**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

Spatial Compounding of Volumetric Data Enables Freehand Optoacoustic Angiography of Large-Scale Vascular Networks

Knauer, Nikolaus ; Dean-Ben, Xose Luis ; Razansky, Daniel

Abstract: Optoacoustic tomography systems have attained un-precedented volumetric imaging speeds, thus enabling insights into rapid biological dynamics and marking a milestone in the clinical translation of this modality. Fast imaging performance often comes at the cost of limited field-of-view, which may hinder potential applications looking at larger tissue volumes. The imaged field-of-view can potentially be expanded via scanning and using additional hardware to track the position of the imaging probe. However, this approach turns impractical for high-resolution volumetric scans performed in a freehand mode along arbitrary trajectories. We have developed an accurate framework for spatial compounding of time-lapse optoacoustic data. The method exploits the frequency-domain properties of vascular networks in optoacoustic images and estimates the relative motion and orientation of the imaging probe. This allows rapidly combining sequential volumetric frames into large area scans without additional tracking hardware. The approach is universally applicable for compounding volumetric data acquired with calibrated scanning systems but also in a freehand mode with up to six degrees of freedom. Robust performance is demonstrated for whole-body mouse imaging with spiral volumetric optoacoustic tomography and for freehand visualization of vascular networks in humans using volumetric imaging probes. The newly introduced capability for angiographic observations at multiple spatial and temporal scales is expected to greatly facilitate the use of optoacoustic imaging technology in pre-clinical research and clinical diagnostics. The technique can equally benefit other biomedical imaging modalities, such as scanning fluorescence microscopy, optical coherence tomography or ultrasonography, thus optimizing their trade-offs between fast imaging performance and field-of-view.

DOI: <https://doi.org/10.1109/TMI.2019.2945297>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-175732>

Journal Article

Accepted Version

Originally published at:

Knauer, Nikolaus; Dean-Ben, Xose Luis; Razansky, Daniel (2020). Spatial Compounding of Volumetric Data Enables Freehand Optoacoustic Angiography of Large-Scale Vascular Networks. *IEEE Transactions on Medical Imaging*, 39(4):1160-1169.

DOI: <https://doi.org/10.1109/TMI.2019.2945297>

Spatial Compounding of Volumetric Data Enables Freehand Optoacoustic Angiography of Large-Scale Vascular Networks

Nikolaus Knauer , Xosé Luís Deán-Ben, and Daniel Razansky

Abstract — Optoacoustic tomography systems have attained unprecedented volumetric imaging speeds, thus enabling insights into rapid biological dynamics and marking a milestone in the clinical translation of this modality. Fast imaging performance often comes at the cost of limited field-of-view, which may hinder potential applications looking at larger tissue volumes. The imaged field-of-view can potentially be expanded via scanning and using additional hardware to track the position of the imaging probe. However, this approach turns impractical for high-resolution volumetric scans performed in a freehand mode along arbitrary trajectories. We have developed an accurate framework for spatial compounding of time-lapse optoacoustic data. The method exploits the frequency-domain properties of vascular networks in optoacoustic images and estimates the relative motion and orientation of the imaging probe. This allows rapidly combining sequential volumetric frames into large area scans without additional tracking hardware. The approach is universally applicable for compounding volumetric data acquired with calibrated scanning systems but also in a freehand mode with up to six degrees of freedom. Robust performance is demonstrated for whole-body mouse imaging with spiral volumetric optoacoustic tomography and for freehand visualization of vascular networks in humans using volumetric imaging probes. The newly introduced

capability for angiographic observations at multiple spatial and temporal scales is expected to greatly facilitate the use of optoacoustic imaging technology in pre-clinical research and clinical diagnostics. The technique can equally benefit other biomedical imaging modalities, such as scanning fluorescence microscopy, optical coherence tomography or ultrasonography, thus optimizing their trade-offs between fast imaging performance and field-of-view.

Key words — Angiographic imaging, animal models and imaging, optoacoustic imaging, motion compensation and analysis, registration, vessels.

Optoacoustic imaging and tomography capitalizes on light-based ultrasound generation, thus enabling non-invasive visualization of optical contrast with high resolution at centimeter-scale depths, far beyond the penetration limits imposed by photon scattering in biological tissues [1]–[3]. The inherent spectroscopic discrimination capacity of multi-spectral optoacoustic tomography (MSOT) further provides unique functional and molecular imaging capabilities by resolving distinctive spectral signatures of intrinsic tissue chromophores, genetic and synthetic labels deep in living tissues [4]–[6]. Hence, the method is growingly being used by the biomedical community to facilitate progress in a variety of fields, from cardiovascular research and neuroimaging to dermatology and cancer research [7]–[10].

Recent technological developments have further enabled significantly increasing the frame rate of two- (2D) and three-dimensional (3D) OA imaging based on dedicated transducer array designs [11]. For instance, real-time 2D imaging has been achieved via simultaneous acquisition of OA signals with concave transducer arrays [12]–[14]. Furthermore, spherical matrix array probes have been developed for volumetric (3D) imaging [15]–[17]. Owing to their broad tomographic coverage, those dedicated array geometries enable rendering significantly better OA image quality in comparison with implementations relying on standard linear arrays commonly employed in pulse-echo ultrasonography. However, the main limitation of concave array geometries is their restricted field-of-view (FOV) that can be covered in each image frame, typically scaling with the number of simultaneously acquired

tomographic projections [18]–[20]. The FOV can be enlarged by scanning the detection array at the expense of reduced imaging frame rate. For instance, whole-body imaging of mice has been achieved by translating a cylindrically-focused transducer array along the elevation direction [13], [21], [22]. A different approach termed spiral volumetric optoacoustic tomography (SVOT) has employed scanning of a spherical detection array along a helical trajectory around the animal. In this way, an unprecedented 3D image quality at the whole-body level has been demonstrated by effectively maximizing the isotropic tomographic coverage [16]. Furthermore, the temporal resolution of 3D imaging can be scaled with the FOV, thus offering new ways to study multi-scale biological dynamics [23]. Finally, high-resolution microscopic observations over large FOVs can be enabled by raster scanning of single or arrayed high-frequency transducers [24]–[26].

The availability of real-time imaging capacity has further facilitated clinical translation of the OA imaging technology [17], [27], [28]. Due to the limited penetration of light into living tissues, freehand scanning remains the most convenient approach for this purpose. However, many potential OA human applications may involve imaging of large areas, e.g. extending over the entire female breast [19], [29], the carotid artery [30], [31] or the prostate [32], [33]. In this scenario, registration of multiple image frames is challenged by the arbitrariness of freehand motion. One potential solution consists in using radio-frequency, optical or inertial measurement unit (IMU) systems to track motion [34], [35]. However, those are unable to deliver sufficient translational and rotational estimation accuracy in 3D in order to match typical spatial resolutions in the 100 μ m range attained by state-of-the-art freehand OA systems [34], [36].

Alternatively, motion of a freehand probe between consecutive frames can potentially be estimated from the imaging data. As opposed to cross-sectional (2D) imaging approaches, clear vascular patterns can be usually identified in volumetric OA images acquired by the spherical-array-based (3D) probes. In this work, we developed a framework for spatial compounding of time-lapse OA data acquired using large-area volumetric scans. The method takes advantage of the frequency-domain properties of vascular networks in OA images, which allows for optimally combining sequential volumetric frames into large area scans without using additional tracking hardware.

I. MATERIALS AND METHODS

A. Real-Time Volumetric Optoacoustic Tomography

OA images were acquired with a spherical matrix ultrasound array schematically depicted in Fig. 1a and described in detail in [15]. Briefly, the 8 cm diameter spherical array consists of 256 densely distributed piezocomposite elements having 4 MHz central frequency and $\sim 100\%$ (half width at half maximum) detection bandwidth, resulting in an effective point spread function (spatial resolution) of 200 μ m around the center of the spherical array geometry. Optical excitation was provided with an optical parametric oscillator (OPO) laser guided with a fiber bundle through a central cavity of the array. The laser further provides fast wavelength tunability

Fig. 1. Fourier-based 3D motion estimation. (a) Lay-out of the 3D optoacoustic tomography probe used in the experimental study, indicating the 6 degrees of freedom corresponding to an arbitrary freehand motion. (b) 2D rotation estimation is based on the maximum intensity projections (MIPs) of the volumetric image frames (step 1), which are reduced to their brightest pixels (step 2). After applying Fourier transform (step 3) and resolution enhancement (step 4), the spectra are correlated for different rotation angles, yielding a RCF for x-, y- and z-axis. Rotation parameters are extracted after least squares approximation (LSA) of the RCF (dashed line). (c) Translation estimation based on two consecutive frames i_1 in blue and i_2 in red, respectively (step 1). The frames are reduced to their relevant structures, represented by the brightest voxels (step 2) before the PoC is computed (step 3, equations. 8 and 9). The translation parameters are finally extracted after Gaussian noise suppression (step 4).

between 680 and 950 nm on a per pulse basis and an additional 1064 nm beam output. A custom-made data acquisition system, triggered with the Q-switch output of the laser, was used to digitize the detected individual pressure waveforms at 40 megasamples/s and simultaneously transfer them via a 1 Gbps Ethernet connection to PC for further processing and storage. The acquired signals were deconvolved with the electric impulse response of the piezoelectric detection elements, band-pass filtered with cut-off frequencies of 0.1 and 6.0 MHz and subsequently processed with a back-projection reconstruction procedure, as described in [37]. A volume of 12 mm $\times$ 12 mm $\times$ 12 mm containing 120 $\times$ 120 $\times$ 120 voxels was reconstructed for each acquired frame.

B. Fourier-Based Rotation Estimation

The relative position and orientation of consecutive 3D images acquired with a real-time OA scanner following

an arbitrary scanning trajectory can in a general form be expressed as superposition of translations and rotations. One may represent the translation by a vector $\mathbf{t} = [t_x, t_y, t_z]^T$ describing the displacements in x-, y- and z-direction, whereas the rotation may be represented by a vector $\boldsymbol{\theta} = [\theta_x, \theta_y, \theta_z]^T$ describing the rotation angles around the corresponding axes (Fig. 1a). This yields a total of six degrees of freedom for any arbitrary motion.

Rotation among subsequent acquisitions can be estimated by cross-correlating rotated Fourier spectra and selecting the rotation angle that attains the best fit, which is based on the PROPELLER method used in MRI and generalized for 3D images and rotations around more than one axis [38]. Let $i_1(x,y,z)$ and $i_2(x,y,z)$ be sequentially acquired images having a sufficient degree of spatial overlap yet acquired at different position and orientation. The 3D-Fourier transform $I_2(f_x, f_y, f_z)$ of $i_2(x,y,z)$ is rotated around different angle combinations $\boldsymbol{\theta} = [\theta_x, \theta_y, \theta_z]^T$ (represented by $I_2^\theta(f_x, f_y, f_z)$) and cross-correlated with the Fourier transform $I_1(f_x, f_y, f_z)$ of $i_1(x,y,z)$. As similarity metric, the normalized cross-correlation (NCC) is calculated via [38] and referred to as the rotation correlation function (RCF)

$$\text{RCF}_{\boldsymbol{\theta}_x, \boldsymbol{\theta}_y, \boldsymbol{\theta}_z} = \frac{1}{\sqrt{P_1 P_2}} \frac{|I_1(f_x, f_y, f_z) \cdot I_2^\theta(f_x, f_y, f_z)|}{I_1^\theta(f_x, f_y, f_z)} \quad (1)$$

where P_1 and P_2 are the image powers (average quadratic image intensity) of frames $i_1(x,y,z)$ and $i_2(x,y,z)$, respectively. Note that $\text{RCF}(\boldsymbol{\theta}_x, \boldsymbol{\theta}_y, \boldsymbol{\theta}_z)$ is not affected by phase differences associated with translational shifts. Any arbitrary rotation between frames is accurately estimated as

$$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_x, \hat{\boldsymbol{\theta}}_y, \hat{\boldsymbol{\theta}}_z^T = \underset{\boldsymbol{\theta}_x, \boldsymbol{\theta}_y, \boldsymbol{\theta}_z}{\text{argmax}} \text{RCF}_{\boldsymbol{\theta}_x, \boldsymbol{\theta}_y, \boldsymbol{\theta}_z} \quad (2)$$

However, (1-2) represent a computationally expensive full 3D approach, impractical for high-resolution data. Thereby, we derive herein an approximated method using maximum intensity projections (MIPs) of the 3D images as an input. According to Euler's rotation theorem, the rotation along an arbitrary direction is equivalent to three subsequent rotations along the corresponding Cartesian coordinates [39]. For small rotation angles ($|\theta_i| < 5^\circ$, $i = \{x, y, z\}$), the coordinates (x', y', z') of a point (x, y, z) after rotation can be approximately calculated using the simplified rotation matrix, i.e. [39].

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & -\theta_z & \theta_y \\ \theta_z & 1 & -\theta_x \\ -\theta_y & \theta_x & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3)$$

One may subsequently estimate how the rotation along an arbitrary axis affects the MIP views of the object along the three Cartesian axes. Without loss of generality, we first consider the top view (along the z direction) of a 3D structure. According to (3), the new coordinates (x', y') of any point in the image can be estimated from the previous (x, y, z) coordinates of this point (before rotation) via

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & -\theta_z \\ \theta_z & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \theta_y \\ -\theta_x \end{bmatrix} z \quad (4)$$

The first term in (4) represents a rotation around the z-axis. The second term may have different effects depending on the orientation of the particular structure. In the visible and near-infrared excitation spectra, endogenous OA contrast is mainly governed by hemoglobin with blood vessels appearing as the predominantly visible structures in the images. For vessels distributed normal to the z direction, the second term in (4) is constant, solely representing translation along x and y. For other vessels also expanding along z, this second term corresponds to non-rigid motion leading to shape distortions as manifested in the top views. As a result, vessels having dominant propagation directions in the x-y plane will be the main overlapping structures in the z-axis MIPs before and after rotation. Following the same reasoning, the correlation of the x- or y-axis MIPs before and after rotation will be mainly for the blood vessels expanding perpendicular to x and y, respectively. The rotation correlation function in (1) can be adapted for a faster MIP-based methodology via

$$\begin{aligned} \text{RCF}_x(\boldsymbol{\theta}_x) &= \frac{1}{\sqrt{P_1 P_2}} \frac{|I_{1, \text{MIP}, x}(f_y, f_z) \cdot I_{2, \text{MIP}, x}^\theta(f_y, f_z)|}{I_{2, \text{MIP}, x}^\theta(f_y, f_z)} \\ \text{RCF}_y(\boldsymbol{\theta}_y) &= \frac{1}{\sqrt{P_1 P_2}} \frac{|I_{1, \text{MIP}, y}(f_x, f_z) \cdot I_{2, \text{MIP}, y}^\theta(f_x, f_z)|}{I_{2, \text{MIP}, y}^\theta(f_x, f_z)} \\ \text{RCF}_z(\boldsymbol{\theta}_z) &= \frac{1}{\sqrt{P_1 P_2}} \frac{|I_{1, \text{MIP}, z}(f_x, f_y) \cdot I_{2, \text{MIP}, z}^\theta(f_x, f_y)|}{I_{2, \text{MIP}, z}^\theta(f_x, f_y)} \end{aligned} \quad (5)$$

where $I_{1, \text{MIP}, i}(f_j, f_k)$ and $I_{2, \text{MIP}, i}^\theta(f_j, f_k)$ are the respective 2D Fourier transforms of the MIPs along the i-th direction before and after rotation. Note that the correlation functions in (5) generally depend on $\boldsymbol{\theta} = [\theta_x, \theta_y, \theta_z]^T$. However, as previously alluded, the correlation is primarily associated with a rotation around the Cartesian axis corresponding to the MIP view and it is further assumed that the correlation functions are solely dependent on this angle. The rotation angles along the x-, y- and z-directions are consequently estimated via

$$\begin{aligned} \hat{\theta}_x &= \underset{\theta_x}{\text{argmax}} \{\text{RCF}_x(\theta_x)\} \\ \hat{\theta}_y &= \underset{\theta_y}{\text{argmax}} \{\text{RCF}_y(\theta_y)\} \\ \hat{\theta}_z &= \underset{\theta_z}{\text{argmax}} \{\text{RCF}_z(\theta_z)\} \end{aligned} \quad (6)$$

The two-dimensional Fourier-based estimation of the rotation angles is showcased in Fig. 1b including the image filtering steps introduced in the Image filtering section.

C. Fourier-Based Translation Estimation

After the rotation between two consecutive frames has been estimated and compensated for, the translation estimation is performed by adapting a phase correlation method [40] with frame $i_2(x,y,z)$ being merely a shifted version of frame $i_1(x,y,z)$. By further taking into account the noise $n(x,y,z)$

and interfering components $d(x,y,z)$, which represent artifacts and transient structures either entering or leaving the effective

FOV of the imaging system, the relationship between the two images is expressed as

$$i_2 = i_1 \cdot x - t_x, y - t_y, z - t_z + n(x, y, z) + d(x, y, z) \quad (7)$$

This relation is transformed into Fourier domain in order to compute the phase correlation function (PCF) defined as [40]

$$\text{PCF}_{f_x, f_y, f_z} = \frac{I_1(f_x, f_y, f_z) \cdot I_2^*(f_x, f_y, f_z)}{|I_1(f_x, f_y, f_z) \cdot I_2^*(f_x, f_y, f_z)|} \quad (8)$$

where the term $I_2^*(f_x, f_y, f_z)$ is the complex conjugate of $I_2(f_x, f_y, f_z)$. The inverse Fourier transform of the PCF yields the phase only correlation (PoC), which corresponds to a Dirac's delta function at translation $\mathbf{t}$, provided the images are not severely affected by noise or interfering transient structures. In our case, these terms reappear as disturbing components $\tilde{n}$ and $\tilde{d}$ in the PoC, i.e.

$$\text{PoC}(x, y, z) = \mathbf{F}^{-1} \text{PCF}_{f_x, f_y, f_z} = \delta(x - t_x, y - t_y, z - t_z) + \tilde{n}(x, y, z) + \tilde{d}(x, y, z) \quad (9)$$

Assuming that the interfering components are negligible in comparison to the Dirac's delta function $\delta(\mathbf{x} - \mathbf{t})$, the translation $\mathbf{t}$ can be estimated from the maximum of the PoC, i.e.

$$\hat{\mathbf{t}} = \begin{bmatrix} \hat{t}_x \\ \hat{t}_y \\ \hat{t}_z \end{bmatrix} = \underset{x, y, z}{\text{argmax}} \{ \text{PoC}(x, y, z) \} \quad (10)$$

A schematic description of the method for translation estimation is depicted in Fig. 1c including image filtering steps.

D. Image Filtering

OA images are known to be affected by noise, negative values associated to limited-view artifacts [41] as well as reflection and reverberation artifacts, which propagate into the Fourier transform of the images. In order to increase the accuracy of both RCF and PoC computation, the 3D images were thresholded so that only the brightest voxels in every image are considered, i.e. 0.1% of voxels with the highest intensity representing the most dominant image features. The noise and undesired artifacts below the threshold levels are discarded (step 2 in Figs. 1b-c). This measure has an additional benefit of eliminating distortions due to an inhomogeneous laser light distribution since the shift and rotation parameters are determined according to geometrical shapes, such as sharp edges along thresholded image features, rather than image intensities.

Since most of the spectral energy in the images is concentrated at low frequencies, only spectrally relevant components were selected for the rotation estimation. The spectral resolution in this region was furthermore enhanced by artificially increasing the number of samples in the Fourier domain via zero-padding of the MIPs (Fig. 1b, step 3 and 4) [42], [43]. The RCF generally has a concave shape with a global maximum. However, in practice it is typically affected by noise, which may shift the actual position of the maximum. To attain a more accurate and robust estimation, a least squares approximation (LSA) of the RCF was performed by means of a higher order polynomial (Fig. 1b, step 5).

We further applied a Gaussian low-pass filter to the PoC in order to mitigate noise associated with $\tilde{n}$ and $\tilde{d}$. In case the PoC does not represent a Dirac's delta function but is distributed over multiple spatially neighboring positions, the low-pass filtering concentrates the distributed energy in its spatial focus, leading to a more accurate estimate based on the filtered PoC (Fig. 1c, step 4).

E. Spatial Compounding Procedure

In total, $N-1$ rotations θ_n and translations $\mathbf{t}_n$ between subsequent frames need to be estimated in order to render a compounded image is from an OA dataset containing N sequentially acquired frames in ($n = 1 \dots N$). For this, the initial frame i_1 is assigned with coordinates $\mathbf{T}_1 = [0, 0, 0]^T$ and orientation $\mathbf{1} = [0, 0, 0]^T$. Since translation and rotation can be superimposed, all the following frames i_n ($n = 2 \dots N$) are assigned with coordinates $\mathbf{T}_n = \mathbf{G}_{i=2 \dots n} \mathbf{t}_i$ and orientation $\mathbf{n} = \mathbf{G}_{i=2 \dots n} \theta_i$. In essence, $\mathbf{T}_n$ represents the 3D trajectory of the transducer motion. Note that the absolute position and orientation of every frame depends on all previous estimates represented by the relative rotation and shift between consecutive frames. Therefore, the estimation inaccuracies and error propagation were further mitigated by an extra estimation step comparing every frame with the compounded image i_s . Once the location and orientation of frame i_n is determined, it will be added to i_s . For this, frame i_n is first rotated to its position $\mathbf{T}_n$. Here the image intensities have simply been added up, even though the redundant intensity values in overlapping frame areas can be combined in several other ways [44]–[46]. All the data / image processing and compounding routines were implemented in MATLAB (Mathworks Inc, Natick, MA, USA) running on a 3.4 GHz Intel i7 3820 CPU with 64 GB of RAM.

F. Imaging Experiments

The suggested Fourier-based motion estimation method was experimentally tested with three independent experiments. First, two phantoms were imaged in a handheld experiment in order to qualitatively validate the accuracy of the motion estimation method. The first phantom consisted of a vessel-mimicking structure with an approximate size of 30 mm $\times$ 30 mm printed with black ink on a white paper and embedded in agar (Fig. 2a). In the second phantom, light absorbing polyethylene microspheres with 100 μm diameter (Cospheric BKPMS 90-106) were randomly distributed in an agar-based substrate (Fig. 2d). The phantoms were scanned by the spherical array probe in a freehand mode by following a random trajectory. The laser was tuned to operate at 720 nm wavelength and 10 Hz pulse repetition frequency. The probe was moved slowly with inter-frame displacements not exceeding several millimeters, thus ensuring sufficient overlap between the consecutive frames.

In the second experiment, a spiral volumetric optoacoustic tomography (SVOT) scan [23] of a female athymic nude-Fox1nu mouse (Harlan Laboratories LTD, Itingen, Switzerland) was performed in order to assess accuracy of the suggested motion estimation method. In this case, the orientation and position of every frame is known in advance,

Fig. 2. Spatial image compounding results for phantom scans. (a) The printed vessel-mimicking structure. (b) Maximum intensity projections (MIP) of a single 3D image frame taken for a single position of the detection array. (c) The corresponding MIP of the compounded volume. (d) Reference image of the microsphere phantom taken by a bright-field microscope. The corresponding MIPs of the single 3D frame and compounded image are shown in (e) and (f), respectively.

which provides a gold-standard reference for validating the algorithm's performance. For this purpose, the spherical matrix transducer array was translated and rotated around the mouse using calibrated stages in steps of 1.5 mm and 3° , respectively (Fig. 3a). The scan consisted of a total of 21 elevational positions and 61 angular positions. The 1064 nm output of the pulsed laser was used for OA signal excitation. During the experiment, the mouse remained in a stationary vertical position inside a water tank heated to 34°C . All procedures involving animals were conducted in full compliance with the institutional guidelines of the Institute for Biological and Medical Imaging and with approval from the Government District of Upper Bavaria.

Lastly, imaging of an arm and palm of a healthy volunteer was performed. For this, the spherical matrix transducer array was first translated across the imaged area using a mechanical stage covering a zigzag-type sampling pattern with 2mm distance between neighboring grid points, thus acquiring a total of 21×21 volumetric image frames. A freehand scan was subsequently performed with the transducer slowly moved over the palm along an arbitrary trajectory with $< 2\text{mm}$ inter-frame displacements. Total of 200 volumetric frames were recorded. For the human experiments, the wavelength was set to 800nm and the laser was operated at pulse repetition frequency of 10 Hz. The laser fluence and average intensity at the skin surface were maintained below 15 mJ/cm^2 and 150 mW/cm^2 , respectively, which is below the safety limits for human skin laser exposure in the near-infrared spectrum [47].

The entire experiment was conducted in full compliance with the American National Standard for safe use of Lasers [47]. Although the human arm was not supposed to be moved during the experiment, small movements are barely avoidable. The associated motion effects are discussed and analyzed in the results section.

II. RESULTS

A. Handheld Phantom Scans

OA images were acquired with a spherical matrix ultrasound array schematically depicted in Fig. 1a and described in detail

in the Materials and Methods section. The relative position and orientation of consecutive 3D images following an arbitrary scanning trajectory was expressed as a superposition of a translation and a rotation. To render a compounded OA volume, the rotation among subsequent acquisitions was estimated by cross-correlating their rotated Fourier spectra (Fig. 1b) whereas the translation fits were performed by a phase correlation method (Fig. 1c), as described in detail in the Materials and Methods section.

Fig. 2 illustrates results of the spatial compounding procedure for the phantom scans. Clearly, the effective FOV in the vessel-mimicking phantom has been increased significantly by the spatial compounding procedure (Fig. 2c), further showing good agreement with the originally printed structure (Fig. 2a). Similarly, good agreement was found between the compounded OA image (Fig. 2f) and the bright-field microscopy image of the microsphere phantom (Fig. 2d). Note that, whereas the sphere location in the reconstructed OA images matches well the ground truth image, the average reconstructed size ($\sim 300 \mu\text{m}$) greatly exceeds the actual sphere diameter ($100 \mu\text{m}$). The discrepancy can be attributed to the convolution with the spatial resolution (point spread function) of the imaging system.

B. The SVOT Scan

The results of the SVOT scan experiment are shown in Fig. 3. In this case, motion of the probe is fully calibrated by the translation and rotation stages so that accurate references are available for both orientation and position of each volumetric image frame (Fig. 3a). Due to the cylindrical scanning geometry, the motion estimations were further transformed into cylindrical coordinates with r , φ and z representing translation coordinates and θ_r , θ_φ and θ_z representing rotation angles with respect to the cylindrical axes. A subsequent analysis of the spatial compounding performance attained standard deviation (STD) of the translation estimation error of 1.29 voxels, 1.10 voxels and 1.06 voxels in the r , φ and z directions, respectively (isotropic voxel size is $100\mu\text{m}$). Note that translation in the r direction implies that only the distance of the probe from the object is altered, thus the imaged structures remain almost unaltered, i.e., the contribution of the d term in (7) is insignificant. Conversely, translation in the φ - z plane results in a more significant alteration of object's illumination and thus more significant alterations to the reconstructed image. This results in a less accurate translation estimation in the φ - z plane as compared to the r axis.

We further performed a quantitative comparison between rotation estimation accuracy with the 3D approach using (2) and the simplified 2D approach using (6), by considering that the actual positions of the array are known. The STD values of the error distributions (in degrees) are provided in the table in Fig. 3, suggesting that the 3D approach attains slightly more accurate results. However, this comes at the expense of significantly higher computational complexity. The estimation of rotation angles between two frames typically takes 0.08 s with the 2D approach versus 49.31 s with the 3D approach.

Fig. 3. Illustration of the Fourier-based spatial compounding for a spiral volumetric optoacoustic tomography (SVOT) scan. (a) Schematic of the SVOT system showing scanning trajectory of the spherical matrix detection array. The inset shows a single reconstructed image frame covering an approximate volume of 1 cm^3 . (b) The known and estimated scanning positions of the center of the image frames are indicated by blue and red dots, respectively, for four different slices (z -coordinates) of the scan. The exact location of the slices is labeled with blue dashed lines in panels (f) and (g). (c) MIP image (along the central mouse axis) of the compounded volume based on the estimated array positions and orientations. (d) The corresponding MIP rendered using the known array positions and orientations. (e) Comparison of the rotation estimation accuracy in the SVOT scan for the 2D and 3D Fourier-based spatial compounding approaches. Standard deviation (STD) values in degrees are provided. (f) and (g) are 3D images of the compounded volume based on the estimated and known array positions and orientations, respectively. (h) and (i) are zoom-ins into the kidney area in (f) and (g).

The 2D approach was therefore selected for compounding the entire mouse volume. Since the probe is solely rotated along the φ -axis during the SVOT scan, the orientation estimation and correction for the r and z axes was neglected. The translation estimation was performed using the two-step process, described in the Materials and Methods section. First, the position of frame i_n was estimated based on the previous frame i_{n-1} . Final adjustment was subsequently performed on the estimated position based on the stitched frame i_s . Moreover, the search range in the PoC was limited. Fig. 3b shows the known (blue dots) and estimated (red dots) positions of the spherical array probe in the r - φ plane for four exemplary slices along the z axis (exact location of the slices is labeled in panels f and g). While the estimation has proven to be accurate on a local scale, relatively large deviations between the known and estimated positions may occur due to accumulated errors. It should be noted that the image-based estimation of the trajectory of the probe may further depend on the animal motion during the scan (e.g. due to breathing), in which case the image-based spatial compounding may in fact partially compensate for the artifacts caused by motion of the imaged object. Nevertheless, a good agreement exists between the SVOT image attained using the proposed Fourier-based spatial compounding procedure (Figs. 3d and 3g) and the corresponding (reference) image obtained by adding the frames using known locations and orientations of the array for all the scanning positions (Figs. 3c and 3f). Yet, some blurring effects and loss of contrast can be recognized when comparing fine vascular structures in the images obtained using the estimated positions and orientations of the array (Fig. 3i) with the “ground truth” images rendered using known array positions (Fig. 3h). A direct comparison of the rendered images is provided in supplementary movie.¹ Quantitatively, the 3D rotation

estimation method performs slightly better than the 2D estimation method, as summarized in the table shown in Fig 3e.

C. Zigzag Scan of a Human Arm

Fig. 4 illustrates results of the spatial compounding procedure for the zigzag scan. The top and lateral MIPs of the compounded volume are displayed in Fig. 4a as well as the projected estimated 3D trajectory, indicated by the white dashed line. The effective FOV has been increased significantly by the spatial compounding procedure from $15 \text{ mm} \times 15 \text{ mm} \times 15 \text{ mm}$ (single volumetric frame) to $500 \text{ mm} \times 500 \text{ mm} \times 8 \text{ mm}$ (compound volumetric frame). The inclination of the image and the projected trajectory in the x - y plane indicates that the orientation of the spherical matrix array does not exactly match the orientation of the translation path. Under this scenario, volumetric image compounding solely based on the known positions of the translation stage would result in severe misalignment artifacts. On the other hand, the Fourier-based compounding procedure is not affected by an insufficiently calibrated orientation, rendering accurately stitched volumes. Note that the spherical matrix array was moved along straight lines by the translation stage, yet some of the trajectories in Fig. 4a are bent despite the fact that the compound volume does not exhibit any signs of misalignment. This further suggests that the marginal motions of the human arm during the acquisition have been corrected implicitly by the spatial compounding procedure. Hence, the single axis trajectories $x(n)$, $y(n)$ and $z(n)$ in Fig. 4b can be interpreted as motion superposition of both the spherical matrix array and the human arm, further emphasizing effectiveness of the developed spatial compounding procedure. The time lapse flow of the spatial compounding process for the freehand scan is further illustrated in supplementary movie² accompanying the manuscript.

Fig. 4. Volumetric optoacoustic angiography of a human arm performed in a zigzag-scan pattern. (a) Top and lateral maximum intensity projections of the compounded volume. Depth axis is color-coded. The white dashed line represents estimated the trajectory of the spherical array transducer during the scan. (b) Estimated single axis trajectories $x(n)$, $y(n)$, and $z(n)$ of the transducer during the zigzag scan. The frame index n represents time via $t = nT$.

D. Arbitrary (Freehand) Motion Scan

Results from the freehand scan of a human palm are shown in Fig. 5. This experiment mimics a clinical imaging scenario where the transducer follows an arbitrary trajectory having all the six degrees of freedom. Fig. 5a shows MIPs of the compounded volume with the projected estimated 3D trajectory indicated by a dashed white line. In this example the effective FOV is increased from $12 \text{ mm} \times 12 \text{ mm} \times 12 \text{ mm}$ (single volumetric frame) to approximately $50 \text{ mm} \times 70 \text{ mm} \times 15 \text{ mm}$ (compound volume). The depth range was changed to cover the detected shift in the vertical (z) direction, which was approximately 3 mm. Fig. 5b shows the estimated transducer orientations for every frame in the three directions. The orientation remains predominantly constant with respect to the x and y axes whereas the z -axis orientation varies between -45° and $+45^\circ$. The translation parameters are further estimated in Fig. 5c. The transducer was primarily moved along the x and y directions with the z -coordinate remaining nearly constant. This is expected considering the relatively flat skin surface in the imaged area. The compound volume appears to accurately represent an actual vascular network with no signs of misalignment perceived in the individual vessels. Note that, according to the estimated probe trajectory, some tissue areas were revisited during the scan.

Fig. 5. Freehand optoacoustic human angiography of a human palm along an arbitrary trajectory. (a) Top and lateral maximum intensity projections of the compounded volume. Depth axis is color-coded. The white dashed line represents the projected 3D trajectory of the spherical array transducer operated in a freehand mode. (b) The estimated rotation angles during the scan (left) and the corresponding estimated transducer positions (right).

Yet, those areas appear seamlessly stitched despite the fact that they were compounded using non-consecutive frames. The time lapse flow of the spatial compounding process for the freehand scan is further illustrated in supplementary movie³ accompanying the manuscript.

III. DISCUSSION

A universal methodology was developed for spatial compounding of volumetric optoacoustic data acquired using either calibrated scanning systems or freehand-mode scans with up to six rotational and translational degrees of freedom. This was accomplished by a purely image-based Fourier domain motion estimation method without using additional hardware for tracking the position and orientation of the detection array, which may turn impractical especially in the case of high-resolution freehand 3D scans. Even though Fourier-based approaches have previously been reported in the context of registration of cross-sectional (2D) image data [48], [49], the combination of both rotation and translation estimation for volumetric image sequences is attempted for the first time. Moreover, our approach outperformed feature-based techniques [50], which may often fail to provide sufficient reliability and accuracy due to an insufficient amount of feature matches in the registration step. Other registration

methods, e.g. based on mutual information or sum of squared differences [51], may yet be considered.

The freehand real-time 3D imaging capacity greatly facilitates clinical utility of the OA imaging technology, with applications currently explored in many areas of clinical diagnostics of skin malignancies, breast tumors, vascular abnormalities, and inflammatory diseases [17], [29], [52], to name a few major examples. The fast imaging performance often comes at the cost of limited field-of-view, which can effectively be compensated by the developed trackerless spatial compounding algorithm. The method's performance was validated by controlled experiments with phantoms, for which a reference image was available, and further supported by *in vivo* data acquired from mice and a human arm, where the positions of the transducer array were known exactly and served as a reference. In all cases the estimation accuracy remained below the diffraction-limited spatial resolution of the imaging system. Freehand scans of two phantoms and large-scale vascular trees in a human arm using arbitrary (unknown) trajectories have similarly rendered compounded image volumes having no visible signs of misalignments or discontinuities.

The algorithm performed optimally for inter-frame displacements in the order of 0.5-2 mm. Considering the 12x12 mm² effective field of view of the spherical array probe, this corresponds to >85% of voxels overlapping between two consecutive frames, thus ensuring they mainly contain similar structures. Accurate translation estimation may yet fail for larger displacements when significantly different images are rendered for consecutive positions or the images lack distinctive features. The maximum scanning speed then depends on the pulse repetition frequency of the laser. The latter is often kept in the 10-20 Hz range in order to conform to the safety limits pertaining average laser intensity on the human skin [47]. A scanning speed below ~2 cm/s would then be sufficient to guarantee the required overlap between consecutive images. Note that here we employed a simple superposition approach by simply adding up the image intensities values of the adjacent frames. In this case, non-uniform spatial sampling density due to e.g. varying motion velocity of the probe may lead to uneven signal intensities in the compounded images. Other compounding methods can alternatively be applied to facilitate accurate image formation, e.g. by weighting each voxel in the final image by the total number of scan frames effectively covering it. However, this would involve a much more significant modeling and computational effort to take into account the actual sensitivity field of the matrix array transducer, wavelength- and tissue-dependent volumetric light distribution in tissues etc. Another important consideration is the effectively covered imaging depth. In this regard, optical wavelengths experiencing weaker attenuation in tissues (e.g. 1064 nm) can be employed to expand the covered depth range, which may in turn facilitate image co-registration as more structures are visible in the images. However, it is additionally important to keep the depth range of interest within the field of view of the transducer during the entire freehand scan.

In the SVOT scan, the probe orientation between subsequent frames varied in the $\pm 3^\circ$ range with respect to all the three axes. In theory, the rotation estimation may remain accurate

for arbitrarily large rotation angles, provided that the consecutive frames show sufficient overlap. In realistic optoacoustic imaging scenarios, however, an abrupt change in transducer orientation is accompanied by a very significant alteration of the effectively imaged field-of-view, resulting in lack of sufficient overlap between subsequent frames. These effects can be mitigated by using laser pulse repetition frequencies in the order of tens of Hz while ensuring that the probe is moved slowly and steadily during the freehand scan, avoiding abrupt leaps or changes in orientation. The experiments were executed using a relatively slow velocity not exceeding several centimeters per second. As a rule of thumb, the velocity scales linearly with the laser pulse repetition frequency and the effectively covered FOV. The accuracy of the motion estimation can be increased when employing higher resolution optoacoustic imaging systems. Recently, real-time 3D OA imaging with spatial resolution in the 35 μ m range has been demonstrated using spherical matrix array transducer with 25MHz usable detection bandwidth [24]. In this case, the effective FOV is scaled down accordingly [24], thus reducing the permissible displacements between consecutive frames.

The developed methodology may enable accurate estimation of other types of motion not necessarily linked to freehand scanning. For instance, motion artefacts are often generated in a sequence of images due to breathing or heartbeat [9], [23]. In general, motion does not lead to blurring in single optoacoustic image frames since optoacoustic excitation is performed with very short (nanosecond duration) laser pulses while the responses are collected simultaneously by all the array elements. Nevertheless, many types of movements such as arterial pulsation is often accompanied by structural tissue deformations on a small spatial scale that cannot accurately be accounted for by assuming rigid motion. The latter approximation may still be valid in some cases where motion affects a region much larger than a single OA image volume where deformations remain minimal on a local scale (e.g. during breathing). In such cases, correcting for motion may help to better identify temporal changes in the signals, such as those associated with contrast agent perfusion [9] or physiological activity [18]. Even if no motion correction is performed, detection and rejection of the frames affected by motion may facilitate more efficient signal averaging, enhancing the spatial resolution and contrast-to-noise ratio of the images [16]. Motion correction is also essential for processing of MSOT data where displacements between the images acquired at different wavelengths may hinder accurate identification (unmixing) of spectrally-distinctive absorbers [53]–[55]. Despite the fast wavelength tunability available with state-of-the-art pulsed lasers [4], even slight (sub-resolution) motion between images taken at different wavelengths may generate significant spectral unmixing artifacts in freehand scans [54]. The OA images are also often afflicted with artefacts related to acoustic scattering or reflections [56]. While those were negligible in the present experiments solely involving soft tissues, areas including bones, lungs or other strongly mismatched regions may need to be cropped before image registration.

In conclusion, we have developed a novel Fourier-based framework for spatial compounding of time-lapse optoacoustic data acquired using large-area volumetric scans. The method allows rapidly combining sequential volumetric frames into large area scans without using additional tracking hardware. The new approach is universally applicable for compounding volumetric data acquired with calibrated scanning systems but also in a freehand mode with up to six rotational and translational degrees of freedom. Robust performance was demonstrated for whole-body mouse imaging with spiral volumetric optoacoustic tomography as well as for freehand visualization of large-scale vascular networks in humans using 3D imaging probes. It is anticipated that the newly introduced capability for angiographic observations at multiple spatial and temporal scales will greatly facilitate the use of optoacoustic imaging technology in pre-clinical research and clinical diagnostics. The volumetric compounding technique can equally benefit other biomedical imaging modalities looking at large-scale vascular network data, such as scanning fluorescence microscopy, optical coherence tomography or ultrasonography, thus optimizing their trade-offs between fast imaging performance and field-of-view.

REFERENCES

- [1] J. Xia, J. Yao, and L. V. Wang, "Photoacoustic tomography: Principles and advances," *Prog. Electromagn. Res.*, vol. 147, pp. 1–22, May 2014.
- [2] S. Zackrisson, S. M. W. Y. van de Ven, and S. S. Gambhir, "Light in and sound out: Emerging translational strategies for photoacoustic imaging," *Cancer Res.*, vol. 74, no. 4, pp. 979–1004, 2014.
- [3] V. Ntziachristos, "Going deeper than microscopy: The optical imaging frontier in biology," *Nature Methods*, vol. 7, pp. 603–614, Jul. 2010.
- [4] X. L. Deán-Ben and D. Razansky, "Adding fifth dimension to optoacoustic imaging: Volumetric time-resolved spectrally enriched tomography," *Light: Sci. Appl.*, vol. 3, p. e137, Jan. 2014.
- [5] N. Beziere *et al.*, "Dynamic imaging of PEGylated indocyanine green (ICG) liposomes within the tumor microenvironment using multispectral optoacoustic tomography (MSOT)," *Biomaterials*, vol. 37, pp. 415–424, Jan. 2015.
- [6] L. R. McNally *et al.*, "Current and emerging clinical applications of multispectral optoacoustic tomography (MSOT) in oncology," *Clin. Cancer Res.*, vol. 22, no. 14, pp. 3432–3439, Jul. 2016.
- [7] S. V. Ovsepian, I. Olefir, G. Westmeyer, D. Razansky, and V. Ntziachristos, "Pushing the boundaries of neuroimaging with optoacoustics," *Neuron*, vol. 96, no. 5, pp. 966–988, Dec. 2017.
- [8] A. B. E. Attia *et al.*, "Noninvasive real-time characterization of non-melanoma skin cancers with handheld optoacoustic probes," *Photoacoustics*, vol. 7, pp. 20–26, Sep. 2017.
- [9] 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.*, vol. 5, no. 1, 2015, Art. no. 010133.
- [10] A. Taruttis, G. M. Van Dam, and V. Ntziachristos, "Mesoscopic and macroscopic optoacoustic imaging of cancer," *Cancer Res.*, vol. 75, no. 8, pp. 1548–1559, 2015.
- [11] A. Taruttis and V. Ntziachristos, "Advances in real-time multispectral optoacoustic imaging and its applications," *Nature Photon.*, vol. 9, pp. 219–227, Apr. 2015.
- [12] 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.*, vol. 18, no. 7, 2013, Art. no. 076014.
- [13] L. Li *et al.*, "Single-impulse panoramic photoacoustic computed tomography of small-animal whole-body dynamics at high spatiotemporal resolution," *Nature Biomed. Eng.*, vol. 1, no. 5, p. 0071, May 2017.
- [14] J. Jose *et al.*, "Passive element enriched photoacoustic computed tomography (PER PACT) for simultaneous imaging of acoustic propagation properties and light absorption," *Opt. Express*, vol. 19, no. 3, pp. 2093–2104, 2011.
- [15] X. L. Deán-Ben and D. Razansky, "Portable spherical array probe for volumetric real-time optoacoustic imaging at centimeter-scale depths," *Opt. Express*, vol. 21, no. 23, pp. 28062–28071, 2013.
- [16] 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*, vol. 3, no. 11, pp. 1153–1159, 2016.
- [17] V. Neuschmelting *et al.*, "Performance of a multispectral optoacoustic tomography (MSOT) system equipped with 2D vs. 3D handheld probes for potential clinical translation," *Photoacoustics*, vol. 4, pp. 1–10, Mar. 2016.
- [18] X. L. Deán-Ben, S. Gottschalk, B. Mc Lamey, S. Shoham, and D. Razansky, "Advanced optoacoustic methods for multiscale imaging of in vivo dynamics," *Chem. Soc. Rev.*, vol. 46, no. 8, pp. 2158–2198, 2017.
- [19] R. A. Kruger, R. B. Lam, D. R. Reinecke, S. P. D. Rio, and R. P. Doyle, "Photoacoustic angiography of the breast," *Med. Phys.*, vol. 37, no. 11, pp. 6096–6100, 2010.
- [20] A. Dima, N. C. Burton, and V. Ntziachristos, "Multispectral optoacoustic tomography at 64, 128, and 256 channels," *J. Biomed. Opt.*, vol. 19, no. 3, 2014, Art. no. 036021.
- [21] J. Xia *et al.*, "Whole-body ring-shaped confocal photoacoustic computed tomography of small animals in vivo," *J. Biomed. Opt.*, vol. 17, no. 5, 2012, Art. no. 050506.
- [22] D. Razansky, A. Buehler, and V. Ntziachristos, "Volumetric real-time multispectral optoacoustic tomography of biomarkers," *Nature Protocols*, vol. 6, no. 8, pp. 1121–1129, 2011.
- [23] 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.*, vol. 6, no. 4, 2017, Art. no. e16247.
- [24] X. L. Deán-Ben, H. López-Schier, and D. Razansky, "Optoacoustic micro-tomography at 100 volumes per second," *Sci. Rep.*, vol. 7, no. 1, p. 6850, 2017.
- [25] M. Omar, D. Soliman, J. Gateau, and V. Ntziachristos, "High-resolution epi-illumination raster-scan optoacoustic mesoscopy for imaging of model organisms and microvessels," *Proc. SPIE*, vol. 9539, Jul. 2015, Art. no. 95390U.
- [26] J. Aguirre *et al.*, "Precision assessment of label-free psoriasis biomarkers with ultra-broadband optoacoustic mesoscopy," *Nature Biomed. Eng.*, vol. 1, no. 5, p. 0068, 2017.
- [27] I. Stoffels *et al.*, "Metastatic status of sentinel lymph nodes in melanoma determined noninvasively with multispectral optoacoustic imaging," *Sci. Transl. Med.*, vol. 7, p. 317ra199, Dec. 2015.
- [28] F. Knieling *et al.*, "Multispectral optoacoustic tomography for assessment of Crohn's disease activity," *New England J. Med.*, vol. 376, no. 13, pp. 1292–1294, Mar. 2017.
- [29] G. Diot *et al.*, "Multispectral optoacoustic tomography (MSOT) of human breast cancer," *Clin. Cancer Res.*, vol. 23, no. 22, pp. 6912–6922, 2017.
- [30] I. Ivankovic, E. Merčep, C.-G. Schmedt, X. L. Deán-Ben, and D. Razansky, "Real-time volumetric assessment of the human carotid artery: Handheld multispectral optoacoustic tomography," *Radiology*, vol. 291, no. 1, pp. 45–50, Apr. 2019.
- [31] A. Dima and V. Ntziachristos, "Non-invasive carotid imaging using optoacoustic tomography," *Opt. Express*, vol. 20, no. 22, pp. 25044–25057, Oct. 2012.
- [32] M. A. Yaseen *et al.*, "Optoacoustic imaging of the prostate: Development toward image-guided biopsy," *J. Biomed. Opt.*, vol. 15, no. 2, Mar. 2010, Art. no. 021310.
- [33] E. V. Petrova, H. P. Brecht, M. Motamedi, A. A. Oraevsky, and S. A. Ermilov, "In vivo optoacoustic temperature imaging for image-guided cryotherapy of prostate cancer," *Phys. Med. Biol.*, vol. 63, no. 6, Mar. 2018, Art. no. 064002.
- [34] C. Salchow-Hömmen, L. Callies, D. Laidig, M. Valtin, T. Schauer, and T. Seel, "A tangible solution for hand motion tracking in clinical applications," *Sensors*, vol. 19, no. 1, p. 208, 2019.
- [35] F. Godenschweiger *et al.*, "Motion correction in MRI of the brain," *Phys. Med. Biol.*, vol. 61, no. 5, p. R32, Mar. 2016.
- [36] A. A. A. Rahni and I. Yahya, "Obtaining translation from a 6-DOF MEMS IMU—An overview," in *Proc. Asia-Pacific Conf. Appl. Electromagn.*, 2007, pp. 1–5.
- [37] 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. Imag.*, vol. 32, no. 11, pp. 2050–2055, Nov. 2013.

- [38] A. A. Tamhane and K. Arfanakis, "Motion correction in PROPELLER and turbo-prop-MRI," *Magn. Reson. Med.*, vol. 62, no. 1, pp. 174–182, 2009.
- [39] J. Diebel, "Representing attitude: Euler angles, unit quaternions, and rotation vectors," *Matrix*, vol. 58, nos. 15–16, pp. 1–35, 2006.
- [40] R. Szeliski, "Image alignment and stitching: A tutorial," *Found. Trends Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–4, 2006.
- [41] X. L. Deán-Ben and D. Razansky, "On the link between the speckle free nature of optoacoustics and visibility of structures in limited-view tomography," *Photoacoustics*, vol. 4, no. 4, pp. 133–140, Dec. 2016.
- [42] K. M. Aamir, M. A. Maud, and A. Loan, "On Cooley-Tukey FFT method for zero padded signals," in *Proc. IEEE Symp. Emerg. Technol.*, Sep. 2005, pp. 41–45.
- [43] M. Wojtkowski, V. J. Srinivasan, T. H. Ko, J. G. Fujimoto, A. Kowalczyk, and J. S. Duker, "Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation," *Opt. Express*, vol. 12, no. 11, pp. 2404–2422, May 2004.
- [44] S. Nitkunanantharajah, C. Hennesperger, X. L. Dean-Ben, D. Razansky, and N. Navab, "Trackerless panoramic optoacoustic imaging: A first feasibility evaluation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 5, pp. 703–711, May 2018.
- [45] V. Rankov, R. J. Locke, R. J. Edens, P. R. Barber, and B. Vojnovic, "An algorithm for image stitching and blending," *Proc. SPIE*, vol. 5701, Mar. 2005, Art. no. 590536.
- [46] S. Samsudin, S. Adwan, H. Arof, N. Mokhtar, and F. Ibrahim, "Development of automated image stitching system for radiographic images," *J. Digit. Imag.*, vol. 26, no. 2, pp. 361–370, Apr. 2013.
- [47] *American national standard for safe use of lasers*, Amer. Nat. Standards Inst., ANSI, Orlando, FL, USA, Mar. 2007.
- [48] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Process.*, vol. 5, no. 8, pp. 1266–1271, Aug. 1996.
- [49] R. C. Hardie, K. J. Barnard, J. G. Bogner, E. E. Armstrong, and E. A. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Opt. Eng.*, vol. 37, no. 1, Jan. 1998, Art. no. 601623.
- [50] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
- [51] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: A review," *Comput. Methods Biomech. Biomed. Eng.*, vol. 17, no. 2, pp. 73–93, 2014.
- [52] A. Taruttis *et al.*, "Multispectral optoacoustic tomography of myocardial infarction," *Photoacoustics*, vol. 1, no. 1, pp. 3–8, 2013.
- [53] A. Taruttis, J. Claussen, D. Razansky, and V. Ntziachristos, "Motion clustering for deblurring multispectral optoacoustic tomography images of the mouse heart," *J. Biomed. Opt.*, vol. 17, no. 1, 2012, Art. no. 016009.
- [54] X. L. Deán-Ben, E. Bay, and D. Razansky, "Functional optoacoustic imaging of moving objects using microsecond-delay acquisition of multispectral three-dimensional tomographic data," *Sci. Rep.*, vol. 4, p. 5878, Jul. 2014.
- [55] B. T. Cox, J. G. Laufer, P. C. Beard, and S. R. Arridge, "Quantitative spectroscopic photoacoustic imaging: A review," *J. Biomed. Opt.*, vol. 17, no. 6, 2012, Art. no. 061202.
- [56] X. L. Deán-Ben, V. Ntziachristos, and D. Razansky, "Artefact reduction in optoacoustic tomographic imaging by estimating the distribution of acoustic scatterers," *J. Biomed. Opt.*, vol. 17, no. 11, Oct. 2012, Art. no. 110504.