Two-photon Bessel beam tomography for fast volume imaging

ANDRES FLORES VALLE AND JOHANNES D. SEELIG *

Center of Advanced European Studies and Research (caesar), 53175 Bonn, Germany
*johannes.seelig@caesar.de

Abstract: Light microscopy on dynamic samples, for example neural activity in the brain, often requires imaging volumes that extend over several $100 \ \mu m$ in axial direction at a rate of at least several tens of Hertz. Here, we develop a tomography approach for scanning fluorescence microscopy which allows recording a volume image in a single frame scan. Volumes are imaged by simultaneously recording four independent projections at different angles using temporally multiplexed, tilted Bessel beams. From the resulting projections, three-dimensional images are reconstructed using inverse Radon transforms combined with convolutional neural networks (U-net).

1. Introduction

In many fluorescence microscopy experiments samples show dynamics which evolves not in a single focal plane but in a volume. For example, imaging neural activity in the brain requires monitoring multiple cells that are distributed in three dimensions [1, 2]. In this situation, the temporal resolution of the microscope needs to match the time scale of the sample dynamics, which for neural activity monitored with calcium indicators, lies in the range of several tens of Hertz [1, 3, 4].

While such imaging rates are standard for scanning a single plane, they are more difficult to reach for volumes. The scan rate is ultimately limited by the requirement of integrating a sufficient number of photons per scanned pixel. To overcome this limit, solutions for volume imaging have been developed that scan multiple pixels in parallel [3, 4]. Multifocal microscopy, for example, images several spots at the same time, but requires detectors with multiple pixels making this approach susceptible to scattering (for example when imaging in tissue) [3, 5]. Multiple focal spots can also be imaged in parallel with single pixel detectors: by introducing temporal offsets between different (typically pulsed) beams, fluorescence emission can be sorted into different temporal channels [6, 7].

To speed up volume imaging rates, extended volumes - instead of multiple focal spots - can be excited at the same time, for example using Bessel beams, Airy beams, line foci or light sheets [4, 8, 9]. Combined with camera detectors at an angle to the axis of excitation, as for example in Bessel beam light sheet microscopy [8], these approaches work best for weakly scattering samples [3, 4], but Bessel beams or line foci [5] can also be used for imaging with non-descanned PMT detection. This is more suitable for imaging in scattering tissue and results in a projection of the entire imaging volume along the beam axis [10–14]. While for Bessel beams this approach offers the advantage of recording from an entire volume in a single frame scan, it comes at the cost of losing all axial resolution.

Axial resolution was restored in experiments that used two Bessel beams which imaged the sample at different angles [11, 12, 14] for generating two different projections. These projections were either recorded sequentially [11, 12] or simultaneously (as required for moving samples such as brain tissue *in vivo*) in a single channel by averaging over two different projections [14]. From these projections axial information was recovered based on the distance between prominent, sparse features such as fluorescent beads or cell bodies [12, 14].
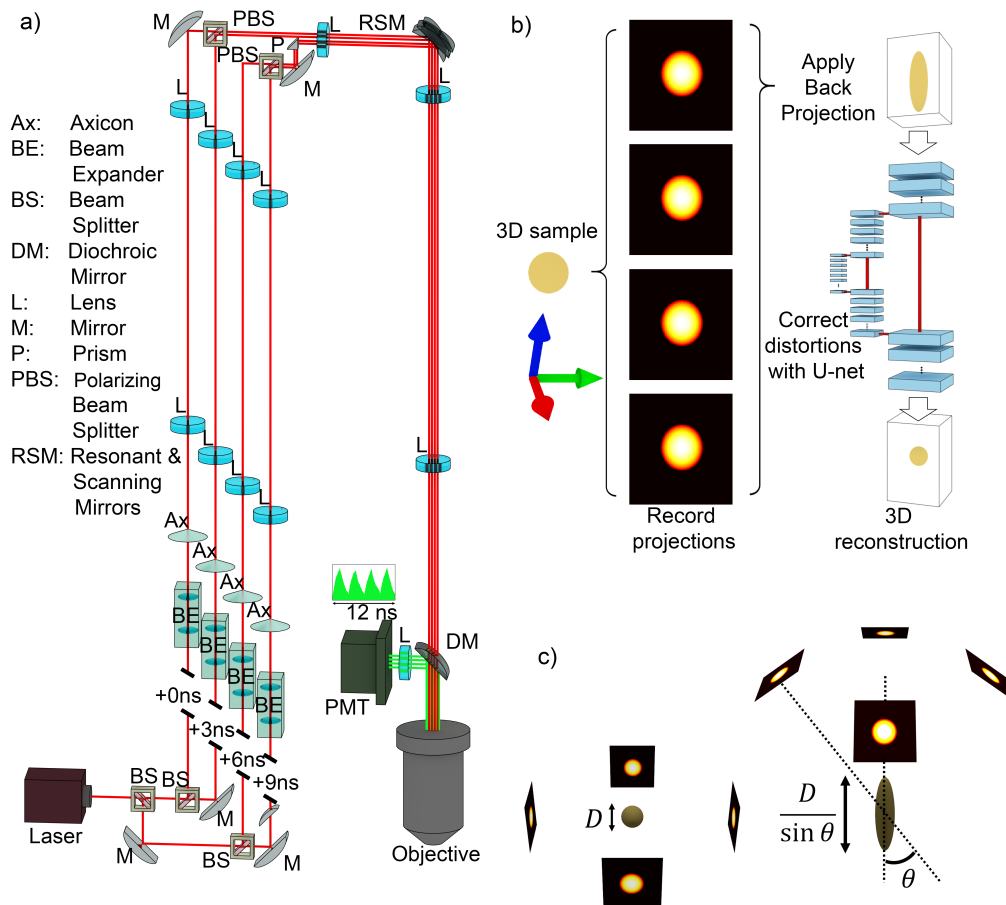
Fig. 1. (a) Schematic of the experimental setup. A laser beam is split into four temporally offset beams using beam splitters (BS). Each beam is converted into a Bessel beam with independently adjusted beam parameters. The Bessel beams excite the volume from four different sides. Inset: Fluorescence emission is detected in four temporally separated channels. (b) Workflow for volumetric reconstruction: four projections are recorded from a three dimensional (3D) sample and back projection is applied to obtain its 3D tomographic reconstruction. Then, a U-net is applied to correct distortions in the reconstructed sample. (c) Illustration of 3D tomography with different viewing angles. Left side: projections parallel to the xy plane allow a perfect reconstruction of a sphere of diameter D. Right side: projections at an oblique angle θ produce a reconstructed sphere elongated in the z axis by a factor of $1/\sin(\theta)$. This is corrected for by using convolutional networks. Projection images are obtained by bidirectional raster scanning of tilted Bessel beams across the sample (see Methods for details).

Recording multiple projections at different viewing angles generally forms the basis for reconstructing volumes in tomographic imaging [15]. While such tomographic approaches have been implemented for fluorescence microscopy by rotating the samples with respect to the illumination direction [16, 17], this is not compatible with most *in vivo* imaging preparations. A tomography approach compatible with fluorescence microscopy for *in vivo* imaging has recently been developed with Gaussian line foci generating projections in the focal plane and was combined with scanning the objective along the axial direction to record volumes [5].
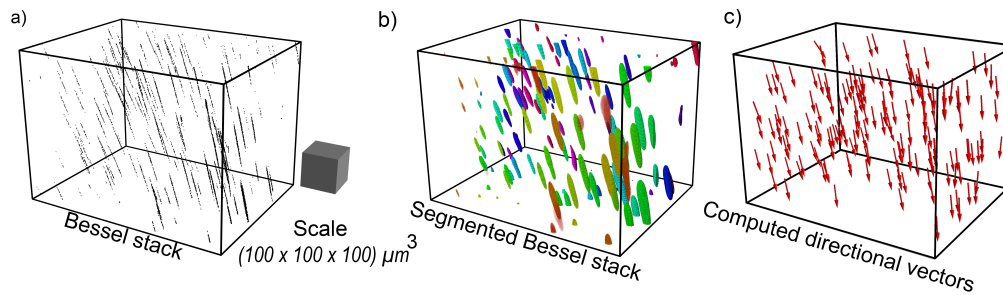
a)   b)   c)

Bessel stack   Scale (100 x 100 x 100) μm^3 Segmented Bessel stack   Computed directional vectors

Fig. 2. Calibration procedure for a single Bessel beam. (a) A volumetric sample of $1\mu m$ diameter beads in Agarose recorded with a Bessel beam with $1\mu m$ step size in the axial direction (z-axis), resulting in a Bessel beam z-stack. (b) For segmenting different beams, we applied a Gaussian filter with standard deviation of 5 pixels ,thresholded the image with its mean value and binarized it. Segmented beams are colored randomly for visualization. (c) Using Principal Component Analysis (PCA) we computed the center of each beam and the directional vector at its center. These vectors are used in linear regression to compute the directional vector field for each the Bessel beam z-stack. Viewing angles were selected for best visibility of the vector fields.

We here develop a tomography method based on Bessel beams. Volume information is obtained from four independent projections recorded at the same time, in a single frame, from four different angles by scanning temporally multiplexed, tilted Bessel beams. For volume reconstruction we combine inverse Radon transforms adapted for Bessel beam scanning with machine learning [15, 18–22]. Machine learning has been shown for example in optical phase imaging to allow high resolution reconstruction from sparse projections at shallow angles similar to the ones used here [20]. We show that three-dimensional convolutional neural networks (U-net) trained on suitable datasets similarly improve reconstructions in Bessel beam tomography. This scanning microscopy method allows volumetric imaging with non-descanned detection in a single frame scan.

2. Results

2.1. Setup

The setup is based on a custom built two-photon scanning microscope controlled by Scanimage (Vidrio) in a resonant scanning configuration with photon counting. The microscope optics was adapted to scan four tilted Bessel beams across the sample as described in the following. Four different projections were generated by using four different Bessel beams tilted with a mean angle of 166.75 (standard deviation 2.75) degrees with respect to the optical axis and an azimuth separation of approximately 90 degrees between neighboring beams (see Table 1 and Fig. 8 for more details and for quantification). The setup is shown in Fig. 1(a) and pictures of the setup are shown in Figs. 9(a) and 9(b). A laser beam (Coherent Discovery, 120 fs pulse width) is split into four beams with non-polarizing plate beam splitters and the (nevertheless polarization dependent) splitting ratio is additionally balanced using a half-wave plate for each beam (not shown). The beams are separated with a delay of 3 ns (corresponding to a 1 m additional path length per beam) [6, 7] which allows distinguishing the different projections with a single detector using photon counting at a time resolution faster than the pulse repetition rate, provided that the fluorescence lifetime is shorter than the pulse separation [6]. The beam width and collimation is adjusted for each beam with a telescope (two achromatic lenses, one mounted on a translation stage).
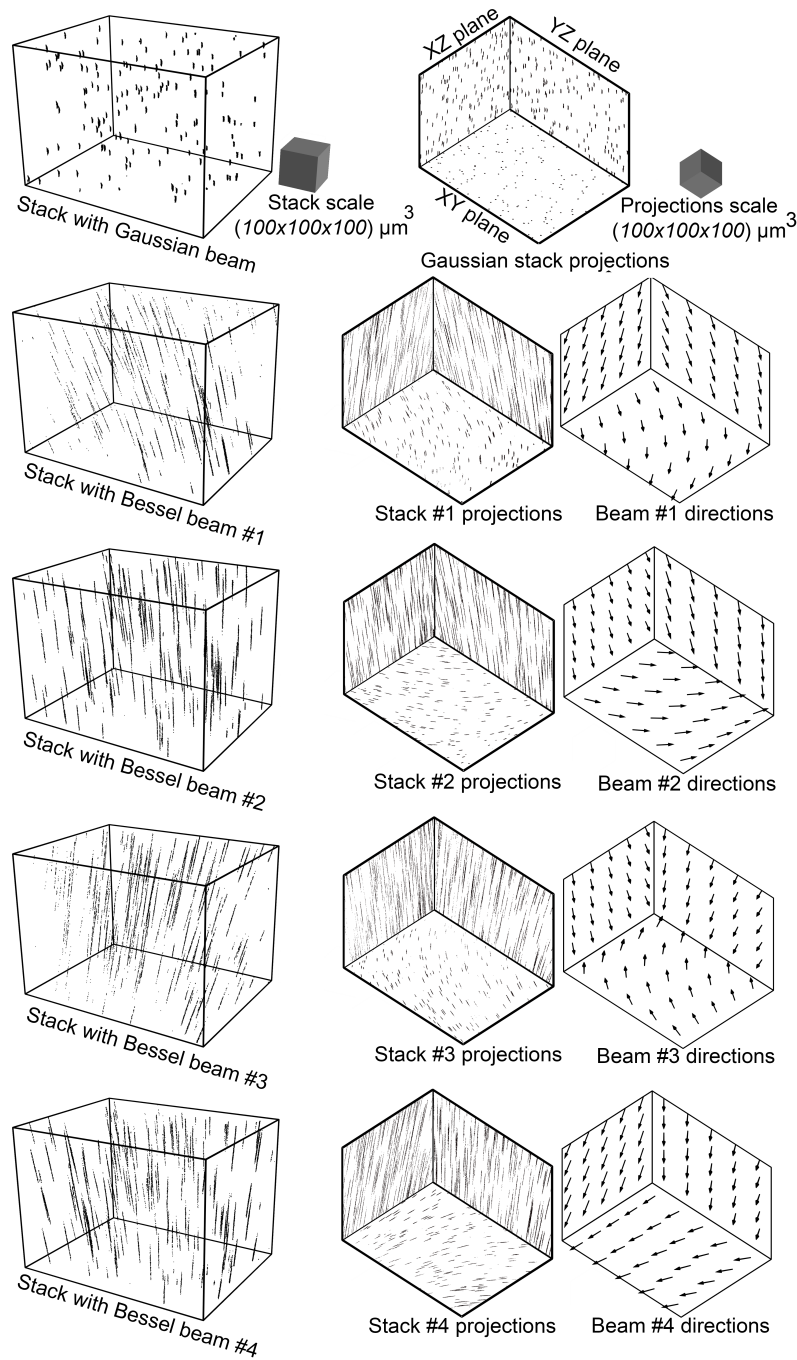
Fig. 3.   Directional vector fields of each Bessel beam measured and computed from a volumetric sample of $1\mu m$ diameter beads. Top row: sample recorded with a Gaussian beam (left side) and projections of the volume in the xy, xz and yz planes, respectively (right side). Second row to bottom row: volumetric samples recorded by the four Bessel beams (left side) and plane projections (center). The extrapolated vector field for each beam (right side) is shown in plane projections. The elevation angle is centered around angle 166.75+/- 2.75 degrees, see Table 1 and Fig. 8 for more details and quantification of vector fields. Viewing angles were selected for best visibility of the vector fields and projections, respectively.
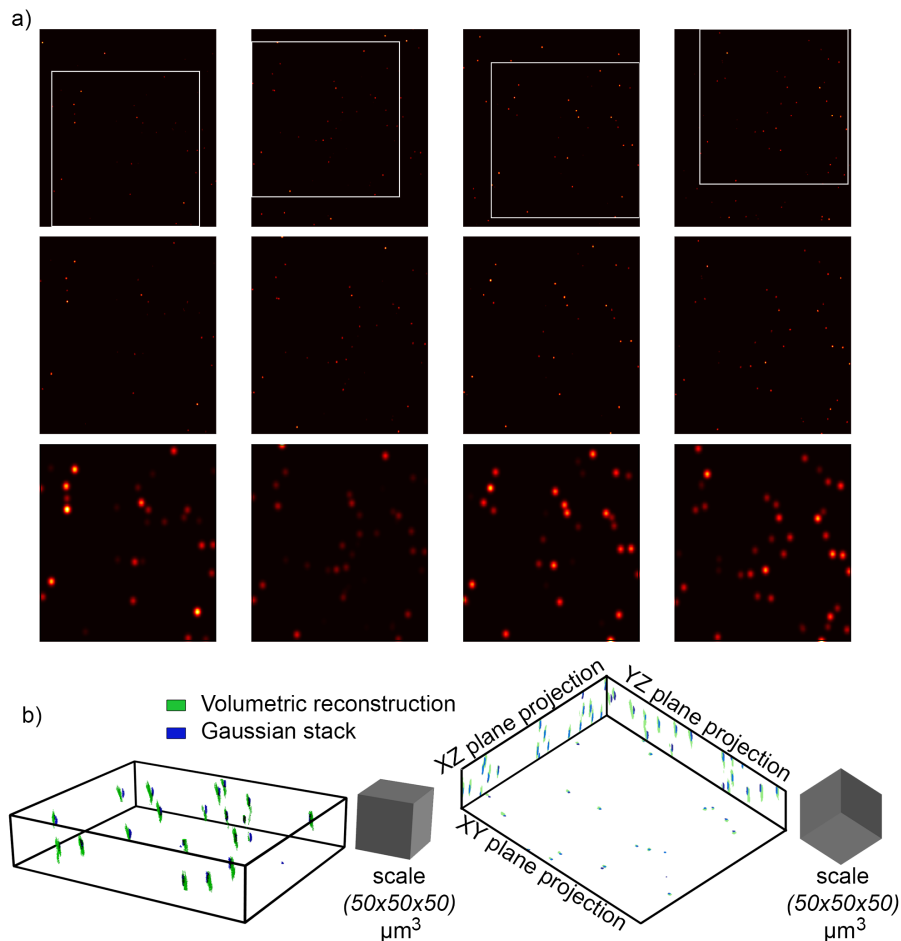
Fig. 4. Tomographic reconstruction of $1 \mu m$ diameter beads. (a) Top row: projections for each Bessel beam of a volumetric sample. The white frames indicate the overlap between the four Bessel beam projections and any bead within this area is visible in all projections. Center row: cropped projections corresponding to the overlapping area (white frames). Bottom row: cropped projections after applying a Gaussian filter with standard deviation of 5 pixels to improve bead visibility. (b) Reconstruction of the beads (using only back projection, no convolutional networks) in green, compared to the volume recorded with the Gaussian beam (ground truth). A single frame required for reconstruction was recorded at the same time in each channel with 512×512 pixels at 30 Hz. As described in the results section, the tomographic reconstruction only shows regions in which all four beams intersect. Due to the offsets between the different projections, some of the beads near the borders of the reference volume are therefore not reconstructed.

Each beam is sent through an axicon (AX252-B, all optical components were from Thorlabs unless otherwise noted) for generating a Bessel beam. The beam is collimated after the axicon with an achromatic lens (AC254-100-B) and the resulting ring is imaged onto the back focal plane of the microscope objective. For this, the ring is first demagnified and imaged onto the scan mirrors using a telescope (AC254-200-B and AC254-100-B), with the latter lens common to all four beam paths. The scan mirror is imaged onto the back focal plane of the objective using a scan lens (AC300-050-B) and tube lens (AC508-300-B).
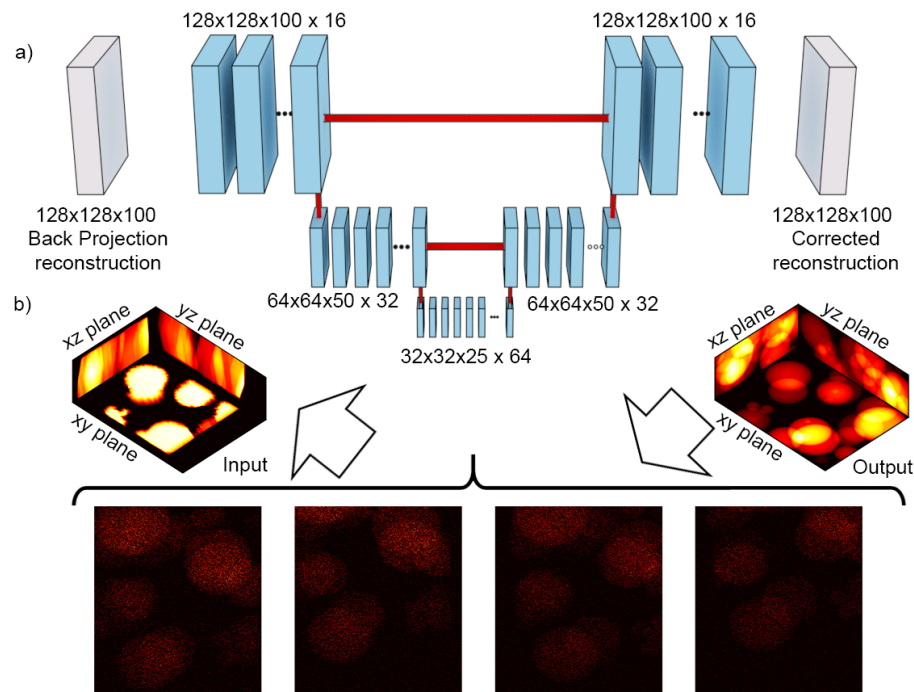
Fig. 5. U-net architecture and simulated data. (a) U-net: each step in the network is composed of 3D convolution, batch normalization and max pooling in 3D in the encoding part while the decoding part is composed of 3D convolution, batch normalization and up sampling in 3D. (b) Simulated data used for training the network. The simulated 3D samples consist of ellipsoids of different sizes and serve as the output of the network. Four projections are generated from each simulated sample volume with addition of shot noise (bottom part). These four projections are produced using the calibration vector fields obtained experimentally. Applying back projection to these four projections yields a volume image. The black region in the reconstructed volume results from the cropping of the projections due to the offset between the different Bessel beam projections. This reconstructed volume serves as the input of the network, while the original volumes serves as the output.

To achieve tilted illumination from four opposing azimuth angles (see Fig. 3), the last lens in each beam path (before the beams are recombined) is displaced laterally and/or vertically for each beam, resulting in an off-center Bessel ring in each quadrant of the back focal plane of the objective. The displacement was in the range of 1-2 mm in the respective direction and was adjusted while observing the projection of the Bessel ring at the back aperture of the objective. The Bessel ring diameters at the back aperture of the objective were slightly below half the objective aperture diameter and allowed laterally displacing the beam without clipping. The four rings projected onto the back aperture of the objective are shown in Fig. 9(c).

Depending on the parameters and methods used to generate Bessel beams, different profiles in terms of width, length and axial intensity distribution result [23]. The axial intensity profile additionally depends on the placement of the collimating lenses [14, 23]. Here, we aimed for a beam length of about 80 μm as a compromise between recording a sufficiently large volume while limiting the power required for two-photon excitation in each beam. The maximum power per beam used in the experiments was 90 mW measured at the entrance aperture of the resonant scanner. In each beam path the collimating lens was placed 11.5 cm behind the axicon, the first lens of the telescope was placed 30 cm behind the collimating lens which was separated by 28
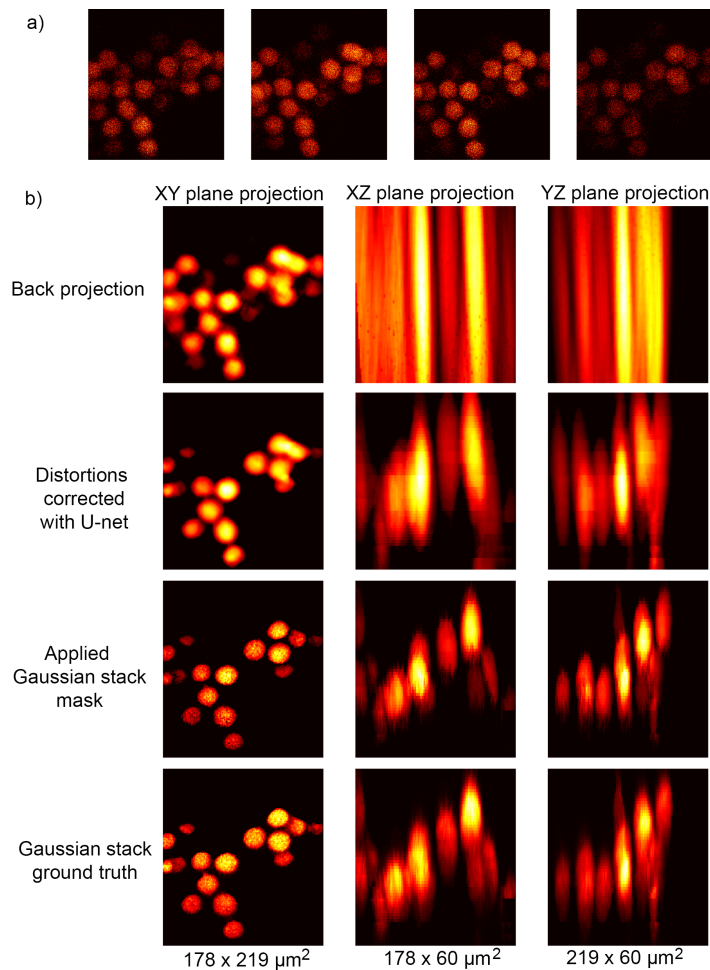
Fig. 6. Tomography on pollen grain sample. (a) Four projections recorded from a 3D sample of pollen grains. (b) Plane projections of the volumetric reconstructions. Top row: back projection shows poor resolution in the z axis since the object approaches the beam length. The second row shows the reconstruction after applying the U-net. The third row additionally applies a mask recorded with the Gaussian beam (see text for details). The bottom row shows the reference Gaussian stack. The dimensions of the images are indicated for each image in the bottom row. A single frame required for reconstruction was recorded at the same time in each channel with 512×512 pixels at 30 Hz.

cm from the lens common to all beam paths. The focal length of each beam was iteratively adjusted based on the two-photon point spread function by changing the beam collimation using the translation stage in the beam expander [14].

The four beams are recombined by first combining pairs of beams - after adjusting their polarizations othogonally with a half-wave plate (not shown) - with a polarizing beam splitter cube. All four beams are then combined with a right angle reflecting prism positioned just before the common telescope lens (that forms the demagnified image of the Bessel ring on the scan mirror). The reflecting prism is positioned laterally into the beam path using a translation stage, reflecting two beams at 90 degrees while allowing the other two beams to pass near the edge of the prism. This is possible without clipping any of the beams by laterally displacing the two

passing beams into the appropriate neighboring quadrants in the objective back focal plane (as required for tilting of the Bessel beams).

An additional Gaussian beam path (not shown) was used for recording reference image z-stacks. A flip mirror was used to direct the beam through a telescope towards a polarizing beam splitter cube mounted on a magnetic mount that was manually inserted before the resonant scanner.

All experiments were performed with a custom-built two-photon microscope controlled with Scanimage (Vidrio) [24] with a resonant scanner (resonant scan box, Sutter Instruments) and components similar to the ones described in [25]. Scanimage allows FPGA-based (National Instruments, PXIe-7975R) photon counting, in our setup at a time resolution of 1.28 GHz (16 time bins between laser pulses at 80 MHz repetition rate). The resonance frequency of the resonant scan mirror was 8 kHz and images were acquired using bidirectional scanning (images are acquired during both movement directions across the sample). Overlaying the four beams resulted in a combined pulse repetition rate of 320 MHz and therefore four time bins per beam. To suppress cross talk between the fluorescence decays excited by the different beams, only three time bins were used for each beam and the intensity in these three bins was averaged into one channel (or image) using the scanning control software. The four different channels were selected with the help of the fluorescence decay time histogram displayed in the scanning control software.

2.2. Tomographic volume reconstruction

A typical 3D tomography approach, for example in medical imaging, acquires 2D projections of a sample by varying the azimuth angle at a fixed elevation of 90 degrees, as illustrated on the left side of Fig. 1(c). Changing the elevation to an oblique angle as shown on the right side of Fig. 1(c) distorts the reconstruction by a factor of $1/\sin(\theta)$. Using tilted Bessel beams to record projections results in a distortion along the z axis and this was (partially) corrected for using machine learning. The workflow of our tomography approach is shown in Fig. 1(b): first, four projections are recorded from a 3D sample by scanning the four different Bessel beams accross the sample, second, back projection is applied resulting in a distorted 3D reconstruction and, and third, a convolutional neural network (3D U-net [26–28]) is used to recover the 3D object. These steps are described in more detail below.

2.2.1. Radon transform for Bessel beam scanning

For volume reconstruction we developed a 2D Radon transform based approach for Bessel beam scanning with tilted beams. In this approach, the amount of light emitted following excitation at a given point $\mathbf{x} \in \mathbb{R}^3$ in a volumetric sample is described with a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$. The Bessel beam is approximated as a line segment of length ℓ with constant intensity centered at a point $\mathbf{x}_c \in \mathbb{R}^3$ and a direction given by the normalized vector $\hat{\mathbf{s}} \in \mathbb{R}^3$ with a parameter t. The total amount of light emitted by the sample due to excitation by the beam is then described in terms of the Radon transform along the segment:

$$[\mathcal{R} f](\mathbf{x_c}, \hat{\mathbf{s}}, \ell) = \int_{-\ell/2}^{\ell/2} f(\mathbf{x_c} + t \cdot \hat{\mathbf{s}}) dt. \tag{1}$$

We define an image of size $(2N+1) \times (2N+1)$ as a function $I : \{\mathbb{Z} \cap [-N, N]\} \times \{\mathbb{Z} \cap [-N, N]\} \rightarrow \mathbb{R}$ representing the intensity of a pixel at a position with integer indices (ν, μ). To create an image, the Bessel beam is scanned with the two scanning mirrors across the field of view and for each pixel the resulting intensity is recorded. When scanning the tilted Bessel beams we observed small changes of both the vector $\mathbf{x}_c$ and the direction vector $\hat{\mathbf{s}}$ across large fields of view compared to a Gaussian reference z-stack, see Fig. 3 and Table 1 for quantification. Both vectors are therefore described as vector fields depending on the pixel coordinates (ν, μ) and axial position z: the offset vector field, $\mathbf{x_c} = \mathbf{x_c}(\nu, \mu, z)$, and the the direction vector field, $\hat{\mathbf{s}} = \hat{\mathbf{s}}(\nu, \mu, z)$. A general expression for the intensity in each pixel is then given by:

$$I(\nu, \mu) = [\mathfrak{R}f](\nu, \mu) = \int_{-\ell/2}^{\ell/2} f\left(\mathbf{x_c}(\nu, \mu) + t \cdot \hat{\mathbf{s}}(\nu, \mu)\right) dt. \tag{2}$$

2.2.2. Calibration of Bessel beam projections

The vector fields $\mathbf{x_c^{(i)}}$ and $\hat{\mathbf{s}}^{(i)}$ are required for reconstructing the volumetric sample $f(\mathbf{x})$ from its projections $I^{(i)}(\nu, \mu; z)$ and are measured experimentally for each projection $I^{(i)}$ with a volumetric sample of 1 μm diameter fluorescent beads (Fluoresbrite, Cat. No. 17154-10, Polysciences) suspended in 2% Agarose (Agarose LE, M 3044, Glenaxxon) (see Fig. 2(a) and Fig. 3). For this, a z-stack was recorded while scanning the four Bessel beams resulting in volumes of 2D projections, $V^{(i)}(\nu, \mu, z) \equiv \{I^{(i)}(\nu, \mu; z)/z \in [z_{min}, z_{max}]\}$. As quantified in Table 1, variations of the vector field in z-direction were small and are ignored in the following.

The recorded volumes were segmented using the mean value after Gaussian filtering with standard deviation of 5 pixels for noise reduction (manually tuned for optimizing segmentation) as threshold, see Fig. 2(a) and (b). The center position of each segmented bead b, $(\nu_c^{(i,b)}, \mu_c^{(i,b)}, z_c^{(i,b)})$ was identified using the *skimage* function *regionprops* and Principal Component Analysis (PCA) was applied to obtain the axis with the highest eigenvalue as the direction vector $\hat{\mathbf{s}}^{(i,b)}$ at the center point $(\nu_c^{(i,b)}, \mu_c^{(i,b)}, z_c^{(i,b)})$ as well as the length of the beam $\ell^{(i,b)}$, see Fig. 2(c) and Fig. 3.

Making use of all beads measured in each z-stack i and the corresponding directional vectors,

$$S^{(i)} \equiv \{\hat{\mathbf{s}}^{(i,b)}(\nu_c^{(i,b)}, \mu_c^{(i,b)}, z_c^{(i,b)})/b = 0, 1, ..., (N_b^{(i)} - 1)\}, \tag{3}$$

we used linear regression to obtain the directional vector field across the entire field of view (ν, μ, z),

$$\hat{\mathbf{s}}^{(i)}(\nu, \mu, z) \approx A_s^{(i)} \cdot (\nu, \mu, z)^T + \mathbf{B_s^{(i)}}, \tag{4}$$

where A_s is a 3×3 matrix while $\mathbf{B_s}$ is a bias vector, which are both parameters fit to the data $S^{(i)}$.

The length of the beam in each z-stack i is considered constant, and it is computed as the average value of the lengths obtained for all beams:

$$\ell^{(i)} = \frac{1}{N_b^{(i)}} \sum_b^{N_b^{(i)}} \ell^{(i,b)}. \tag{5}$$

Due to the selected beam length and collimation of the tilted Bessel beams, the different beams do not meet at the center of the field of view (similar to [14]). This offset leads to cropping of the projections, see Fig. 4(a) top row for full frames and Fig. 4(a) center row for cropped frames (as indicated by the white frames). (An additional cropping is introduced by the offset between reference z-stack recorded with a Gaussian beam used for calibration and the z-stack recorded with the Bessel beam (referred to as a Bessel beam z-stack below).) For tomographic reconstructions, this offset was corrected for by computing an offset vector field. For this, one Bessel beam z-stack was used as a reference. For each centroid of each bead in the segmented Bessel beam z-stack the offset vector to the same bead imaged in the three other segmented Bessel beam z-stacks was computed. Interpolating the offset vectors using linear regression for all three z-stacks results in the offset vector field:

$$\mathbf{x_c^{(i)}}(\nu, \mu, z) \approx A_c^{(i)} \cdot (\nu, \mu, z)^T + \mathbf{B_c^{(i)}}. \tag{6}$$

2.2.3. Back projection

Using the calibrated vector fields $\mathbf{x_c}^{(i)}$, $\hat{\mathbf{s}}^{(i)}$, and length $\ell^{(i)}$ for each projection i, the inverse Radon transform, $\mathfrak{R}^{(-1)}I^{(i)}$, can be applied to any projection $I^{(i)}(\nu, \mu)$, where the pixels of $I^{(i)}$

are projected along the corresponding direction vector $\hat{\mathbf{s}}^{(\mathbf{i})}$ and translated with its corresponding offset vector $\mathbf{x_c^{(i)}}$. This results in a 3D image:

$$\left[\mathfrak{R}^{(-1)}I^{(i)}\right](v, \mu, t) = \begin{cases} I^{(i)}(v, \mu) - \mathbf{x_c^{(i)}}(v, \mu, c_z) - t \cdot \hat{\mathbf{s}}^{(\mathbf{i})}(v, \mu, c_z) & \text{if} \quad t \in \left[-\frac{\ell^{(i)}}{2}, \frac{\ell^{(i)}}{2}\right] \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The parameter t needs to be discretized: $t = t_{min}, (t_{min} + z_{res}), ..., (t_{max} - z_{res}), t_{max}$, where z_{res} gives the z resolution. We obtain an approximated reconstruction of the volume as the sum of the the inverse Radon transforms (back projection):

$$R(v, \mu, t) = \sum_i \left[\mathfrak{R}^{(-1)}I^{(i)}\right](v, \mu, t). \tag{8}$$

For small objects, such as the fluorescent beads in Fig. 4), the back projections were additionally filtered by setting the 3D image to zero if not all four inverse Radon transforms overlapped in a given position (v, μ, t). For this filtering, a binary image was computed from each inverse Radon transform (obtained from equation (7)) by applying a Gaussian filter with 1 pixel standard deviation and thresholding to the mean value of the inverse Radon transform. This was done for each projection, and the binary inverse Radon transforms were then summed to form a thresholded back projection. The maximum value of the thresholded back projection is 4 (only 4 projections). This image was binarized, setting values below 4 to zero, creating a back projection mask with a value of 1 where the four beams intersect. The actual back projection obtained from equation (8) was finally multiplied element-wise with the back projection mask so that the density of the resulting image was non-zero only where the 4 beams intersected. The recorded projections are shown in Fig. 4(a) and the resulting reconstructed volume with only intersecting regions of the four beams as well as a comparison with the Gaussian reference z-stack is shown in Fig. 4(b) (no convolutional networks were applied for this reconstruction).

2.2.4. Neural networks for distortion correction

For objects that are of similar size to the Bessel beam focal length, the back projected volume will extend over the entire axial range of the sampled volume due to the shallow projection angles. Neural networks were used to correct for distortions and to recover axial resolution in this situation, (see Fig. 1(b) and (c)).

A three dimensional U-net (Fig. 5(a)) was trained using simulated 3D samples consisting of ellipsoids of different sizes with homogeneous constant density (ranging from $1\mu m$ to 40 μm) (Fig. 5(b)). Using the Radon transform approach, four projections were generated from each simulated 3D sample using the vector fields measured during calibration and shot noise was added to simulate detection noise (Fig. 5(b), bottom row). From these four projections a (distorted) 3D volume was reconstructed using back projection and a back projection mask (as described above) to obtain reconstructions based on all four beams (Fig. 5(b), left side). This reconstructed volume served as input to a 3D U-net (dimension $128x128x100$ pixels) and the original simulated 3D sample (same dimension of $128x128x100$) served as the output. The number of data samples generated was 1000.

Networks were built and trained using Python and *keras* with *tensorflow* as the backend [29]. The 3D U-net had 804809 parameters and was composed of blocks formed in the first part by 3D convolutional layers, batch normalization and 3D max pooling. After the bottleneck we used a combination of 3D convolutional layers, batch normalization and 3D upsampling layers to deconvolve the 3D image and recover the corrected reconstruction (Fig. 5).

The model was trained on a desktop computer running Ubuntu with an Intel Xeon (R) CPU E5-1650 v4 @ 3.60GHz x 12 processor and a TITAN XP graphics card, that allowed to train the
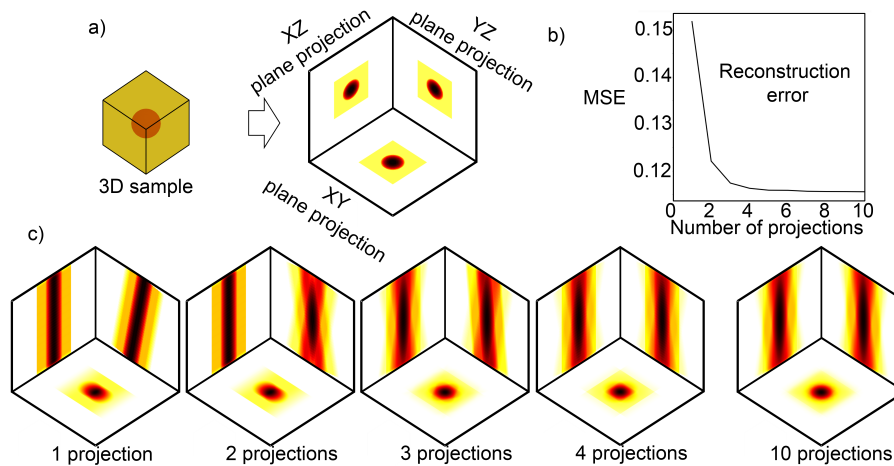
Fig. 7. Simulated reconstruction error of 3D tomography depends on number of projections for the given isolated sample. (a) A homogeneous cube that contains a homogeneous sphere with higher density and its projection in the xy, xz and yz planes (for visualization). (b) Mean Square Error (MSE) between the simulated 3D sample and the reconstructed sample using back projection as a function of the number of recorded projections using a constant elevation angle of 170 degrees. (c) Projections of the reconstructed sample considering 1, 2, 3, 4, and 10 projections, respectively.

network for 100 epochs with batches of 5 samples in about 5 hours. The loss function was the mean squared error between the original simulated image and the network output and we used Adam optimizer to train the neural network. Results of applying this network on a pollen grain volume sample (Mixed Pollen Grains, part number 304264, Carolina) are shown in Fig. 6, where Fig. 6(a) shows the four recorded projections and Fig. 6(b), top row, shows the reconstruction using back projection. The back projection of the pollen grains was again filtered by a back projection mask to base the reconstruction on all four projections. The Gaussian filter of the back projected mask had again 1 pixel standard deviation and the threshold was the mean of the inverse Radon transform of each projection. Due to the (compared with the beam length) large size of the pollen grains, almost no axial resolution is recovered using back projection (Fig. 6(b), top row). However, axial information can be partially recovered with the U-net (Fig. 6(b), second row, compare to Fig. 6(b), last row for ground truth).

For samples which do not change their structure during the observation time but only their intensity a Gaussian reference z-stack could first be recorded for obtaining structural information and Bessel beam tomography for subsequently monitoring activity at high temporal resolution [23]. We tested this approach by generating a structure-mask from the Gaussian z-stack by filtering it in the xy-plane with a Gaussian filter with 1 pixel standard deviation, thresholding it at the mean value, binarizing and downsampling it to 128 pixel resolution by averaging over four pixels. The resulting mask was multiplied with the volume reconstructed using combined back projection and U-net filtering (Fig. 6(b), second row) and resulted in the reconstruction shown in Fig. 6(b), third row. The accuracy of the reconstructed intensities in this case still depends on the quality of the tomographic reconstruction.

3. Discussion

We introduced a tomography method for scanning fluorescence microscopy: volumes were imaged in a single frame scan using temporally multiplexed and tilted Bessel beams for recording

independent projections at four different viewing angles with a single objective. A tomographic reconstruction approach suitable for scanning with tilted Bessel beams was developed combining Radon transforms with three dimensional convolutional networks (3D U-net). We recovered volume information using fluorescent beads (Fig. 4) and pollen grain samples (Fig. 6). The frame rate depended on the resolution of the image and was 30 Hz for images with 512×512 pixel resolution (and increases with fewer pixels, for example 90 Hz for 128 pixel resolution) and the largest reconstructed volumes was $179 \mu m \times 219 \mu m \times 60 \mu m$ (Fig. 6).

Imaging using extended beams, lines or light sheets offers a promising approach for increasing the imaging rate beyond point scanning methods [1–4]. Different from approaches that detect fluorescence using a camera, Bessel beam tomography uses non-descanned PMT detection. This is an advantage for imaging in tissue, since, although scattering affects the infrared excitation light, it does not lead to cross talk of the more strongly scattering visible emission between different detection channels or pixels [3, 9].

The depth of the field of view depends on the length of the Bessel beam and can be adjusted. For for *in vivo* two-photon imaging the power that can be sent into the sample is a limiting factor. Here we used a maximum of 90 mW per beam at a combined pulse rate of 360 MHz. Fluorescence excitation could likely be optimized by using lasers with higher energy per pulse and lower repetition rate without exceeding the damage threshold [4]. Lower repetition rates would also reduce crosstalk between channels and would allow imaging with fluorophores with longer lifetimes, such as typical genetically encoded calcium indicators. Additionally, using higher numerical aperture objectives (we here used 0.8) or using a larger fraction of the available N.A. at the cost of smaller tilt angles could further improve the lateral beam resolution and two-photon excitation efficiency. Bessel beam tomography could also be performed using single photon excitation, where laser power would not be a problem, or using three-photon excitation.

Different from previous approaches with tilted Bessel beams [11, 12, 14] we simultaneously recorded independent projections using temporal multiplexing. The advantage of simultaneous instead of sequential recordings [11,12] lies in the increased frame rate and makes the measurement less sensitive to motion artefacts (which would interfere with the tomographic reconstruction) as typically observed in *in vivo* imaging. Increasing the number of projections to four is advantageous for volume reconstruction at shallow angles as illustrated in Fig. 7 and adding more projections is expected to only add a limited benefit for an isolated object. Different from another recent tomography approach [5], projections here were recorded along the optical axis instead of across the focal plane. While the axial projections at shallow angles come at the cost of lower resolution it is compatible with a standard 2D scanning configuration and doesn't require axial scanning.

For volume reconstruction we used Radon inverse transforms combined with machine learning and adapted it for the presented Bessel beam scanning configuration. We modeled the Bessel beam as a constant line segment for computational simplicity, but models of the actual point spread function could be used for computing the inverse Radon transform.

Approaches combining imaging physics and machine learning have been successfully applied in other tomography techniques for improving the reconstruction from sparse, shallow angle projections [15, 18–22]. Overall, the quality of the reconstruction depends on the neural network model and training data. Our networks were limited by GPU memory and therefore only low resolution images (128 pixel resolution instead of 512) were used. Training the network models on higher resolution images will likely improve the reconstruction. Additionally, we used a limited training data set based on simulated, randomly distributed homogeneous spheres to approximate the pollen grain sample. Training a network on this data set was sufficient for partially reconstructing axial information, but as expected didn't recover higher resolution features of the sample, such as small spikes on the pollen grains. The quality of the reconstruction is expected to improve if the training data is more similar to the sample. Therefore, increasing the training data set to include large numbers of more realistic samples and, if necessary, increasing

the model complexity will allow to further improve the reconstruction. It will be interesting to explore in how far denser and less regular samples can be reconstructed with Bessel beam tomography and suitable training data sets. While for an isolated object four projections suffice for reconstruction (see Fig. 7), for denser samples, the reconstruction would be expected to improve with the number of projections.

Overall, Bessel beam tomography is suitable for fast imaging of sparse volumetric samples that can be accessed only with a single objective, a situation that is commonly encountered when imaging neural activity from cell bodies in the brain.
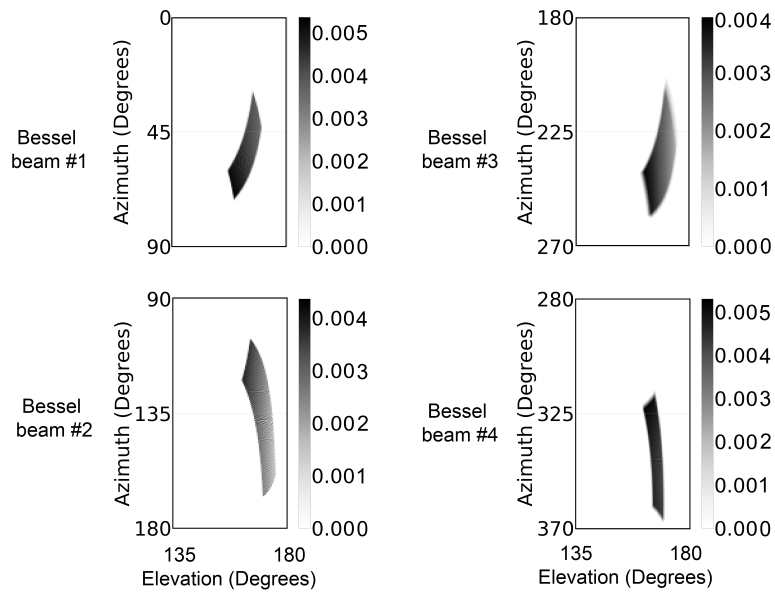
Appendix

Fig. 8. 2D normalized distribution of the elevation (horizontal axis) and azimuth (vertical axis) angles for each Bessel beam. The distribution is computed from the interpolation equation (4), using a total of $N = 10^6$ vectors. See Table 1 for mean values of the distributions.
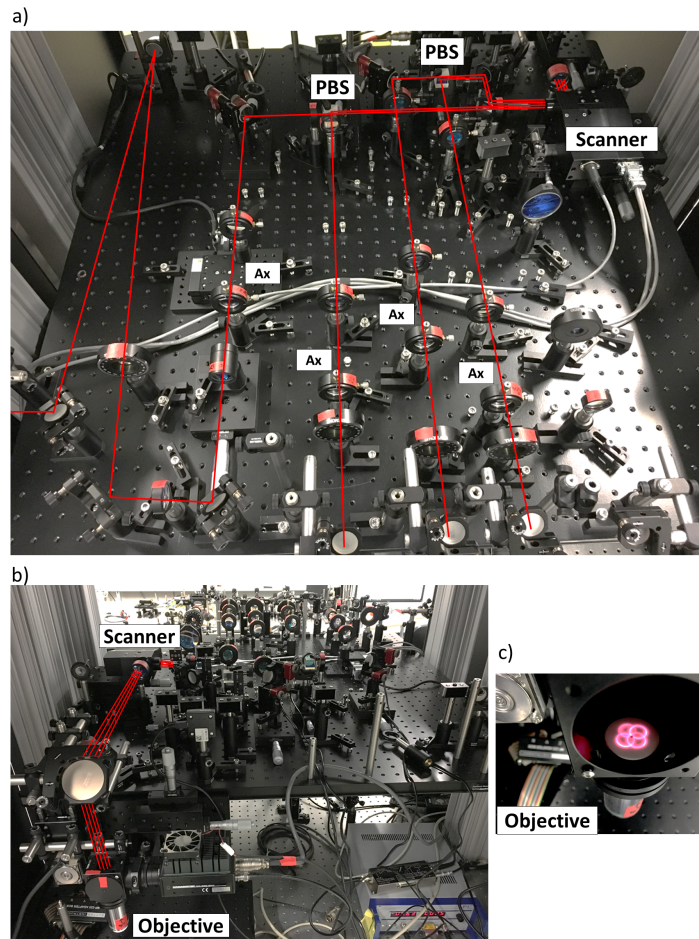
a)

b)

c)

Fig. 9. Pictures of the setup. (a) Top view of beam path before entering the resonant scanner and after the delay lines. Four independent Bessel beams are generated with four axicons and combined with polarizing beams splitters (PBS) and a reflecting prism (to combine all four beams) as shown schematically in Figure 1. Delay lines are not shown. Red lines schematically illustrate the beam paths and the number of parallel lines illustrate the number of combined beams. Rotation mounts contain half-wave plates and are not shown in the schematic in Fig. 1. (b) View of microscope with beam path illustrated between the scanner and the microscope objective, passing through scan lens and tube lens and reflected with a mirror into the objective. (c) Picture of the four Bessel beams at the back aperture of the objective with the laser tuned to 680 nm. A 1 inch diameter alignment disk was placed in the back aperture for taking the picture. The diameter of the back aperture is 20 mm. The back focal plane of the objective (Nikon MRP07220 - CFI-75 LWD 16x/ 0.80/ 3,00 water dipping) is at 42.6mm from the specimen surface. The collimation of each beam was iteratively adjusted to obtain four beams with similar focal length. This results in the four rings having different diameters, which is due to the different initial collimations of the beams resulting from the 1 m additional path lengths of each beam.

Table 1 shows the variation of the elevation and azimuth angles for each Bessel beam across an entire z-stack. For each beam, a total of 10^6 vectors were generated from the linear interpolation equation (4) after calibration, and the mean and standard deviation (std) of the beam angle with respect to the x-axis (azimuth) and z-axis (elevation) were computed for each of these vector

fields. In the third column the change of direction of each Bessel beam with respect to the z axis (over a larger volume than used in experiments) is shown. Generating 10^4 vectors in the top layer ($z = 200\mu m$) and 10^4 vectors in the bottom layer $z = 0\mu m$ from equation (4), the angle between each pair of vectors corresponding to the same pixel locations as well as the mean and standard deviation for all vectors were computed. The mean difference for each beam was close to zero and therefore any z-dependence of the vector fields could be ignored in the reconstruction.

Table 1. Variation of elevation and azimuth angle of Bessel beam vector fields

	Elevation mean ± std (degrees) $N = 10^6$	Azimuth mean ± std (degrees) $N = 10^6$	Top-bottom angle difference mean ± std (degrees) $N = 10^4$
Bessel beam #1	164 ± 3	53 ± 10	0, 54 ± 0.01
Bessel beam #2	169 ± 3	133 ± 15	0, 11 ± 0, 01
Bessel beam #3	168 ± 3	237 ± 12	1, 71 ± 0.11
Bessel beam #4	166 ± 2	310 ± 21	0, 75 ± 0.04

Disclosures

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

References

1. P. J. Keller and M. B. Ahrens, "Visualizing whole-brain activity and development at the single-cell level using light-sheet microscopy," Neuron **85**, 462–483 (2015).
2. J. A. Calarco and A. D. Samuel, "Imaging whole nervous systems: insights into behavior from worms to fish," Nat. Methods **16**, 14 (2019).
3. S. Weisenburger and A. Vaziri, "A guide to emerging technologies for large-scale and whole-brain optical imaging of neuronal activity," Ann. Rev. Neurosci. **41**, 431–452 (2018).
4. E. M. Hillman, V. Voleti, K. Patel, W. Li, H. Yu, C. Perez-Campos, S. E. Benezra, R. M. Bruno, and P. T. Galwaduge, "High-speed 3d imaging of cellular activity in the brain using axially-extended beams and light sheets," Curr. Opin. Neurobiol. **50**, 190–200 (2018).
5. A. Kazemipour, O. Novak, D. Flickinger, J. S. Marvin, J. King, P. Borden, S. Druckmann, K. Svoboda, L. L. Looger, and K. Podgorski, "Kilohertz frame-rate two-photon tomography," bioRxiv p. 357269 (2018).
6. W. Amir, R. Carriles, E. E. Hoover, T. A. Planchon, C. G. Durfee, and J. A. Squier, "Simultaneous imaging of multiple focal planes using a two-photon scanning microscope," Opt. Lett. **32**, 1731–1733 (2007).
7. A. Cheng, J. T. Gonçalves, P. Golshani, K. Arisaka, and C. Portera-Cailliau, "Simultaneous two-photon calcium imaging at different depths with spatiotemporal multiplexing," Nat. Methods **8**, 139 (2011).
8. J. Nylk and K. Dholakia, "Light-sheet fluorescence microscopy with structured light," in *Neurophotonics and Biomedical Spectroscopy,* (Elsevier, 2019), pp. 477–501.
9. F. O. Fahrbach, P. Simon, and A. Rohrbach, "Microscopy with self-reconstructing beams," Nat. Photonics **4**, 780 (2010).

10. G. Thériault, Y. De Koninck, and N. McCarthy, "Extended depth of field microscopy for rapid volumetric two-photon imaging," Opt. Express **21**, 10095–10104 (2013).

11. G. Thériault, M. Cottet, A. Castonguay, N. McCarthy, and Y. De Koninck, "Extended two-photon microscopy in live samples with bessel beams: steadier focus, faster volume scans, and simpler stereoscopic imaging," Front. Cell. Neurosci. **8**, 139 (2014).

12. Y. Yang, B. Yao, M. Lei, D. Dan, R. Li, M. Van Horn, X. Chen, Y. Li, and T. Ye, "Two-photon laser scanning stereomicroscopy for fast volumetric imaging," PLOS ONE **11**, e0168885 (2016).

13. R. Lu, W. Sun, Y. Liang, A. Kerlin, J. Bierfeld, J. D. Seelig, D. E. Wilson, B. Scholl, B. Mohar, M. Tanimoto *et al.*, "Video-rate volumetric functional imaging of the brain at synaptic resolution," Nat. Neurosci. **20**, 620 (2017).

14. A. Song, A. S. Charles, S. A. Koay, J. L. Gauthier, S. Y. Thiberge, J. W. Pillow, and D. W. Tank, "Volumetric two-photon imaging of neurons using stereoscopy (vtwins)," Nat. Methods **14**, 420 (2017).

15. T. G. Feeman, *The mathematics of medical imaging* (Springer, 2010).

16. J. Sharpe, U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sørensen, R. Baldock, and D. Davidson, "Optical projection tomography as a tool for 3d microscopy and gene expression studies," Science **296**, 541–545 (2002).

17. M. Rieckher, U. J. Birk, H. Meyer, J. Ripoll, and N. Tavernarakis, "Microscopic optical projection tomography in vivo," PLOS ONE **6**, e18963 (2011).

18. X. F. Ma, M. Fukuhara, and T. Takeda, "Neural network ct image reconstruction method for small amount of projection data," NUCL INSTRUM METH A **449**, 366–377 (2000).

19. F. Thaler, C. Payer, and D. Štern, "Volumetric reconstruction from a limited number of digitally reconstructed radiographs using cnns," in *Proc. of the OAGM Workshop,* (2018), pp. 13–19.

20. A. Goy, G. Roghoobur, S. Li, K. Arthur, A. I. Akinwande, and G. Barbastathis, "High-resolution limited-angle phase tomography of dense layered objects using deep neural networks," arXiv preprint arXiv:1812.07380 (2018).

21. T. C. Nguyen, V. Bui, and G. Nehmetallah, "Computational optical tomography using 3-d deep convolutional neural networks," Opt. Eng. **57**, 043111 (2018).

22. D. Pelt, K. Batenburg, and J. Sethian, "Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks," J. Imaging **4**, 128 (2018).

23. R. Lu, M. Tanimoto, M. Koyama, and N. Ji, "50 hz volumetric functional imaging with continuously adjustable depth of focus," Biomed. Opt. Express **9**, 1964–1976 (2018).

24. T. A. Pologruto, B. L. Sabatini, and K. Svoboda, "Scanimage: flexible software for operating laser scanning microscopes," Biomed. Eng. Online **2**, 13 (2003).

25. J. D. Seelig, M. E. Chiappe, G. K. Lott, A. Dutta, J. E. Osborne, M. B. Reiser, and V. Jayaraman, "Two-photon calcium imaging from head-fixed drosophila during optomotor walking behavior," Nat. Methods **7**, 535 (2010).

26. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Med. Image. Comput. Comput. Assist. Interv.*, (Springer, 2015), pp. 234–241.

27. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: learning dense volumetric segmentation from sparse annotation," in *Med. Image. Comput. Comput. Assist. Interv.*, (Springer, 2016), pp. 424–432.

28. M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, and E. W. Myers, "Content-aware image restoration: pushing the limits of fluorescence microscopy," Nat. Methods **15**, 1090 (2018).

29. F. Chollet, "Keras," https://keras.io (2015).