VCD network could obtain optimized **n, f, s** parameters for each layer and efficiently converge to a well-trained status, at which the network was capable of transforming the synthetic light-field inputs back to 3D outputs accurate enough as compared to the source 3D images. At reconstruction stage, unlike the traditional iterative reconstruction methods, the trained VCD network directly inferred a sequence of 3D images from the input experimental light-field video that typically contained thousands of LF frames recording the dynamic biological process. Thus, the reconstruction throughput could be very high, for example, reconstructing *C. elegans* at ~16.5 volumes (size: 924×924×31) s^{-1} by using a workstation equipped with four GPUs.

PSF measurement

Before light-field video of highly dynamic processes, e.g., behavior of whole acting *C. elegans*, being captured, we first used both light-field and 3D wide-field (z-scan) modes in our epi-illumination LFM system, to image sub-diffraction fluorescent beads (0.5 μm Lumisphere, BaseLine) distributed in a piece of hydrogel (0.7% low melting agarose solution, BBI Life Sciences), and thereby characterize the performance of VCD-LFM. 40×/0.8w objective (LUMPlanFL, Olympus) was used to image the same volume of beads across 325 μm × 325 μm × 60 μm FOV. 60-μm depth stacks were recovered from the recorded light-field image using LFD and VCD approaches. Their results were compared with wide-field image stacks to show the better reconstruction fidelity of VCD. The PSFs of LFDM and VCD-LFM at certain depth were quantified via plotting the line profiles of the same resolved beads. The achieved lateral and axial resolutions at certain depth by each method were then indicated through calculating the averaged FWHM values of the resolved beads, e.g., z=11 μm and n=15 beads. The axial performance of VCD-Net *versus* LFD at different depths were further analyzed via measuring the FWHMs of resolved beads (synthetic) across a 90-μm depth, with their variation indicating the non-uniformity of light-field recovery.

C. elegans experiments

All *C. elegans* were cultured on standard Nematode Growth Medium (NGM) plates seeded with OP50 and maintained at 22 °C incubators. The strain ZM9128 *hpIs595*[*Pacr-2(s)*::GCaMP6(f)::wCherry] of *C. elegans*, which expressed the calcium indicators to A-and B- class motor neurons, was used to detect the Ca^{2+} signals in moving animals. For imaging neural activities, late L4 stage worms were loaded into a microfluidic chip with chamber size 300 μm × 300 μm × 50 μm, allowing the worm to act within the FOV of 40× objective (LUMPlanFLN 40×/0.8W, Olympus). The camera recorded the light-field video of GFP calcium signals of acting worm at 100 Hz frame rate. The recording time varied from 1 to 5 minutes depending on the status of the worm. A 2× pixel binning was applied to improve the SNR of captured calcium signals. To obtain HR 3D data for network training, worms were anesthetized by 2.5 mM levamisole (Sigma-Aldrich) in M9 buffer. The RFP background signals of labelled neurons in L4 worm larva were captured using a 40×/0.95 objective on a confocal microscope (FV3000, Olympus). To validate network performance, the anesthetized strain QW1217 *hpIs491*[*Prgef-1*::GCaMP6(f)];*hpIs467*[*Prab-3*::NLS::RFP] was *insitu* imaged by both light-field and z-scan wide-field modes using the same 40×/0.8 objective. The light-field image was then processed by both VCD network and LFD, with their results being further compared with 3D wide-field image (**Supplementary Fig. 9**).

Zebrafish Experiment

Transgenic zebrafish lines *Tg*(cmlc2:gfp-gata1a:dsRed) with cardiomyocytes tagged with GFP and RBCs tagged with dsRed were used in our experiment (see **Supplementary Fig. 12,13** for other fish lines used for imaging). The two lines were crossbred and micro-injected with gata2a morpholino oligonucleotide to reduce hematopoiesis, thereby, resulting in a mild reduction in the number of red blood cells. Embryonic stage fish were maintained until 4 days post fertilization in standard E3 medium, which was additionally supplemented with PTU (Sigma Aldrich, MO) to inhibit melanogenesis. Then, the larvae were anesthetized with tricaine (3-amino benzoic acidethylester, Sigma Aldrich, MO) and immobilized in 1% low-melting-point agarose inside FEP (Fluorinated Ethylene Propylene) tube for further imaging. All the experiments were performed in compliance with UCLA IACUC protocols.

For light-field imaging of the blood flow and heartbeat, the middle of ventricle was moved to the focal plane so that the light-rod (532-nm) could selectively illuminate RBCs flowing through the heart while the light-sheet (488-nm) provided optical sectioning of cardiomyocytes deep inside the heart. The cameras at DsRed and GFP channels simultaneously recorded light-field video of rapidly flowing RBCs and wide-field video of cardiac cycles at certain plane of beating heart, respectively, thereby providing temporal reference (time stamp) for dual-color synchronization afterward. A movie stack of light-sheet images was then acquired with containing 60-70 movies at different planes (z spacing=2 μm) that covered the complete myocardial structures. Both light-field and light-sheet movies was recorded at 200 Hz frame rate, with 768 × 768 frame size that corresponds to a lateral FOV of ~250 μm x 250 μm. The movies covered 4-5 cardiac cycles with containing 450 frames (5 ms exposure for each frame). To obtain high-resolution images of RBCs for network training, deeply anesthetized fish larvae with stationary 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) and dsRed channel.

Implementation details of VCD-LFM

The optical parameters and data implementation details of VCD-LFM at imaging / training / reconstruction stages for fluorescence bead, *C.elegans* and zebrafish heart were further shown as **Supplementary Table 2**.

5D post-reconstruction in toto visualizing the cardiac hemodynamics

To reconstruct the blood flow from acquired light-field images, we trained the VCD-Net using confocal image stacks of static red blood cells from 16 zebrafish embryos. When generating synthetic light-field images for training, a light-field PSF that matched the imaging system was chosen. It's designed to yield 101 planes from -50 μm to 50 μm depth of focal plane. The network converged after 100 epochs training, which took around 3.5 hours on a GPU workstation (GTX 2080 Ti, Nvidia). Then the trained network merely spent 52.3 s to reconstruct 400 consecutive volumes from acquired light-field videos. This 4D reconstruction throughput was compared to 10.62 hours (38232 s) by conventional LFD method (8 iterations) with using the same light-field PSF to reconstruct the same amount of volumes and based on the same computer.

4D reconstruction of periodic heartbeat was done by post-synchronization of the acquired SPIM movie stack[17, 19, 20], in which one plane movie was recorded simultaneously with the light-field movies of RBCs. The dynamics of myocardial structures at this plane then provided a sequence of time markers to permit accurate temporal alignment of entire 4D beating heart reconstructed by post-synchronization, with the 4D blood flow reconstructed by VCD network, thereby forming the 5D visualization that could completely visualize the cardiac hemodynamics. The aforementioned post-computation was run on a workstation equipped with Intel(R) Core(TM) i9-7900X CPU @3.3GHz, 64G RAM, and GeForce RTX 2080Ti graphic cards.

Quantitative Analysis of neural activities and behavior of acting C. elegans .

Based on the 4D image of the worm by VCD-LFM, we performed semi-automated tracking of the movement and intensity fluctuation of each individual neuron using TrackMate Fiji Plugin[21] . Neurons were segmented automatically in each volume with applying a circular ROI through the difference of gaussian (DoG) detector and then tracked using Kalman filter which predicted the most probable position range of the neuron in the next frame. In a few cases when this automatic tracking failed because of fast movement of the neuron, the missing detections and tracking mistakes were manually corrected. For each neuron in each volume, all the pixels within the ROI were averaged to produce a single fluorescence value F. The F values in the time series was re-normalized using a linear-decay function, to remove the effect of photobleaching. To extract $\Delta F/F0$, we calculated $\Delta F/F0 = (F(t)-F0)/F0$ with F(t) being the fluorescence intensity at certain time point, and F0 being the mean value of background intensity of F(t).

Worm behavior analysis was then implemented based on deep learning processing at high throughput. We successfully inferred the 2D worm outlines throughout the whole time period using a U-Net based image segmentation[22], by which the center lines could be further extracted to calculate the changing curvatures and motion velocities[23]. More details are given in **Supplementary Fig. 11** and its note. The body curvature map was shown in Fig. 2e, indicating the time-varying worm behaviors during the 1-minute acquisition. The velocity curves shown in Fig. 2h were smoothed using curve-fitting toolbox and "smoothing spine" function of MATLAB.

365 *PIV analysis of RBCs and displacement analysis of cardiomyocytes in zebrafish heart*

366 We developed a 3D Particle Image Velocimetry (PIV) code to map the velocity field of RBCs in adjacent
367 frames. The method was based on the pixel-intensity-similarity between images, where for each ROI in the
368 first frame we searched for the correspondence that had the highest correlation with the second frame. To
369 minimize the error from pixel coordinate system, a continuous polynomial function was applied to the 3D
370 correlation map to find the optimal coordinates of the corresponding ROI. The shift between ROIs, which
371 was assumed to represent the motion of RBC flow, was converted to the velocity field, given the known
372 frame rate and pixel size. We further used the RBC image as a mask to filter out noise vectors on background
373 in visualization. The displacement field of cardiomyocytes was generated by deformable image registration,
374 for which we chose Demons Algorithm[24, 25]. It estimated the displacement vectors that could map each pixel
375 location in the first frame to the next frame of the heart images. Like aforementioned PIV analysis, the vectors
376 were used to represent the movement of the beating heart, after filtering the background noise. The analysis
377 was written in Matlab and the vector field was visualized by Mayavi[26].
378

379 **Funding**

380 National Key R&D program of China (2017YFA0700500), the National Natural Science Foundation of
381 China (21874052, and 31671052), National Institute of Health (NIH, 5R01HL129727-04), the National
382 Science Foundation of Hubei (2018CFA039), Wuhan Morning Light Plan of Youth Science and Technology
383 (2017050304010295) and the Junior Thousand Talents Program of China (Peng Fei and Shangbang Gao).
384

385 **Acknowledgements**

386 We thank Jau-Nian Chen and Yuan Dong for providing the zebrafish and advices on fish experiment. We
387 thank Haiwen Li for technical assistance on *C.elegans* behavior analysis. We thank Yanyi Huang and Jianbin
388 Wang for discussions and comments on the manuscript. Confocal laser scanning microscopy (SP8-
389 STED/FLIM/FCS) was performed at the Advanced Light Microscopy/Spectroscopy Laboratory and the
390 Leica Microsystems Center of Excellence at the California NanoSystems Institute at UCLA with funding
391 support from NIH Shared Instrumentation Grant S10OD025017 and NSF Major Research Instrumentation
392 grant CHE-0722519.
393

394 **Disclosures**

395 The authors declare that there are no conflicts of interest related to this article.

References

1. Pawley, J. Handbook of biological confocal microscopy. (Springer Science & Business Media, 2010).
2. Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J. & Stelzer, E.H.K. Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy. *Science* **305**, 1007-1009 (2004).
3. Keller, P.J., Schmidt, A.D., Wittbrodt, J. & Stelzer, E.H.K. Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy. *Science* **322**, 1065-1069 (2008).
4. Planchon, T.A. et al. Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. *Nature methods* **8**, 417 (2011).
5. Levoy, M., Ng, R., Adams, A., Footer, M. & Horowitz, M. Light field microscopy. *ACM Transactions on Graphics (TOG)* **25**, 924-934 (2006).
6. Prevedel, R. et al. Simultaneous whole-animal 3D-imaging of neuronal activity using light-field microscopy. *Nature Methods* **11**, 727-730 (2014).
7. Pégard, N.C. et al. Compressive light-field microscopy for 3D neural activity recording. *Optica* **3**, 517 (2016).
8. Broxton, M. et al. Wave optics theory and 3-D deconvolution for the light field microscope. *Optics express* **21**, 25418-25439 (2013).
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. *Nature methods* **14**, 811-818 (2017).
11. Wagner, N. et al. Instantaneous isotropic volumetric imaging of fast biological processes. *Nature methods* **16**, 497-500 (2019).
12. Cohen, N. et al. Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express* **22**, 24817-24839 (2014).
13. Fei, P., Wang, Z., Zhang, H., Yang, Y. & Li, Y. Deep learning light field microscopy for rapid four-dimensional imaging of behaving animals. (2018).
14. 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).
15. Wen, Q., Gao, S. & Zhen, M. Caenorhabditis elegans excitatory ventral cord motor neurons derive rhythm for body undulation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **373** (2018).
16. Leifer, A.M., Fang-Yen, C., Gershow, M., Alkema, M.J. & Samuel, A.D. Optogenetic manipulation of neural activity in freely moving Caenorhabditis elegans. *Nature methods* **8**, 147 (2011).
17. Mickoleit, M. et al. High-resolution reconstruction of the beating zebrafish heart. *Nature methods* **11**, 919 (2014).
18. Truong, T.V. et al. Selective volume illumination microscopy offers synchronous volumetric imaging with high contrast. *bioRxiv*, 403303 (2018).
19. Lee, J. et al. 4-Dimensional light-sheet microscopy to elucidate shear stress modulation of cardiac trabeculation. *The Journal of clinical investigation* **126**, 1679-1690 (2016).
20. Liebling, M., Forouhar, A.S., Gharib, M., Fraser, S.E. & Dickinson, M.E. Four-dimensional cardiac imaging in living embryos via postacquisition synchronization of nongated slice sequences. *Journal of biomedical optics* **10**, 054001 (2005).
21. Tinevez, J.Y. et al. TrackMate: An open and extensible platform for single-particle tracking.

Methods **115**, 80-90 (2017).

22. Ronneberger, O., Fischer, P. & Brox, T. 234-241 (Springer International Publishing, Cham; 2015).
23. Restif, C. et al. CeleST: computer vision software for quantitative analysis of C. elegans swim behavior reveals novel features of locomotion. *PLoS computational biology* **10**, e1003702 (2014).
24. Thirion, J.-P. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical image analysis* **2**, 243-260 (1998).
25. Vercauteren, T., Pennec, X., Perchant, A. & Ayache, N. Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage* **45**, S61-S72 (2009).
26. Ramachandran, P. & Varoquaux, G. Mayavi: 3D visualization of scientific data. *Computing in Science & Engineering* **13**, 40-51 (2011).