



## Optoacoustic imaging at kilohertz volumetric frame rates

ALI ÖZBEK,<sup>1,2</sup> XOSÉ LUÍS DEÁN-BEN,<sup>1</sup> AND DANIEL RAZANSKY<sup>1,2,\*</sup>

<sup>1</sup>Institute for Biological and Medical Imaging (IBMI), Helmholtz Center Munich, D-85764 Neuherberg, Germany

<sup>2</sup>School of Medicine and School of Bioengineering, Technical University of Munich, D-81675 Munich, Germany

\*Corresponding author: dr@tum.de

Received 27 December 2017; revised 3 April 2018; accepted 9 April 2018 (Doc. ID 318376); published 12 July 2018

State-of-the-art optoacoustic tomographic imaging systems have been shown to attain three-dimensional (3D) frame rates of the order of 100 Hz. While such a high volumetric imaging speed is beyond reach for other bio-imaging modalities, it may still be insufficient to accurately monitor some faster events occurring on a millisecond scale. Increasing the 3D imaging rate is usually hampered by the limited throughput capacity of the data acquisition electronics and memory used to capture vast amounts of the generated optoacoustic (OA) data in real time. Herein, we developed a sparse signal acquisition scheme and a total-variation-based reconstruction approach in a combined space–time domain in order to achieve 3D OA imaging at kilohertz rates. By continuous monitoring of freely swimming zebrafish larvae in a 3D region, we demonstrate that the new approach enables significantly increasing the volumetric imaging rate by using a fraction of the tomographic projections without compromising the reconstructed image quality. The suggested method may benefit studies looking at ultrafast biological phenomena in 3D, such as large-scale neuronal activity, cardiac motion, or freely behaving organisms. © 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

**OCIS codes:** (100.6890) Three-dimensional image processing; (110.5120) Photoacoustic imaging; (110.5125) Photoacoustics; (100.0118) Imaging ultrafast phenomena; (320.7080) Ultrafast devices.

<https://doi.org/10.1364/OPTICA.5.000857>

### 1. INTRODUCTION

Imaging dynamics is essential for understanding the complex biology of living organisms. Optoacoustic tomography (OAT) holds significant advantages for monitoring of fast biological events, since the entire imaged volume can be excited with a single pulse of light, thus potentially enabling very high imaging rates ultimately limited by the time-of-flight of the generated ultrasound waves [1]. However, in contrast to previously demonstrated high frame rate two-dimensional (cross-sectional) imaging implementations [1], high-resolution three-dimensional (3D) optoacoustic (OA) imaging implies the acquisition of signals from a much larger number of positions around the imaged object [2–6], thus imposing a practical limitation for the temporal resolution or imaging rate that can be achieved. Indeed, even if sufficient numbers of OA signals are simultaneously collected with a suitable ultrasound array, data transmission, storage, and processing may turn very challenging as the frame rates increase. Recently, volumetric data acquisition (DAQ) and real-time 3D image rendering at 100 Hz frame rates have been demonstrated with high-end OAT systems employing matrix detection arrays and GPU-based processing of the generated data flows in the gigabits per second range [7–9]. With this volumetric imaging speed it became possible to efficiently monitor fast biological

phenomena, such as calcium transients in the brain [10] or cardiovascular dynamics [9,11], whereas spiral volumetric optoacoustic tomography (SVOT) further enabled scaling of temporal resolution with the field of view to attain multi-scale dynamic imaging capabilities [12]. However, some biological events may still occur on a faster time scale. For instance, voltage transients (action potentials) in neurons have durations of the order of milliseconds and the recently developed voltage indicators have enabled both fluorescence- and absorption-based readings of those transients in one or two dimensions [13,14]. In addition, investigation of brain activity patterns is often done in freely behaving animals whose rapid movements may necessitate much faster volumetric imaging rates in order to avoid motion-related artifacts [15].

One possible way to boost the temporal resolution without increasing the average data flow consists of using partial DAQ and reconstruction procedures based on compressed sensing (CS) approaches. CS has attracted vast attention in the past decade, suggesting that it is possible to fully recover a signal from a lower number of samples than required according to the Shannon–Nyquist sampling criterion [16,17]. To this end, tomographic imaging modalities have greatly benefited from CS-based methods by reducing the amount of data needed for accurate reconstructions [18]. In OAT, recent works have showcased reconstructions from partial data using CS [19,20]. In this work, we take this approach

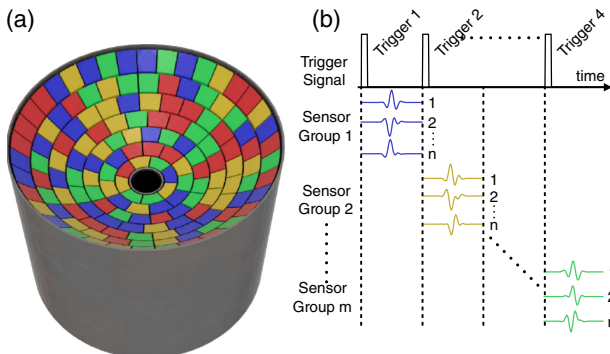
one step further by incorporating a total variation (TV) regularization term in the combined spatio-temporal domain in order to achieve significant acceleration of 3D imaging frame rates with negligible effects on the resulting image quality.

## 2. METHODS

### A. Experimental Setup

The DAQ protocol is schematically depicted in Fig. 1. The OAT system used for real-time acquisition of volumetric data employed a custom-made 30 mm radius spherical matrix transducer array (Imasonic SaS, Voray, France) spanning an angle of  $140^\circ$  ( $1.3\pi$  solid angle) and consisting of 512 piezocomposite elements having a size of  $\sim 2.5$  mm. The 10 MHz central frequency and  $>80\%$  detection bandwidth of the individual detection elements provided an isotropic  $75 \mu\text{m}$  spatial resolution near the center of the spherical acquisition geometry and approximately  $6 \text{ mm} \times 6 \text{ mm} \times 6 \text{ mm}$  effective field of view [21]. Light excitation with  $\sim 10$  ns duration pulses at 532 nm wavelength was provided with a diode-end-pumped Nd:YAG Q-switched laser (Model IS8II-E; EdgeWave GmbH, Würselen, Germany) whose pulse repetition frequency can be set up to 10 kHz. The light beam was guided through a custom-made fiber bundle (Ceramoptec GmbH, Bonn, Germany) providing approximately  $1 \text{ mJ}/\text{cm}^2$  fluence in the imaged area. A dedicated multi-channel DAQ electronics was designed to simultaneously digitize the detected OA signals at a sampling rate of 40 megasamples per second. The DAQ allows for the acquisition of 494 samples per channel per laser pulse. The digitization was delayed by  $13.3 \mu\text{s}$  with respect to the laser excitation pulse in order to efficiently capture signals originating from the imaged area located around the effective field of view of the spherical array, i.e., 30 mm away from its active surface. The average data throughput of the DAQ is limited by its 1 Gbps Ethernet interface. Thus, only a 100 Hz frame rate can be achieved when acquiring data from all 512 channels simultaneously for each laser pulse, corresponding to an average data flow of  $\sim 400$  Mbps including the transmission protocol overhead. As a result, the imaging frame rate can be increased only by reducing the number of channels recorded for each laser pulse, i.e., spatial subsampling of the tomographic data.

The sparse acquisition scheme was facilitated by splitting the detection elements into  $m$  randomly distributed groups of  $n$  elements, as labeled with different colors in Fig. 1(a), with  $m \cdot n = 512$  being the total number of the array sensors. The designed DAQ system is capable of sequential acquisition of signals



**Fig. 1.** Schematic description of the DAQ protocol. (a) Division of sensors into acquisition groups. (b) Temporal division of the acquisition groups.

from each group, triggered with the excitation laser pulses [Fig. 1(b)]. After signals from all 512 channels have been acquired, the same acquisition sequence was repeated. In this manner, the amount of data that needs to be transmitted per each laser pulse is reduced, thus allowing increasing the imaging frame rate by a factor of  $m$  without altering the average data transfer rate.

### B. Image Reconstruction

The TV-based image reconstruction procedure was derived from the recently introduced model-based OA reconstruction approach [22]. The pressure waveforms  $\mathbf{p}_i$  detected by the transducer elements at the  $i$ th acquisition were modeled in a vector form via

$$\mathbf{p}_i = \mathbf{C}_i \mathbf{A} \mathbf{b}_i, \quad (1)$$

where  $\mathbf{b}_i$  is the distribution of the absorbed energy density in a grid of voxels enclosing the sample for the  $i$ th acquisition,  $\mathbf{A}$  is the model matrix derived from the discretization of the OA forward model considering all transducer elements [23], and  $\mathbf{C}_i$  is a selection matrix that picks up only signals of the corresponding group. A uniform Grüneisen parameter was assumed for simplicity [22]. For  $l$  consecutive acquisitions, the full theoretical model is expressed as

$$\mathbf{p} = \mathbf{C} \mathbf{A}_{\text{tot}} \mathbf{b}, \quad (2)$$

where  $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_l)^T$ ,  $\mathbf{C} = \text{diag}(\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_l)$ ,  $\mathbf{A}_{\text{tot}} = \mathbf{I}_l \otimes \mathbf{A}$ ,  $\mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_l)^T$ , and  $\mathbf{I}_l$  is the  $l$  by  $l$  identity matrix. According to Nyquist sampling theory, accurate reconstruction of the full sequence of images  $\mathbf{b}$  would in principle necessitate significantly more information than is actually available through our sparse sampling approach. However, it has been previously shown that sparsity of the solution can be exploited to recover information from fewer samples using CS principles [19]. For this, the signal being recovered must exhibit sparsity in some domain and the sampling matrix must have restricted isometry property (RIP). The latter property is generally hard to prove, but the probability of a randomized matrix to fulfill the RIP condition is known to be relatively high [19]. Since our acquisition electronics does not permit randomizing the temporal sampling space, we attempt to achieve RIP for matrix  $\mathbf{C} \mathbf{A}_{\text{tot}}$  by randomizing the spatial subsampling imposed by the  $\mathbf{C}$  matrix. Although here the RIP property was not rigorously demonstrated, such randomization is known to facilitate attaining RIP conditions [24]. Consequently, the CS-based reconstruction problem is defined as

$$\mathbf{b}_{\text{sol}} = \underset{\mathbf{b}}{\text{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{p}_m - \mathbf{C} \mathbf{A} \mathbf{b} \right\|_2^2 + \lambda \|\mathbf{b}\|_{\text{TV}} \right\}, \quad (3)$$

with  $\mathbf{p}_m$  being the measured pressure signals and  $\lambda$  a regularization parameter.  $\|\mathbf{b}\|_{\text{TV}}$  is the spatio-temporal TV of the image sequence defined as

$$\|\mathbf{b}\|_{\text{TV}} = \left\| \sqrt{\left(\frac{\partial \mathbf{b}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{b}}{\partial y}\right)^2 + \left(\frac{\partial \mathbf{b}}{\partial z}\right)^2 + \left(\frac{\partial \mathbf{b}}{\partial t}\right)^2} \right\|_1, \quad (4)$$

where  $x$ ,  $y$ , and  $z$  are spatial dimensions and  $t$  is the temporal dimension.

The TV norm represents a sparsity-enforcing regularization term that can be efficiently exploited with CS-based methods [17]. Equation (4) further leverages the temporal sparsity of the captured images. Here we created a custom solver for Eq. (3) based on the sub-gradient descent algorithm [25]. The latter was selected owing to its straightforward implementation.

Yet other approaches based on, e.g., accelerated primal dual methods may provide better convergence [26], while more advanced algorithms for reconstructing spatio-temporal data are also available [27,28].

The gradient required for the sub-gradient descent method was expressed analytically as the divergence of the normalized gradient of the image:

$$\nabla_a \|\mathbf{h}\|_{\text{TV}} = \text{div} \left( \frac{\left( \frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z}, \frac{\partial h}{\partial t} \right)}{\sqrt{a^2 + \left( \frac{\partial h}{\partial x} \right)^2 + \left( \frac{\partial h}{\partial y} \right)^2 + \left( \frac{\partial h}{\partial z} \right)^2 + \left( \frac{\partial h}{\partial t} \right)^2}} \right), \quad (5)$$

where  $\text{div}$  is the divergence operator [29] and  $a$  is a smoothing term that was added to the denominator to avoid indefiniteness when the gradient equals 0. This approximation of the gradient  $\nabla_a$  results in a smooth function getting closer to the original gradient as the value of  $a$  approaches 0. The resulting iteration scheme is expressed as

$$h_{k+1} = h_k - \gamma_k \nabla f(h_k), \quad (6)$$

where

$$\nabla f(h) = \mathbf{A}_{\text{tot}}^T \mathbf{C}^T (\mathbf{p}_m - \mathbf{C} \mathbf{A}_{\text{tot}} \mathbf{h}) + \lambda \nabla_a \|\mathbf{h}\|_{\text{TV}}, \quad (7)$$

and the step size for the descent algorithm was determined according to the Barzilai–Borwein method. The algorithm is iterated until a satisfactory image quality is achieved. Application of TV regularization in the temporal domain has recently been implemented in the field of MRI [28]. In our case, weighing of temporal domain gradients against their spatial counterparts was performed via

$$\|\mathbf{h}\|_{\text{TV}} = \left\| \sqrt{w_s \left( \left( \frac{\partial h}{\partial x} \right)^2 + \left( \frac{\partial h}{\partial y} \right)^2 + \left( \frac{\partial h}{\partial z} \right)^2 \right) + w_t \left( \frac{\partial h}{\partial t} \right)^2} \right\|_1. \quad (8)$$

Assuming small frame-to-frame variations in the imaged object, it was heuristically determined that satisfactory image quality is achieved when  $w_s = w_t$ , an assumption that significantly decreases the algorithm's complexity.

Note that due to the high data volumes resulting from the volumetric image acquisition, the suggested approach is not capable of reconstructing long image sequences in a single step—a 16000 frame sequence of 3D images containing  $128 \times 128 \times 64$  single precision voxels would require 312.5 gigabytes of memory. To overcome this memory-related limitation, we retained only the central image as an output from the  $L$  images reconstructed with the spatio-temporal TV-regularized inversion using  $L$  consecutive frames. In the subsequent sections, this method will be referred to as  $\text{TV} \times L$ , with  $L$  representing the size of the temporal window.

The reconstructions were implemented using OpenCL framework and executed in Matlab (MathWorks, Natick, Massachusetts) as a mex function on Nvidia 900 Series Titan X GPU. A single 3D image on a  $128 \times 128 \times 64$  grid was rendered in 50 ms by using the accelerated 3D back-projection (BP) algorithm [30]. The TV-based reconstructions were done using 10 iterations with a custom sub-gradient descent solver using the same image grid. The most computationally extensive matrix–vector multiplications required by the solver were the multiplication with the model matrix ( $\mathbf{A}$ ) and its transpose.

The latter were calculated on the fly using the recently reported GPU implementation [31]. The remaining calculations were done on the Intel i7 3820 CPU. The reconstruction times were 6, 9, 15, 24, and 55 s per individual 3D frame when considering 1, 2, 4, 8, and 16 consecutive acquisitions with 512, 256, 128, 64, and 32 random channels per acquisition, respectively. Note that the total reconstruction time is chiefly conditioned by the number of computationally expensive matrix–vector multiplications, which is in turn proportional to the number of consecutive frames considered in the TV procedure.

### C. Numerical Simulations

Performance and validity of the algorithms were first verified in numerical simulations. For the purpose of assessing the temporal and spatial resolution performance, a  $75 \mu\text{m}$  diameter sphere (approximately the size of the spatial resolution of the system) was simulated moving at various velocities. Simulations were based on analytical calculation of OA signals generated by a sphere having a 3D parabolic absorption profile, for which the generated time-resolved pressure signal can be calculated analytically [32], thus providing a reliable metric for analyzing the system's performance. It was assumed that the sphere was illuminated uniformly in the entire reconstructed 3D volume by a laser pulse whose temporal profile was approximated by a Dirac delta function. The matrix array transducer geometry was approximated by 512 point detectors evenly distributed over the  $140^\circ$  spherical aperture. The individual signals received by each transducer were calculated using a Poisson-type solution to the OA wave generation equation [33]. The detectors were assumed to have 10 MHz central frequency and  $>80\%$  bandwidth. The 512 detected signals were then combined to create a sinogram. An additive Gaussian white noise was added to the sinogram, resulting in SNR of 20 dB. For sparse acquisition, only the active channels were simulated. For each subsampling scheme, a new simulation was created, resulting in a slightly different set of signals due to the added noise. Each 3D image frame was reconstructed on a  $128 \times 128 \times 64$  grid enclosing a  $1 \text{ mm} \times 1 \text{ mm} \times 0.5 \text{ mm}$  volume. For TV-based reconstructions, the  $\lambda$  value in Eq. (3) was selected manually to render the best perceived image.

### D. In vivo Imaging of Zebrafish Larvae

Experiments were performed with five days post-fertilization (dpf) wild-type zebrafish larvae. The reason for selecting zebrafish larvae for experimental demonstration was twofold. First, their swimming speed in water is high enough to account for significant displacements (on a spatial resolution scale) within a few milliseconds. For instance, when the zebrafish tail moves with 100 mm/s velocity, the inter-frame motion may reach 1 mm for a 100 Hz frame rate [34], thus exceeding by more than 1 order of magnitude the point spread function of our imaging system ( $75 \mu\text{m}$ ). Imaging at kilohertz rates becomes then essential to smoothly track the larval motion. Second, their size is small compared to the effective field of view of the imaging system, which allowed efficiently exploiting image sparsity during the reconstruction procedure, as detailed in the next section. During the experiments, the spherical transducer array was pointing upward and the volume between the array and the imaged specimen was filled with agar gel to guarantee good acoustic coupling. A 1 cm diameter cavity was carved into the agar at approximately the center of the spherical geometry, where the larva was allowed to swim.

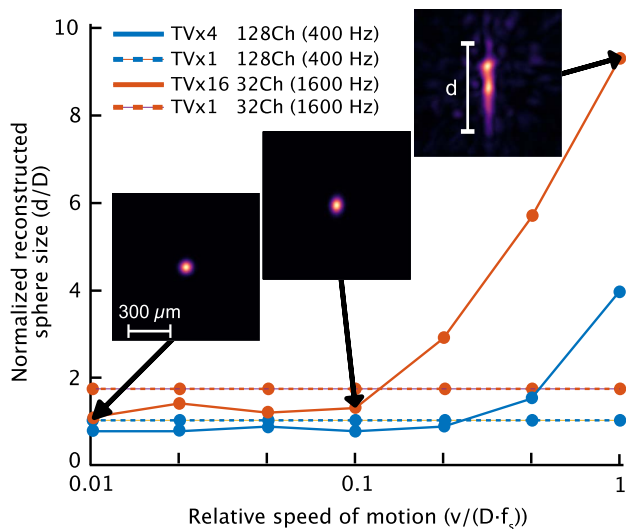


### 3. RESULTS

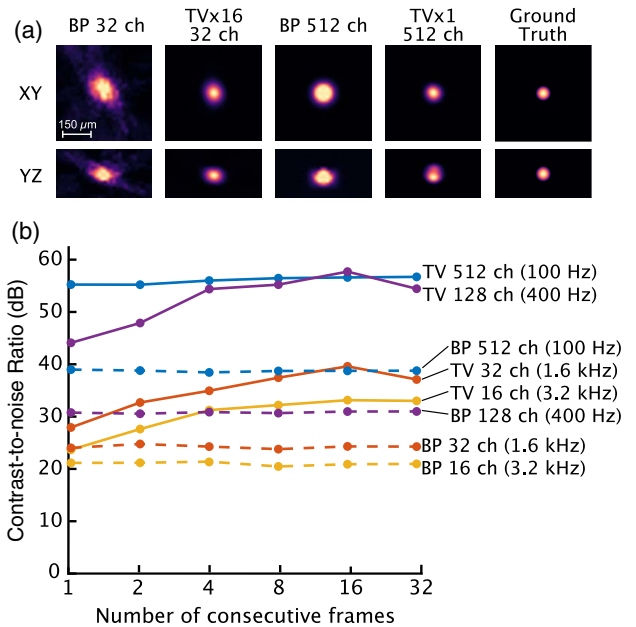
#### A. Numerical Simulations

Performance of the sparse acquisition TV-based reconstruction algorithm was first validated by numerically simulating volumetric imaging of the moving absorbing sphere. Since the algorithm exploits signal sparsity by finding piecewise constant solutions to the reconstruction problem, fast motion would unavoidably result in reduced frame-to-frame sparsity of the data. Furthermore, for rapidly moving absorbers, the given temporal sampling frequency may become insufficient to capture a complete tomographic information. Figure 2 shows dependence of the reconstructed size of the moving sphere (FWHM) versus its normalized (dimensionless) velocity calculated via  $\bar{v} = v/(f_s \cdot D)$ , where  $v$  [m/s] is the actual velocity,  $f_s$  [Hz] is the imaging frame rate, and  $D$  [m] is the size of the sphere. Normalization with  $D$  was necessary to create an independent variable for the motion speed that neutralizes the effect of the physical object size on the spatial resolution performance of the system.

Clearly, for very slow relative velocity values, the TV algorithm is able to recover the correct size of the absorber using a very low number of random channels per acquisition with both TV  $\times 4$  (128 random channels per frame) and TV  $\times 16$  (32 random channels per frame) algorithms performing equally well. As the normalized velocity approaches 1, the reconstructed sphere gets smeared along the dimension of motion with its reconstructed size appearing  $\sim 4$  times longer that its actual diameter when using TV  $\times 4$  and  $\sim 9$  times longer when using the TV  $\times 16$  algorithm. These blurring artifacts are absent in reconstructions where no temporal sparsity is enforced (TV  $\times 1$ ; dashed lines in Fig. 2). This is expected since single frame OA reconstruction is not



**Fig. 2.** Numerical simulations of the TV-based reconstructions of a moving absorbing sphere showcasing spatial resolution degradation as a function of the relative motion speed. The data is presented for the TV algorithm using 16 consecutive frames with 32 random channels per frame (TV  $\times 16$ ), four consecutive frames with 128 random channels per frame (TV  $\times 4$ ), and one frame (TV  $\times 1$ ) with 128 or 32 random channels. The reconstructed relative sphere size, i.e., the reconstructed full width at half-maximum ( $d$ ) for the longest dimension divided by the actual sphere diameter, is plotted against its relative speed of motion measured as the ratio between the inter-frame displacement of the sphere and its diameter.



**Fig. 3.** Comparison between the TV and BP reconstructions for sparsely sampled data. The ground truth image represents an absorbing sphere with 3D parabolic absorption distribution. (a) 3D reconstructions of an absorbing sphere with normalized velocity  $\bar{v} = 0.1$  made with the full 512 channel data (BP 512ch and TV  $\times 1$  512ch) versus 32 random channels (BP 32ch and TV  $\times 16$  32ch). (b) The corresponding CNR of the images reconstructed with the different algorithms plotted against the number of consecutive frames.

affected by motion. However, while the TV  $\times 1$  method introduces less spatial blurring for fast motion speeds as compared with the spatio-temporal TV  $\times 16$  and TV  $\times 4$  methods, it becomes inferior to the latter methods when the motion is relatively slow. This effect is expected to be even more pronounced when the images exhibit less sparsity in the spatial domain. Clearly, spatial blurring can be completely avoided by employing simultaneous sampling with all 512 channels, which, however, results in temporal blurring, as the imaging frame rate is limited to 100 Hz.

In Fig. 3, reconstructions of the absorbing sphere made with the TV algorithm are further compared to the conventional BP algorithm, both for full channel sampling and sparse sampling. In this case, the sphere was moving at a normalized velocity of  $\bar{v} = 0.1$  for all consecutive image frames. While the TV-based reconstruction is able to render a sphere of correct size and shape for both sparse and full sampling, the BP algorithm results in a noisy reconstruction containing artifacts resulting from sub-sampled data when using only 32 random tomographic projections [Fig. 3(a)]. Figure 3(b) further illustrates the dependency of contrast-to-noise-ratio (CNR) on the number of consecutive frames considered during the TV-based reconstruction. For the single frame reconstruction, the TV-based approach attains a better CNR as compared to BP for the same number of channels. As expected, the CNR increases as the number of consecutive image frames in the TV inversion increases. Note that combining multiple frames of a moving object into a single BP reconstruction would readily result in spatial blurring, thus only single frame BP reconstructions were considered in the CNR plot.

## B. Imaging of Freely Swimming Zebrafish Larvae

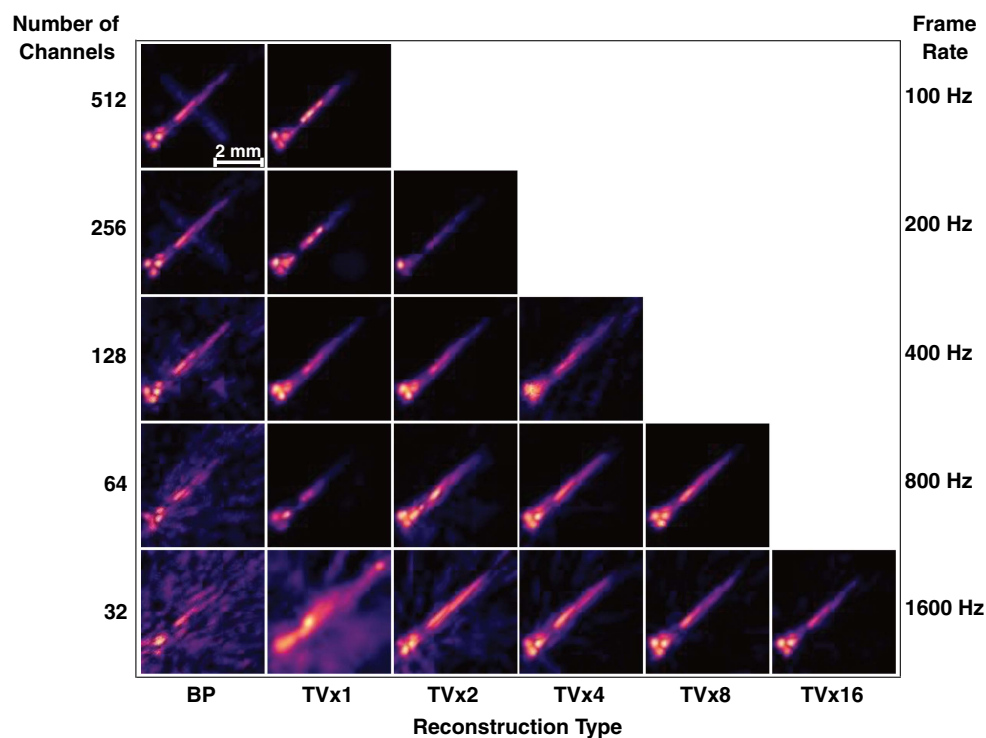
The reconstructed 3D images of the swimming zebrafish larva as a function of the number of random tomographic projections (detection channels) per frame are displayed in Fig. 4. The leftmost column showcases the images reconstructed with the BP algorithm, whose image quality significantly deteriorates for partial DAQ, with the larva barely visible when acquiring with groups of 32 channels per laser pulse (fifth row). The maximum frame rate that can be achieved for the given average data transmission limit of the system is inversely proportional to the number of channels acquired per frame (laser pulse), as indicated on the vertical axes in Fig. 4. The images obtained with the newly developed CS-based method are showcased in the other columns of Fig. 4 as a function of the number of acquisitions employed for the reconstruction. The reconstructions in each column were performed by considering  $L$  consecutive frames (denoted as  $TV \times L$ ). The image artifacts are readily mitigated even when using the single-frame TV algorithm for the partial data reconstructions ( $TV \times 1$  in Fig. 4), which corresponds to TV regularization only in the spatial domain. Indeed, high contrast and good spatial resolution images are achieved when at least 64 random channels are used with  $TV \times 1$ . Image quality further improves once spatio-temporal TV regularization is used. For example, the shape of the larva can be accurately reconstructed when considering two consecutive acquisitions with 32 random channels each ( $TV \times 2$ ). Image contrast and overall quality for a given number of channels per group increase with the number of frames used in the CS-based reconstruction, indicating the importance of TV regularization in the spatio-temporal domain.

A 3D movie of the swimming larva generated with the  $TV \times 16$  algorithm is available as [Visualization 1](#). Note that the video is displayed at 30 frames per second for a better visualization of the motion, while the actual 3D acquisition frame rate is 1.6 KHz. It can be clearly seen that the suggested CS approach is able to fully recover temporal information without introducing spatial blurring, as long as the images are sparse in the spatio-temporal domain. Each video frame was generated by simultaneously reconstructing a stack of  $L$  frames and then selecting the middle frame for display. Note that isotropic TV regularization was used, which resulted in a smooth video without jumps or frozen segments.

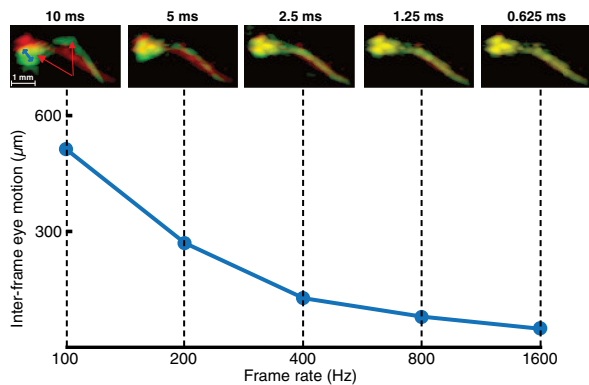
To further showcase the spatio-temporal resolution performance of the method, Fig. 5 displays two consecutive images of the swimming larva (in red and green) during an abrupt turn. The image sequence was reconstructed using TV regularization with different acquisition frame rates ranging from 100 Hz to 1.6 kHz, while the corresponding optimal number of channels per acquisition group ranged from 512 to 32. While a significant shift of the larva's head and tail is clearly visible for images acquired at 100 Hz acquisition rate, the two consecutive images progressively overlap as the frame rate increases. Indeed, while the inter-frame eye motion accounts for 512  $\mu\text{m}$  at 100 Hz, it diminishes to 50  $\mu\text{m}$  at 1.6 kHz, i.e., below the effective spatial resolution of the system. Thus, kilohertz frame rates are essential to accurately track the fish motion.

## 4. DISCUSSION AND CONCLUSIONS

The presented results demonstrate basic feasibility of OA imaging at an unprecedented 1.6 kHz volumetric frame rate.



**Fig. 4.** Comparison between 3D BP and the proposed TV-based CS reconstruction. Maximal intensity projection (MIP) images are shown. The vertical axis represents the acquisition frame rate, which is inversely proportional to the number of random detector positions used per frame (single laser pulse). Different reconstruction methods are shown on the horizontal axis.  $TV \times L$  corresponds to the TV-based reconstruction employing  $L$  consecutive frames.



**Fig. 5.** Spatio-temporal resolution performance of the TV-based CS approach. Two consecutive images of a swimming larva rendered at different frame rates are superimposed in red and green. While a significant shift in the larva position is clearly visible for frames recorded at 100 Hz acquisition rate, the two consecutive images progressively overlap as the frame rate increases. The inter-frame motion of the left eye, marked with blue arrows, is plotted against the frame rate. The red arrows indicate parts of the fish that moved significantly between the two consecutive frames.

The proposed fast sparse acquisition imaging method was shown to capture accurate images without loss of spatial resolution when the inter-frame speed of object motion remained significantly below the available (diffraction-limited) spatial resolution of the system. For instance, when sparsely sampling 32 channels per frame and using the TV  $\times 16$  algorithm for image reconstruction, the normalized object velocity  $\bar{v}$  should remain below 0.1 (Fig. 2). In other words, for the 75  $\mu\text{m}$  resolution system running at a volumetric frame rate of 1.6 kHz, the object motion between the consecutive image frames should remain below  $0.1 \cdot 75 \mu\text{m} \cdot 1.6 \text{ kHz} = 12 \text{ mm/s}$  in order to avoid degradation of the spatial resolution. For a denser sampling with 128 channels and reconstruction with TV  $\times 4$ , the  $\bar{v}$  value can be increased to 0.2, leading to a maximal permissible object velocity of 24 mm/s provided the spatial and temporal resolution parameters remain unchanged.

The ultrafast volumetric imaging capability introduced in this work can greatly benefit applications involving tracking of functional activity in freely behaving organisms [35]. Furthermore, in cardiac imaging the heart rate may reach 600 beats per minute in small animals, thus accounting for significant motion within several milliseconds. Thus, 3D imaging at kilohertz rates can facilitate accurate characterization of the cardiac cycle and valve motion [11], detection of functional alterations due to arrhythmias, and other cardiac diseases [36]. Since blood flow may reach significant velocities in major vessels, fast imaging frame rates can be used to facilitate more accurate spectral unmixing of oxygenation status and perfusion [37]. The methodology introduced in this work may further serve applications involving imaging and tracking of circulating cells or other extrinsically administered agents, and facilitate enhancement of the spatial resolution beyond the diffraction limit [38].

The ultrafast 3D imaging capacity is not only limited to investigations involving physical motion but may equally find use in other applications where visualization of (sub-) millisecond biological dynamics is of importance. One particular example is the field of functional neuroimaging with large-scale brain activity

typically occurring on multiple temporal and spatial scales. While imaging the relatively slow hemodynamic processes remains an important application of OA imaging in neuroscience [8], functional OA neuro-tomography has been recently shown capable of detecting rapid millisecond-scale dynamics using calcium and voltage-sensitive indicators [14,21].

In conclusion, the newly introduced 3D OA imaging capacity at unprecedented kilohertz rates holds promise to provide new insights into biological function not attainable with the existing bio-imaging techniques.

**Funding.** H2020 European Research Council (ERC) (ERC-2015-CoG-682379); National Institutes of Health (NIH) (R21-EY026382).

## REFERENCES

1. K. Sivasubramanian and M. Pramanik, "High frame rate photoacoustic imaging at 7000 frames per second using clinical ultrasound system," *Biomed. Opt. Express* **7**, 312–323 (2016).
2. R. A. Kruger, C. M. Kuzmiak, R. B. Lam, D. R. Reinecke, S. P. Del Rio, and D. Steed, "Dedicated 3d photoacoustic breast imaging," *Med. Phys.* **40**, 113301 (2013).
3. A. P. Jathoul, "Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter," *Nat. Photonics* **9**, 239–246 (2015).
4. S. Ermilov, "Three-dimensional optoacoustic and laser-induced ultrasound tomography system for preclinical research in mice: design and phantom validation," *Ultrason. Imaging* **38**, 77–95 (2016).
5. J. Aguirre, "Precision assessment of label-free psoriasis biomarkers with ultra-broadband optoacoustic mesoscopy," *Nat. Biomed. Eng.* **1**, 0068 (2017).
6. T. F. Fehm, X. L. Deán-Ben, S. J. Ford, and D. Razansky, "In vivo whole-body optoacoustic scanner with real-time volumetric imaging capacity," *Optica* **3**, 1153–1159 (2016).
7. A. Buehler, M. Kacprowicz, A. Taruttis, and V. Ntziachristos, "Real-time handheld multispectral optoacoustic imaging," *Opt. Lett.* **38**, 1404–1406 (2013).
8. S. Gottschalk, T. F. Fehm, X. L. Deán-Ben, V. Tsytsarev, and D. Razansky, "Correlation between volumetric oxygenation responses and electrophysiology identifies deep thalamocortical activity during epileptic seizures," *Neurophotonics* **4**, 011007 (2017).
9. L. Li, "Single-impulse panoramic photoacoustic computed tomography of small-animal whole-body dynamics at high spatiotemporal resolution," *Nat. Biomed. Eng.* **1**, 0071 (2017).
10. X. L. Deán-Ben, S. Gottschalk, G. Sela, S. Shoham, and D. Razansky, "Functional optoacoustic neuro-tomography of calcium fluxes in adult zebrafish brain in vivo," *Opt. Lett.* **42**, 959–962 (2017).
11. X. L. Deán-Ben, S. J. Ford, and D. Razansky, "High-frame rate four dimensional optoacoustic tomography enables visualization of cardiovascular dynamics and mouse heart perfusion," *Sci. Rep.* **5**, 10133 (2015).
12. X. L. Deán-Ben, T. F. Fehm, S. J. Ford, S. Gottschalk, and D. Razansky, "Spiral volumetric optoacoustic tomography visualizes multi-scale dynamics in mice," *Light: Sci. Appl.* **6**, E1624710 (2017).
13. Y. Gong, "High-speed recording of neural spikes in awake mice and flies with a fluorescent voltage sensor," *Science* **350**, 1361–1366 (2015).
14. B. Rao, R. Zhang, L. Li, J.-Y. Shao, and L. V. Wang, "Photoacoustic imaging of voltage responses beyond the optical diffusion limit," *Sci. Rep.* **7**, 2560 (2017).
15. P. Symvoulidis, A. Lauri, A. Stefanou, M. Cappetta, S. Schneider, H. Jia, A. Stelzl, M. Koch, C. C. Perez, A. Myklatun, S. Renninger, A. Chmyrov, T. Lasser, W. Wurst, V. Ntziachristos, and G. G. Westmeyer, NeuTracker—imaging neurobehavioral dynamics in freely behaving fish, *Nat. Method* **14**, 1079–1082 (2017).
16. D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
17. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**, 489–509 (2006).

18. C. G. Graff and E. Y. Sidky, "Compressive sensing in medical imaging," *Appl. Opt.* **54**, C23–C44 (2015).
19. M. Sandbichler, F. Kraher, T. Berer, P. Burgholzer, and M. Haltmeier, "A novel compressed sensing scheme for photoacoustic tomography," *Siam J. Appl. Math.* **75**, 2475–2494 (2015).
20. S. Arridge, "Accelerated high-resolution photoacoustic tomography via compressed sensing," *Phys. Med. Biol.* **61**, 8908–8940 (2016).
21. X. L. Deán-Ben, G. Sela, A. Lauri, M. Kneipp, V. Ntziachristos, G. G. Westmeyer, S. Shoham, and D. Razansky, "Functional optoacoustic neuro-tomography for scalable whole-brain monitoring of calcium indicators," *Light: Sci. Appl.* **5**, e16201 (2016).
22. L. Ding, X. L. Deán-Ben, and D. Razansky, "Efficient three-dimensional model-based reconstruction scheme for arbitrary optoacoustic acquisition geometries," *IEEE Trans. Med. Imaging* **36**, 1858–1867 (2017).
23. G. Paltauf, J. A. Viator, S. A. Prael, and S. L. Jacques, "Iterative reconstruction algorithm for optoacoustic imaging," *J. Acoust. Soc. Am.* **112**, 1536–1544 (2002).
24. R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.* **28**, 253–263 (2008).
25. N. Z. Shor, *Minimization Methods for Non-differentiable Functions* (Springer, 2012).
26. A. Chambolle and T. Pock, "An introduction to continuous optimization for imaging," *Acta Numer.* **25**, 161–319 (2016).
27. M. Holler and K. Kunisch, "On infimal convolution of TV-type functionals and applications to video and image reconstruction," *SIAM J. Imaging Sci.* **7**, 2258–2300 (2014).
28. M. Schloegl, M. Holler, A. Schwarzl, K. Bredies, and R. Stollberger, "Infimal convolution of total generalized variation functionals for dynamic MRI," *Magn. Reson. Med.* **78**, 142–155 (2017).
29. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D Nonlinear Phenom.* **60**, 259–268 (1992).
30. X. L. Deán-Ben, A. Ozbek, and D. Razansky, "Volumetric real-time tracking of peripheral human vasculature with GPU-accelerated three-dimensional optoacoustic tomography," *IEEE Trans. Med. Imaging* **32**, 2050–2055 (2013).
31. L. Ding, X. L. Deán-Ben, and D. Razansky, "Real-time model-based inversion in cross-sectional optoacoustic tomography," *IEEE Trans. Med. Imaging* **35**, 1883–1891 (2016).
32. D. Queirós, X. L. Deán-Ben, A. Buehler, D. Razansky, A. Rosenthal, and V. Ntziachristos, "Modeling the shape of cylindrically focused transducers in three-dimensional optoacoustic tomography," *J. Biomed. Opt.* **18**, 076014 (2013).
33. A. Rosenthal, D. Razansky, and V. Ntziachristos, "Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography," *IEEE Trans. Med. Imaging* **29**, 1275–1285 (2010).
34. X. L. Deán-Ben, H. López-Schier, and D. Razansky, "Optoacoustic micro-tomography at 100 volumes per second," *Sci. Rep.* **7**, 6850 (2017).
35. P. Symvoulidis, A. Lauri, A. Stefanoiu, M. Cappetta, S. Schneider, H. Jia, A. Stelzl, M. Koch, C. C. Perez, A. Myklatun, and S. Renninger, "NeuTracker—Imaging neurobehavioral dynamics in freely behaving larva," *Nat. Methods* **14**, 1079–1082 (2017).
36. H. C. Lin, X. L. Deán-Ben, M. Kimm, K. Kosanke, H. Haas, R. Meier, F. Lohöfer, M. Wildgruber, and D. Razansky, "Non-invasive volumetric optoacoustic imaging of cardiac cycles in acute myocardial infarction model in real-time," *Theranostics* **7**, 4470–4479 (2017).
37. A. Taruttis, E. Herzog, D. Razansky, and V. Ntziachristos, "Real-time imaging of cardiovascular dynamics and circulating gold nanorods with multi-spectral optoacoustic tomography," *Opt. Express* **18**, 19592–19602 (2010).
38. X. L. Deán-Ben and D. Razansky, "Localization optoacoustic tomography," *Light Sci. Appl.* **7**, 18004 (2018).