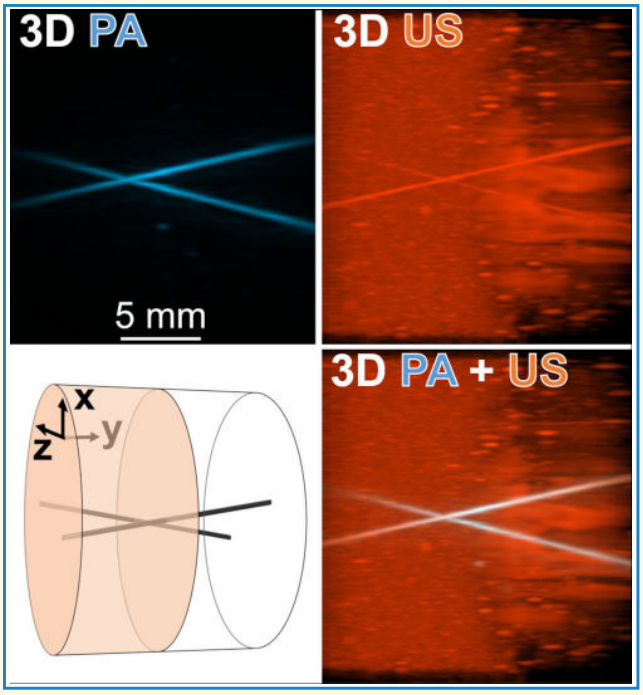
Volumetric and Simultaneous Photoacoustic and Ultrasound Imaging With a Conventional Linear Array in a Multiview Scanning Scheme

Clément Linger, *Member, IEEE*, Yoann Atlas, Rémy Winter, Marine Vandebrouck, Maxime Faure, Théotim Lucas, S. Lori Bridal, *Senior Member, IEEE*, and Jérôme Gateau

***Abstract*—Volumetric, multimodal imaging with precise spatial and temporal coregistration can provide valuable and complementary information for diagnosis and monitoring. Considerable research has sought to combine 3-D photoacoustic (PA) and ultrasound (US) imaging in clinically translatable configurations; however, technical compromises currently result in poor image quality either for PA or ultrasonic modes. This work aims to provide translatable, high-quality, simultaneously coregistered dual-mode PA/US 3-D tomography. Volumetric imaging based on a synthetic aperture approach was implemented by interlacing PA and US acquisitions during a rotate-translate scan with a 5-MHz linear array (12 angles and 30-mm translation to image a 21-mm diameter, 19 mm long cylindrical volume within 21 s). For coregistration, an original calibration method using a specifically designed thread phantom was developed to estimate six geometrical parameters and one temporal offset through global optimization of the reconstructed sharpness and superposition of calibration phantom structures. Phantom design and cost function metrics were selected based on analysis of a numerical phantom and resulted in a high estimation accuracy for the seven parameters. Experimental estimations validated the calibration repeatability. The estimated parameters were used for the bimodal reconstruction of additional phantoms with either identical or distinct spatial distributions of US and PA contrasts. The superposition distance of the two modes was within <10% of the acoustic wavelength, and a wavelength-order uniform spatial resolution was obtained. This dual-mode PA/US tomography should contribute to more sensitive and robust detection and follow-up of biological changes or the monitoring of slower-kinetic phenomena in living systems such as the accumulation of nanoagents.**



***Index Terms*—Photoacoustic (PA), rotate-translate scan, simultaneous dual imaging, tomography, ultrafast ultrasound (US) imaging, volumetric imaging.**

I. Introduction

VOLUMETRIC and simultaneously coregistered multimodal imaging is increasingly developing in biomedical imaging to more fully exploit the growing range of rich, multiplexed, and complementary anatomical–functional information that can be precisely correlated in time and with respect to spatial position [1]. One of the first such bimodal imaging systems arrived in the clinical armamentarium in the 1990s. This pioneering multimodal approach, positron emission tomography (PET) coupled with computed tomography (CT), showed the interest of simultaneous coregistered tomography providing the anatomical context of CT to better interpret metabolic information offered by PET [2]. Many other original volumetric multimodal tomography combinations have followed. For example, PET has been integrated with magnetic resonance imaging (MRI) [1], [2]. Photoacoustic imaging (PAI) and optical coherence tomography (OCT) have

Manuscript received 14 February 2023; accepted 13 April 2023. This work was supported in part by Sorbonne University under the Program Emergence Sorbonne Université 2019–2020; in part by the Centre national de la recherche scientifique (CNRS) through the Mission pour les Initiatives Transverses et Interdisciplinaires (MITI) Programs (Defi Imag'IN, 80 Prime); in part by the Gefluc Paris, Ile de France; and in part by the France Life Imaging under Grant ANR-11-INBS-0006. *(Corresponding author: Jérôme Gateau.)*

Please see the Acknowledgment section of this article for the author affiliations.

This article has supplementary downloadable material available at https://doi.org/10.1109/TUFFC.2023.3268820, provided by the authors.

Digital Object Identifier 10.1109/TUFFC.2023.3268820

Highlights

- **High quality, volumetric and simultaneous photoacoustic (PA) and ultrasound (US) imaging was obtained by rotate-translate scanning of a conventional linear ultrasound array.**
- **A calibration method enabled to estimate 6 geometrical parameters and 1 temporal off-set to obtain co-registered PA/US images with a superposition distance within $<$10% of the acoustic wavelength.**
- **This dual-mode PA/US tomography should contribute to more sensitive and robust monitoring of slower-kinetic molecular or functional phenomena in living systems such as the accumulation of nanoagents.**

been superimposed [3] to enable novel extraction of tissue characteristics like chromophore concentration. PET has been combined with Doppler ultrasound (US) imaging [4] to relate metabolism and blood flow.

Multimodal tomography that provides a simultaneous and coregistered view of different aspects within the body can provide manifold advantages. By obtaining more information within a single imaging session, patients, researchers, and clinical-management teams benefit due to both reduced examination times and more comprehensive characterization of the examined region. Precise coregistration in space and time is essential for a characterization that makes the most of each modality. In addition, volumetric imaging provides a detailed view of regions under examination from various orientations for improved diagnosis and facilitates comparisons during longitudinal studies or monitoring of therapeutic response.

US and PA imaging modalities display complementary information such that when combined, US imaging delineates organs and lesions to provide the anatomical reference frame for the PA-based information, which can range from the characterization of hemoglobin oxygenation to the detection of molecular and nanoparticular contrast agents [5], [6]. Although the recorded ultrasonic signals are generated in situ by the optical absorption of a laser excitation for PAI, while, for US, they are created by transmitting an ultrasonic pulse that is then backscattered from structures with different acoustic impedances, both modalities rely on the detection of ultrasonic signals. This is an advantage for bimodal PA/US imaging because signals can be received using the same US detector for simultaneous coregistration. Such simultaneous coregistration has previously been demonstrated in 2-D using a single US detector array [7], [8] and using a dual array configuration [9].

Bimodal PA/US in 2-D, however, is hampered due to several limitations. First, out-of-plane artifacts can be stronger in 2-D PAI than in 2-D US imaging. The reason is twofold. Because light is strongly scattered within biological tissue (photon diffusion regime [10]), the optical excitation used in PAI is intrinsically 3-D in nature. Furthermore, the US focusing in the elevation direction is weaker for PAI because it only occurs in reception mode. Second, elongated structures like blood vessels have a strong directionality in PAI, so they may not be visible in the limited-view configuration offered by most 2-D PAI systems if the vessels emit outside of the limited angular aperture of the detector.

Spherical US detector matrices (2-D arrays) provide a large angular aperture and have been developed specifically for 3-D PAI using transducers that can also be used in pulse-echo mode to create US images [11], [12]. Given the limited number of elements used to cover the spherical surface, spherical US matrices can be considered sparse for ultrasonic imaging. Pulse-echo US images of biological tissue are, however, densely filled with signals from echogenic structures such that high-quality, 3-D US imaging requires detector matrices with a higher spatial sampling than the spherical matrices designed for PAI. Thus, although a large, sparse angular aperture can provide high-quality photoacoustic (PA) images, especially for angiographic applications [11], [13], [14] and, although such spherical arrays have been shown to perform well when used to extract US Doppler signals [15], [16] or the signals from sparsely distributed US contrast agents [15], their use for biological tissue anatomical imaging can result in strong side lobes and grating lobes artifacts and poor contrast.

Planar US transducer matrices can offer a less sparse transducer distribution. They have been developed for 3-D US imaging and have been tested for 3-D PAI [17], [18]. While US transducer matrices enable high-frame rate 3-D US imaging and associated imaging modes (shear wave elastography, Doppler, localization microscopy, etc.) [19], the limited angular aperture of such matrices and the poor sensitivity of the relatively small elements lead to limited-view artifacts, limited spatial resolution and poor sensitivity for 3-D PAI.

Thus, at the present time, simultaneous coregistration of volumetric PA and US imaging with either spherical or planar US matrices results in degraded image quality for one of the two modalities. Specific mechanical scanning patterns with US linear arrays that have been designed for 2-D imaging may provide an alternative solution toward high-quality imaging with both PA and US techniques. Several systems based on translating a linear array transducer along the elevational direction (perpendicularly to the imaged plane) have been developed recently, and their ability to provide simultaneously coregistered 3-D US and PA imaging has been demonstrated [20], [21], [22], [23]. The 2-D images are then stacked to obtain a volume. This approach is relatively easy to implement and transfer to a clinical environment. The angular aperture in the elevational direction of a linear US array is, however, very limited because of the weak elevational focus implemented to improve the sensitivity and limit the elevational thickness of the imaged slice along a large range of depths. In PAI, directional structures could emit outside of this limited angular aperture which can cause strong limited-view artifacts. This issue is not addressed with the translational scan,

even when multiple scan positions are used to reconstruct each slice (synthetic aperture focusing) [24], [25], because the orientation of the angular aperture remains constant for all the scan positions. Moreover, the spatial resolution in the translational direction is strongly degraded compared to the in-plane resolution both for PA and US imaging.

Adding a rotational motion and implementing a rotate-translate synthetic aperture scanning of a linear US array has been shown to effectively increase the angular aperture and to highly improve the volumetric image quality and elevational resolution compared to a purely translational scan [26], [27]. Synthetic aperture scanning refers here to the fact that signals acquired at several scan positions are coherently combined during the reconstruction process to synthetize a larger angular aperture. The rotate-translate approach has previously been demonstrated independently on two different imaging systems for 3-D PAI [26] and for 3-D US imaging [27], but the data acquired with these systems has not yet been coregistered to provide dual-mode imaging.

In this article, we propose and demonstrate a dual-mode PA and US imaging approach with a linear US array that simultaneously provides high-quality coregistered volumetric data for both modes throughout a volume in the order of 7 cm^3 and within 21 s. A thread phantom calibration identifies the seven required reconstruction parameters based on cost-function global optimization. We present here the first dual PA-US 3-D calibration method. Calibration is conceived to be durable as long as the transducer is not repositioned in its support, even if the optical source is realigned. Since this dual-mode PA and US imaging approach is based on a single rotate-translate scan of a linear US array and does not call upon fiducial points for coregistration, this technique could be readily adapted to in vivo imaging constraints for a variety of applications.

II. Materials and Methods

A. Experimental Setup

The experimental setup presented in Fig. 1(a) consists of: 1) the optical excitation system comprised of a nanosecond laser and an optical fiber bundle; 2) the US excitation and acquisition system composed of a US linear array connected to a programmable US platform; 3) the scanning system made up of two motorized stages operated by a motion controller; and 4) the synchronization based on a programmable trigger generator. The entire acquisition process was automated.

An optical excitation at 700 nm was generated using an optical parametric oscillator laser (SpitLight 600 OPO, Innolas Laser GmbH, Krailling, Germany) delivering <8 ns pulses with a pulse repetition frequency (PRF) of 20 Hz. A bifurcated fiber bundle (CeramOptec GmbH, Bonn, Germany) guided the light toward the imaged volume to provide bilateral illumination. The mean laser energy at each fiber output was estimated to be around 6 mJ. The fixed PRF of the laser sets the time base for the acquisition sequence. The pulse energy was recorded using a pyrometer incorporated into the laser.

A 128-element linear US transducer array (L7–4, ATL, Seattle, WA, USA) was driven by a programmable, 64-channel

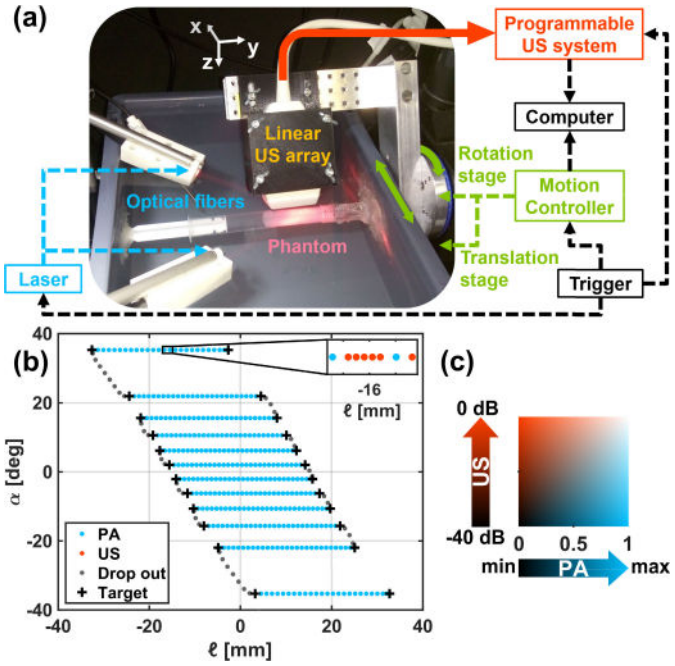


Fig. 1. (a) Annotated picture of the experimental setup. (b) Motor positions for the PA acquisitions over one scan. For better readability, positions corresponding to the US and PA acquisitions are indicated only for the inset in the upper right corner of the graph. Positions shown in gray (dropout) were not used for the reconstruction but were acquired due to the continuous motion of the motors. (c) Colormap used for all images: PA images are represented in shades of blue and US images in shades of orange, the sum of the two leading to white. PA signals are presented on a linear scale, while US signals are presented within the range of −40 to 0 dB.

US machine (Vantage, Verasonics, WA, USA). For all the transmit events and all the receive events, only the 64 central elements of the array were used. Each laser pulse triggered a receive-only event to record the PA data. The 50-ms interval between two laser pulses was divided into equal 6.25-ms parts for eight events: the initial PA acquisition, then a 6.25-ms pause followed by five US "plane waves" (a beam that is unfocused in the lateral direction of the array) transmit-receive events at a PRF of 160 Hz [28] and at steered angles of −4°, −2°, 0° (the 64 elements in the transducer array were fired at the same time), 2° and 4°, respectively, a second 6.25 ms pause and finally a new PA acquisition. The pauses just before and after laser pulses prevented possible interference between the US fields generated by the laser pulse (PA) and by the plane wave emissions (US). Since the array is continuously scanned, the low PRF for the US acquisitions enables further diversification of the spatial positions [27]. Transmitted US pulses were one cycle long with a center frequency of 5.2 MHz. Pulse-echo US signals and PA signals were recorded at a sampling frequency of 20 and 62.5 MS/s, respectively. The gain was adjusted to a constant value so that all received PA and US signals had a sufficient amplitude for good digitalization without risk of saturation. No time gain compensation was applied because of the weak US attenuation of the imaged samples.

The rotate-translate scanning system described previously in [27] consisted of a US array fixed to a rotation stage connected to a translation stage (Physik instrumente, Karlsruhe,

Germany) so that the rotation axis was perpendicular to the translation axis. The array axis (axis along the row of its elements) was aligned with the rotation axis. The rotation stage was moved to 12 different angles with a nominal angular sampling step of $\Delta\alpha = 4°$. At each angular position, the imaged volume was scanned with a translation range $L = 30$ mm and a linear sampling step $\Delta\ell = 1$ mm. Relative to parameters used in [27], the scan parameters, in particular the translation velocity, had to be adjusted to adapt to the PRF of the laser. The motion was continuous to provide more time-efficient scanning compared to stepped motion [29]. For each event producing US signals (laser emission and series of five US plane wave emissions), the motor positions were recorded and stored in the motor controller to be used for the image reconstruction. The scan was automated and initiated with an external trigger sent simultaneously by a generator (BNC Model 577, Berkeley Nucleonics, San Rafael, CA, USA) to the laser, the programmable US system, and the motion controller to synchronize all devices and the emission-acquisition sequence. The motor positions are presented graphically in Fig. 1(b).

The array elements were acoustically coupled to the imaged sample in a tank filled with tap water. The tips of the fiber bundles were also underwater to provide close illumination of the sample. The water temperature was monitored based on readings from a thermometer (HI98509, Hanna Instruments, Lingolsheim, France) to estimate the speed of sound in the water path [30].

To scan a full volume, 411 laser/US transmit sequences were fired in 21 s. After scanning, radio frequency signals received after PA and US events, corresponding motor positions, and pyrometer values were recorded for subsequent processing and image reconstruction. Video 1 is provided to show the movement of the transducer and the flashes for PAI.

B. Image Reconstruction

1) Image Reconstruction Algorithm: The 3-D image grid was defined in a fixed Cartesian coordinate system (O, $\mathbf{e_x}$, $\mathbf{e_y}$, $\mathbf{e_z}$). The vectors $\mathbf{e_y}$ and $\mathbf{e_z}$ correspond to the rotation axis and the radial direction, respectively, when the rotation angle equals zero. The vector $\mathbf{e_x}$ completes the orthonormal basis. The origin O is chosen as the orthogonal projection of the center of the array on the rotation axis when the translation stage is at its central position ($\ell = 0$). In this grid, the voxel dimensions ($p_x \times p_y \times p_z$) were selected to be $71 \times 143 \times 71$ μm to correspond to the expected anisotropic resolution along each axis. The image volume is defined by a diamond-shaped, cross-sectional area [DSCA, Fig. 2(c)] in the xz-plane [length of the diagonal $L = 30$ mm, and centered at ($x = 0$, $z = 25$ mm)] and the active length of the array (19 mm) along the y-axis.

PA signals were divided by the corresponding pyrometer value to compensate for the pulse-to-pulse energy fluctuations of the laser and bandpass filtered between 2 and 10 MHz (Butterworth, order 3).

Image reconstruction was performed with delay-and-sum beamforming algorithms. The one-way (PA) and two-way (US) travel times between the US transducer element positions

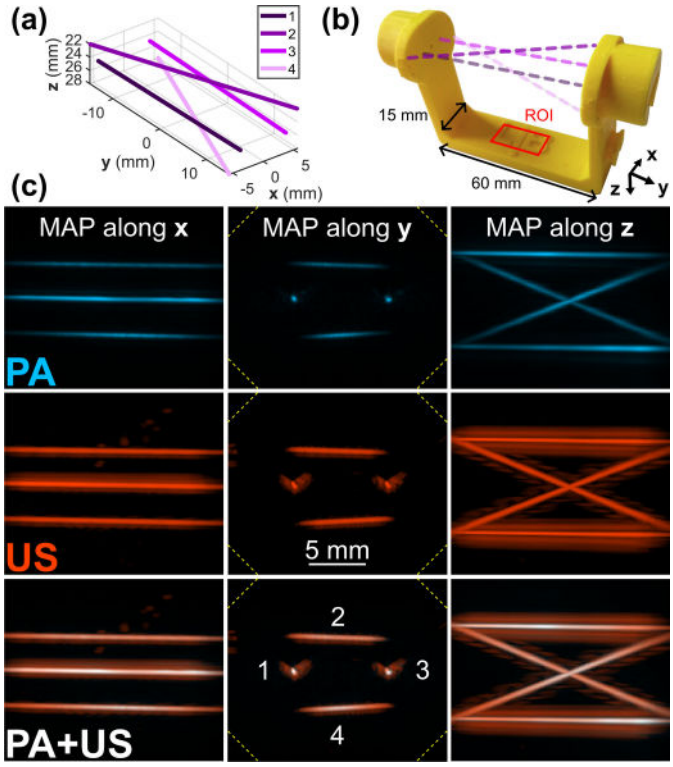


Fig. 2. (a) Imaged region [also called region of interest (ROI)] of the calibration phantom in a 3-D coordinate system with dimensions in mm. (b) Picture of the calibration phantom. Four 20-μm nylon threads are mounted on a yellow 3-D-printed frame. To ease the readability, the threads have been highlighted with the same color as in (a). A projection of the ROI is shown in red. (c) Volumetric images reconstructed with a set of optimized parameters: first row: PA image; second row: US image; and third row: combined PA/US image. The colorscale is presented in Fig. 1(c). Each image is a MAP image. The visible part of the DSCA is shown with yellow dashed lines in the MAP images along y. First column images are 19-mm wide (y-axis) $\times$ 16-mm high (z-axis) and centered at ($y = 0$ mm and $z = 25$ mm); second column images are 19-mm wide (x-axis) $\times$ 16-mm high (z-axis) and centered at ($x = 0$ mm and $z = 25$ mm); third column images are 19-mm wide (y-axis) $\times$ 16-mm high (x-axis) and centered at ($x = 0$ mm, $y = 0$ mm).

(x_n, y_n, z_n) and each imaged voxel (x_p, y_p, z_p) were computed, assuming a constant speed of sound, c, in the medium. Each voxel's value resulted from the sum of signals received by all array elements and tomographic positions arriving within the time-of-flight range estimated to correspond to propagation times to the voxel. The US image reconstruction algorithm was detailed in [27] and assumes that the "plane wave" emissions by the linear array correspond to cylindrical waves in the 3-D space. For both PA and US image reconstruction, the apodization included a dynamic aperture along the axis of the array to maintain a constant angular aperture (lateral f-number of 1.3 with a Hamming window) and a mask to account for the elevational focus of the array (elevational thickness of 1.2 mm with a 20% cosine taper).

Three-dimensional, envelope-detected images were obtained and were displayed using maximum amplitude projection (MAP) along the coordinate system axes. Rotating MAP images around the z-axis were obtained with the 3-D project option of ImageJ [31]. The color scale used for the images is displayed in Fig. 1(c).

2) Time Delays for the Reconstruction: The recording of the US signals was set to start at the same time as the laser emission for the PA acquisition or at the same time as the US emission for the US acquisition. Because of the lens effects [32] and additional time delays induced by the acquisition hardware, we found that two effective temporal offsets, $t_{0\,\mathrm{PA}}$ and $t_{0\,\mathrm{US}}$, were, however, needed to be determined to convert the voxel-associated travel times into the time indexes of the recorded signals. The offset, $t_{0\,\mathrm{PA}}$, was determined experimentally. Two 20-μm diameter black nylon threads (NYL02DS, Vetsuture, France) were positioned perpendicularly to the imaging plane of the array and in the vicinity of the elevation focus (around 25 mm from the face of the array). They were illuminated with the laser light and PA signals were recorded. Given their small diameter with regards to the wavelength at the center frequency of the array ($\Lambda_c \approx 300\ \mu$m), the threads can be assumed to be point absorbers in 2-D. Two-dimensional images ($p_x \times p_z = 71 \times 71\ \mu$m) were beamformed for different values of $t_{0\,\mathrm{PA}}$, using the temperature-adjusted speed of sound in the water bath. The one-dimension Brenner's gradient of the images was computed along the lateral dimension of the array

$$F_{\mathrm{Brenner\ 1D}} = \sum_{y,z} (f(y+2,z) - f(y,z))^2 \tag{1}$$

where $f(y,\ z)$ represents the gray level intensity in the 2-D image. The Brenner's gradient provides a quantitative measurement of image sharpness and has been shown to be an efficient metric for the speed of sound calibration in PA [33]. It was maximized for $t_{0\,\mathrm{PA}} = -1.3\ \mu$s. In the recorded signals, $|t_{0\,\mathrm{PA}}|$ also corresponded to the time of arrival of PA signals generated by a metalized mylar film (space blanket) pressed against the face of the array.

3) Spatial Transformation From Motor to Array Positions: For each tomographic position, the experimental acquisition system records the motor positions in their own coordinate systems (integrated position sensors). The reconstruction algorithm, however, requires the position (x_n, y_n, z_n) and orientation of the array elements within the fixed coordinate system (O, $\mathbf{e_x}$, $\mathbf{e_y}$, $\mathbf{e_z}$) describing the image grid.

Each tomographic position is described by the linear and angular motor positions (ℓ, α), respectively. The translation length equals zero ($\ell = 0$) when the translation stage is at the center position, and the rotation angle equals zero ($\alpha = 0$) when the radial axis is vertical. The rotation axis is, by definition, parallel to $\mathbf{e_y}$. We name $\mathcal{R}_\alpha$ the rotation matrix around the y-axis due to the motion of the rotation motor. The unit translation vector $\mathbf{t}$ is expected to be $\mathbf{e_x}$, but mechanical misalignments require the addition of two angular parameters: φ, the (azimuthal) angle measured from $\mathbf{e_x}$ to the orthogonal projection $\mathbf{t_p}$ of $\mathbf{t}$ on the xz plane, and θ, the angle from $\mathbf{t_p}$ to $\mathbf{t}$. Therefore,

$$\begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} = \begin{pmatrix} \cos(\theta)\cdot\cos(\varphi) \\ \sin(\theta) \\ \cos(\theta)\cdot\sin(\varphi) \end{pmatrix}. \tag{2}$$

The central point of the active transducer aperture is named O_a, and is located on the interface between the transducer array and the water. The coordinates of O_a in (O, $\mathbf{e_x}$, $\mathbf{e_y}$, $\mathbf{e_z}$) depend on (ℓ, α). We define $\Delta x = x_{Oa}(\ell = 0,\ \alpha = 0)$ and $\Delta z = z_{Oa}(\ell = 0,\ \alpha = 0)$

$$\begin{pmatrix} x_{Oa}(\ell,\alpha) \\ y_{Oa}(\ell,\alpha) \\ z_{Oa}(\ell,\alpha) \end{pmatrix} = \mathcal{R}_\alpha \cdot \begin{pmatrix} \Delta x \\ 0 \\ \Delta z \end{pmatrix} + \ell \cdot \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}. \tag{3}$$

Additionally, the spatial transformation matrix from the mobile Cartesian coordinate system (O_a, $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$) attached to the transducer array to the fixed coordinate system (O, $\mathbf{e_x}$, $\mathbf{e_y}$, $\mathbf{e_z}$) needs to be assessed. The vector $\mathbf{u}$ corresponds to the elevation direction, the vector $\mathbf{v}$ to the long axis along the row of the elements, and the vector $\mathbf{w}$ to the axial direction of the linear array. The transformation matrix is written as

$$\begin{pmatrix} u_x(\alpha) & v_x(\alpha) & w_x(\alpha) \\ u_y(\alpha) & v_y(\alpha) & w_y(\alpha) \\ u_z(\alpha) & v_z(\alpha) & w_z(\alpha) \end{pmatrix} = \mathcal{R}_\alpha \cdot \mathcal{R}_{\mathrm{Pitch}} \cdot \mathcal{R}_{\mathrm{Yaw}} \cdot \mathcal{R}_{\mathrm{Roll}}. \tag{4}$$

$\mathcal{R}_{\mathrm{Roll}}$, $\mathcal{R}_{\mathrm{Pitch}}$, and $\mathcal{R}_{\mathrm{Yaw}}$ are rotation matrices around the first axis, the second axis and the third axis, respectively. The product of these rotation matrices models the misalignment of the array axis $\mathbf{v}$ compared to the rotation axis of the stage $\mathbf{e_y}$.

Therefore, the position of the center of the element number n of the array ($n \in [\![1, N]\!]$ with $N = 64$) in (O, $\mathbf{e_x}$, $\mathbf{e_y}$, $\mathbf{e_z}$) can be decomposed as

$$\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = \begin{pmatrix} x_{Oa}(\ell,\alpha) \\ y_{Oa}(\ell,\alpha) \\ z_{Oa}(\ell,\alpha) \end{pmatrix} + \left(n - 1 - \frac{N-1}{2}\right) \cdot p \cdot \begin{pmatrix} v_x(\alpha) \\ v_y(\alpha) \\ v_z(\alpha) \end{pmatrix} \tag{5}$$

with p the interelement spacing of the array. Here, $p = 298\ \mu$m.

Finally, a total of seven geometrical parameters independent of (ℓ, α) need to be determined: Roll, Pitch, Yaw, Δx, Δz, θ, φ.

C. Calibration Method

Inadequate estimation of reconstruction parameters results in degraded image quality in terms of sharpness and will lead to misalignment of the PA and US images. Therefore, we developed a calibration method based on a specific phantom and an optimization algorithm to determine the required seven reconstruction parameters: $t_{0\,\mathrm{US}}$, Roll, Pitch, Yaw, Δx, Δz, θ. The parameter φ was set to be equal to zero because the perpendicularity between the translation axis and $\mathbf{e_z}$ was precisely ensured by the mechanical design and because solutions for our optimization algorithm are equivalent for different sets of (Pitch, φ) values.

The calibration method only needs to be repeated if the position of the array within the scanning system is readjusted (for example, by removing the transducer from the support). The calibration, however, remains valid if the position of the illumination is modified relative to the US array.

1) Calibration Phantom: The calibration phantom needs to be simple to build, easy to use, and should not require an absolute and tedious positioning procedure. We, therefore, developed a wire phantom inspired by phantoms used for the calibration of freehand 3-D US systems such as Z-fiducial

phantoms [34]. Wires or threads have several advantages. First, black threads have strong contrast compared to water for both US and PAI, and are therefore, expected to be clearly visible for both modes and superimposed on the dual-modality images. Second, straight threads provide elongated and uniform structures that can be easily intersected and identified in a volumetric image, even with a sparse sampling in one direction (as opposed to small spheres, for instance). Third, their orientation can be varied. Finally, by using a few noncrossing and well-separated threads, the segmentation of the 3-D image allows local assessments of the image quality by characterizing the image sharpness in specific zones that intersect a thread.

Our calibration phantom is presented in Fig. 2. It is comprised of four threads with two orientations: two threads parallel to one another and positioned in a horizontal plane (Thread 1 and 3), and two threads rotated by $\pm 22°$ placed in parallel planes below and above, respectively (Thread 2 and 4). The angle was chosen to provide good sensitivity to the different parameters to be estimated. The spacing between the threads was chosen so that the phantom fits inside the imaged volume when the parallel threads are roughly aligned along the y-axis and at $z \approx 25$ mm. The orientation of the phantom was chosen so that the threads appear as point-like shapes in xz planes whose reconstruction sharpness is highly sensitive to the tomographic positions.

The experimental calibration phantom was implemented with 20-μm diameter black nylon threads (NYL02DS) mounted on a 3-D-printed frame [see Fig. 2(b)].

2) Numerical Calibration Phantom: Because the values of the parameters estimated experimentally with the calibration phantom could not be determined with a reference measurement, we validated the calibration method using a numerical simulation in which the set of parameters $\chi_c = (t_{0\,\mathrm{USc}}, \mathrm{Roll}_c, \mathrm{Pitch}_c, \mathrm{Yaw}_c, \Delta x_c, \Delta z_c, \theta_c)$ is an input.

For the input set of parameters $\chi_c, \varphi_c = 0$ and a given motor position (ℓ, α), the orientation $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ of the array and the position (x_n, y_n, z_n) of each element were computed from (2) to (5). An experimental dataset of motor positions was used to simulate the entire scan.

The simulation of the US signals (PA or US) for a given motor position and a given element of the array was based on a semianalytical method [35] for linear wave propagation in a 3-D homogeneous medium without attenuation. To limit the frequency content of the simulated signals, we calculated the convolution of the impulse response directly for a specific waveform: a one-cycle sinusoidal signal at 5 MHz with a Gaussian envelope for the US simulation, and the derivative of a Gaussian pulse (standard deviation 17 ns) for the PA simulation. Simulated data were generated at the same sampling frequency as the experimental data.

The 3-D spatial impulse response of the transducer was computed based on the discretization of the Rayleigh integral over the surface of the transducer. The finite size (250 μm-width and 7.5 mm-height) and cylindrical focusing of one array element (focal length 25 mm) were modeled for a cylindrical surface discretized with point transducers spaced by $\Lambda_c/5$. Λ_c is the US wavelength at the center frequency of the transducer ($\Lambda_c \approx 300$ μm). The surface was oriented according to the basis $(\mathbf{u}, \mathbf{v}, \mathbf{w})$, and its center was positioned in (x_n, y_n, z_n). We computed the impulse response of this array element for a given point (x_s, y_s, z_s) in the imaged volume by summing the impulse responses over all the point transducers. For the PA simulation, the point (x_s, y_s, z_s) was considered as a PA source. For the US simulation, pulse-echo signals of (x_s, y_s, z_s) were computed for a titled "plane wave" emission. The "plane wave" generation was computed by summing spatial impulse responses (delayed to induce the steered angle) over all the array elements. The pulse-echo signal for the element number n was then obtained by convolving this sum with the spatial impulse response of n.

To simulate the acquisition of the PA and US signals for the calibration phantom, the threads were discretized along their length into point sources (PA) or point scatterers (US) spaced by $\Lambda_c/5$. The diameter of the threads was not modeled. The superposition principle was used to compute the signal for one element of the array at a given position: the simulated signals were added over all the points (x_s, y_s, z_s) corresponding to the threads.

The input of the numerical simulation is χ_c, the motor positions (ℓ, α) and their corresponding events (PA acquisition or US acquisition with a tilted "plane wave" emission at a given steering angle), the speed of sound and the positions of the point sources (PA) or point scatterers (US). The output is the US signals corresponding to all the PA and US events received by each element of the array. Simulated signals were reconstructed using the method presented in Section II-B1. Point-like objects and threads were found to be reconstructed at their defined positions both in PA and US images, which validated the simulation method.

The numerical calibration phantom reproduced the spatial arrangement of the threads in the experimental calibration phantom [see Fig. 2(a)].

3) Calibration Algorithm: Cost Function: The calibration algorithm is based on an optimization algorithm that minimizes a cost function. The cost function was assessed using reconstructed volumetric images from one complete data acquisition sequence, comprised of PA and US (five steered angles) data sets for the calibration phantom.

a) Calibration images and segmentation of the threads: Each evaluation of the cost function for a set χ of the seven parameters required volumetric image reconstruction. The time needed for this was reduced using sparse sampling of the imaged volume along the y-axis. We reconstructed only five slices located at $y = i \times 3$ mm with $i \in [\![-2; 2]\!]$, i.e., centered and distributed to avoid edge effects. For each slice, we reconstructed two 30-mm width square images (one PA and one US) centered around $(x = 0, z = 25$ mm) and with pixel sizes $p_x \times p_z = 71 \times 71$ μm using the whole dataset and the 3-D reconstruction algorithm described in Section II-B1.

Threads were segmented and identified to calculate several image metrics. Each envelope-detected image was thresholded at one-fourth of its maximum pixel value to produce a binary image. The fourth largest connected components of the binary image were identified as regions, and each region was expected to correspond to one thread. The cost was directly set to zero, meaning that the set χ is rejected if less than four connected components were counted. For each region, the

centroid position (x_j, z_j) was calculated with weights based on the grayscale image intensity value. The distance between the centroids determined in the US image and in the PA image in slice i was

$$d_{\mathrm{US-PA}}(i,j) = \sqrt{\left(x_j^{\mathrm{US}} - x_j^{\mathrm{PA}}\right)^2 + \left(z_j^{\mathrm{US}} - z_j^{\mathrm{PA}}\right)^2}. \quad (6)$$

In each image, the four centroids were sorted by increasing angular position, and each was attributed to one of the threads.

b) Selection process for image-based metrics with the numerical calibration phantom: Since no standard cost function exists for this seven-parameter optimization problem, the cost function was built empirically from observations of how some variable metrics, computed from the images, depended on the different parameters. Our practical goal was to obtain an efficient cost function rather than the optimal one.

An efficient cost function needs to have a global minimum for the targeted set of parameters. Verifying this condition requires that the expected parameter values are known. Consequently, the selection of metrics was performed on the numerical calibration phantom for which χ_c is controlled.

Metrics based on geometric criteria (angles between the threads, straightness of the threads) and metrics of image quality (image amplitude, sharpness) on a local scale around each thread and on the entire image were tested for the PA image and the US image. Metrics describing the PA and US image superposition (image cross correlation, distance between the reconstructed threads) were also considered. The initial screening of the metrics implied graphical visualization of the metric variations with the parameter values. To facilitate this visualization, only one parameter of χ_c was varied while all the other parameters were held constant and equal to their expected values. In this first stage, we chose to set all the parameters of χ_c to zero. Then, the angular parameters were varied over a range of $[-5°, 5°]$ by steps of 0.5°, the distances over a range of $[-3, 3$ mm] by steps of 0.3 mm, and the time parameter over a range of $[-0.5, 0.5\ \mu$s] by steps of 0.1 μs.

We found that the cost function could not be defined with a single metric because none of the tested metrics had a global extremum for all the parameters. Therefore, we performed a careful selection of metrics that were subsequently combined to obtain the cost function. Because of the diversity of metrics, we chose to give a score to each of them and for each parameter. Criteria for the score were the presence of one global extremum at the expected value and steep variations. A selection among the tested metrics was then performed, keeping the ones with the highest scores while making sure that all the parameters were covered by at least one metric.

The selected metrics are presented below (see Section II-C3.c). We combined them with a product. Compared to a weighted sum, the combination with a product enables mixing metrics with different scales without the need to perform a prior normalization but may be more sensitive to fluctuations (noise). With the same procedure as for the test of each metric, we tested that the defined cost function had a minimum at the expected value when each parameter was varied individually (all the other parameters remained constant). For this validation, the set of parameters was fixed to nonzero values to avoid bias.

c) Selected metrics for the cost function: In this section, we present the image-based metrics selected to build the cost function. We determined that some metrics were more sensitive to parameter variations for Threads 2 and 4. Thereby, several quantities were computed for these two threads. First, a linear regression was performed with the five centroids (one per slice) of each thread, and the coefficient of determination was computed. The mean coefficient of determination over the two threads gave R^2_{US} and R^2_{PA} for the US images and the PA images, respectively. R^2_{US} and R^2_{PA} equal one when the threads are reconstructed as straight structures in the volume and when the image quality allows adequate segmentation of the threads. R^2_{US} and R^2_{PA} were found to be sensitive to errors for the parameters Pitch, Δx, and Δz. Second, the mean distance of $d_{\mathrm{PA-US}}$, named $D_{\mathrm{PA-US}}$, was computed over the two threads and the five slices. $D_{\mathrm{PA-US}}$ evaluated the superposition of the PA and US images. $D_{\mathrm{PA-US}}$ is expected to be equal to zero for the correct set of parameters. $D_{\mathrm{PA-US}}$ was found to be sensitive to errors in the parameter $t_{0\,\mathrm{US}}$. Additionally, a thresholding of $D_{\mathrm{PA-US}}$ was found efficient to reject sets of parameters leading to a degraded line spread function (LSF) either in US or in PA. $D_{\mathrm{PA-US}}$ is expressed in mm, and the cost was set equal to zero for $D_{\mathrm{PA-US}} > 1$ mm (meaning that the set χ is rejected). Finally, the local normalized variance of the US images was computed in a square region of 2 mm-width (twice the translation step) around the centroid. The normalized variance quantifies variations in the pixel values about the mean. It is equal to the variance of the pixel values over their mean. This measurement of the image sharpness was reported for an autofocus method [33]. The mean of the local normalized variance over the two threads and over the five slices was named $N_{\mathrm{V\,US}}$. $N_{\mathrm{V\,US}}$ is expected to be maximal for the correct set of parameters. $N_{\mathrm{V\,US}}$ was found to be sensitive to errors in the parameters: Pitch, Yaw, Δx, Δz, θ.

Finally, a normalized 2-D squared gradient of the entire US images (four threads) was computed as a sharpness metric

$$S_{N\,2\mathrm{D}} = \frac{1}{\sum_{x,z} f(x,z)} \left(\sum_{x,z} \left(f(x+1,z) - f(x,z)\right)^2 + \sum_{x,z} \left(f(x,z+1) - f(x,z)\right)^2 \right) \quad (7)$$

where $f(x, z)$ represents the gray level intensity in the 2-D US image. The mean of the normalized squared gradient over the five slices was named $S_{\mathrm{N\,US}}$. $S_{\mathrm{N\,US}}$ was sensitive to errors in the parameters: Roll, Pitch, Yaw, Δx, Δz, and $t_{0\,\mathrm{US}}$.

The selection of the metrics $N_{\mathrm{V\,US}}$ and $S_{\mathrm{N\,US}}$ highlights that the normalized variance and normalized squared gradient were more efficient for US images than for PA images. This could be explained by the fact that the US image reconstruction combines data for five steered angles, which may induce stronger variations when the parameters are away from their expected values.

We combined R^2_{US}, R^2_{PA}, $D_{\mathrm{PA-US}}$, $N_{\mathrm{V\,US}}$, and $S_{\mathrm{N\,US}}$ with a product, and the cost was then defined by

$$\mathrm{Cost} = \begin{cases} 0, & \text{for } D_{\mathrm{PA-US}} \geq 1\mathrm{mm} \\ -R^2_{\mathrm{US}} \cdot R^2_{\mathrm{PA}} \cdot (1 - D_{\mathrm{PA-US}})^2 \cdot N_{\mathrm{V\,US}} \cdot S_{\mathrm{N\,US}}. \end{cases} \quad (8)$$

4) Calibration Algorithm: Optimization Algorithm: The calibration algorithm relies on the cost function. An initial combination of seven parameters was given as input. Two steps were then applied. First, several combinations of parameters were proposed, in which each parameter was drawn at random, following a normal distribution around the initial guess. Costs of these combinations were computed until reaching a total of 100 combinations with nonzero cost. The combination with the smallest cost was used for the second step: the application of a Particle Swarm optimization algorithm. This heuristic algorithm was chosen over a convergent iterative method because of the global optimization problem at hand.

With the numerical calibration phantom, we found variability of the output combination and of the associated cost when the calibration algorithm was repeated on the same dataset. This indicates the local minima of the cost function. To mitigate this variability, the optimization algorithm was repeated 20 times on the same dataset. The 20 combinations were then sorted by increasing cost, and the final combination was obtained by calculating the median of each parameter on the combinations with the five lowest costs. The median was chosen for its ability to limit the weight of strong outliers on the solution.

5) Metrics for the Variability:

a) Acceptability range: The variability of the determination of the parameters was compared to an acceptability range. Three acceptability ranges were defined depending on the unit of the parameter. The acceptability range for length parameters (Δx, Δz) was set equal to the wavelength Λ_c. For $t_{0\,\mathrm{US}}$, we used the wave period at the central frequency of the transducer: 0.2 μs. For the angles, we considered an axial deviation of Λ_c seen from the lateral aperture of the array as a significant error. As the interelement spacing of the array and the element size can be considered equal to Λ_c, the lateral aperture dimension is equal to $64 \cdot \Lambda_c$ and the acceptability range for angles was then set to $\sin^{-1}(\Lambda_c/64 \cdot \Lambda_c) = 15.6$ mrad (0.9°). This value also corresponds to the angular resolution of the array in the lateral direction [36].

b) Variability quantifications: To assess the variability induced by the optimization algorithm, the numerical calibration phantom was used. The simulated set of parameters χ_c reflected the experimental data (mean over ten experiments). The absolute difference between obtained and expected (χ_c) parameters was computed and divided by the acceptability threshold to be expressed as a percentage. This metric evaluated the accuracy of the calibration method and is, therefore, named Ac.

To assess interacquisition variability on experimental data, parameters were obtained for ten acquisitions. These parameters were compared to the mean parameter over the ten acquisitions. For each parameter, the mean absolute difference of the obtained outputs compared to the mean value was computed (also named mean absolute deviation) and divided by the acceptability threshold to be expressed as a percentage. This parameter assesses the repeatability of the entire calibration process and is called Rp.

6) Metrics for the Spatial Resolution and Superposition: We quantified the superposition of the US and PA images using a second phantom comprised of threads (see Ph1, in Section II-D) with a dual contrast. The quantification method assesses the distance between the images of each thread in US and PA imaging. As for the image processing used in the calculation of $D_{\mathrm{US-PA}}$ on the calibration phantom, we determined the positions of the centroids (both in the US images and the PA images) for each thread and five slices located at $y = i \times 3$ mm with $i \in [\![-2; 2]\!]$. Each slice was reconstructed with pixel sizes $p_x \times p_z = 71 \times 71$ μm. The mean distance of $d_{\mathrm{US-PA}}$ [see (6)] and its standard deviation over the slices and over the threads of the same material (nylon or polyester) measured the superposition quality.

For each nylon thread, the spatial resolution was estimated by fitting the images in each slice i with a 2-D-Gaussian model equation

$$f(x, z) = A \cdot \exp\left(-\left(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(z - z_0)^2}{2\sigma_z^2}\right)\right) \quad (9)$$

where A is the amplitude, x_0, z_0 are the centroid positions and σ_x, σ_z are the standard deviations along x and along z. The full width half maximum was calculated along x and z as $\mathrm{FWHM}_i = 2(2\ln 2)^{1/2} \times \sigma_i$. Because nylon threads can be considered small compared to Λ_c, the FWHM is an estimate of the width of the LSF.

D. Imaging Phantoms

The first phantom (Ph1) was a wire phantom designed to evaluate the coregistration capabilities. Ph1 differs from the calibration phantom in the spatial arrangement of the threads and because half of the threads are of a different material. Three 20-μm diameter black nylon threads (NYL02DS) and three black polyester threads (Coat Epic 150) were mounted on a 3-D-printed frame similar to Fig. 2(b). They were arranged symmetrically with respect to the center of the frame and with various angles.

The second phantom (Ph2) was made by suspending 2% w/v agar powder (A1296, Sigma Aldrich, St. Louis, MO, USA) and 1% w/v cellulose powder (Sigmacell cellulose Type 20, Sigma Aldrich) in water to form a gel cast in a cylindrical mold (20-mm diameter) with three mold-length, cylindrical solid inclusions that were 5 mm in diameter. Within this gel, the cellulose forms particles of approximately 20 μm, which act as US scatterers to mimic the US scattering properties of biological tissues. The agar-cellulose solution was heated to 85 °C and poured into the mold. When the mold was half full, 100-μm-diameter black polyethylene microspheres (BKPMS 90–106 μm, Cospheric, Santa Barbara, CA, USA) were spread on the superior interface and were trapped at the interface during the solidification of the gel. These black microspheres act as optical absorbers that can generate a PA signal. The mold was then filled with the hot agar solution. When the gel solidified, the cylindrical inclusions were removed and filled with water. The microspheres remained embedded in the gel. Ph2 was placed so that the cylindrical holes and the plane of spheres were parallel and perpendicular to the rotation axis, respectively.

The third phantom (Ph3) was prepared with agar powder (2% w/v) and cellulose powder (1% w/v) in water for the first half and with agar powder (2% w/v) for the second half. Two

TABLE I
ACCURACY AND REPEATABILITY STUDY ON THE CALIBRATION

	Yaw	Pitch	Roll	θ	Δx	Δz	$t_{0\,US}$
Ac (%) [a]	2.8	18	55	20	25	32	24
Rp (%) [a]	28	8.3	8.6	8.5	17	39	1.1

[a] Ac and Rp are expressed in percentage of the acceptability ranges.

TABLE II
RESOLUTION AND SUPERPOSITION ASSESSMENT

	FWHM nylon (μm)		Superposition (μm)	
	US	PA	nylon	polyester
along x (mean ± std)	212 ± 18	387 ± 30	15 ± 9	23 ± 17
along z (mean ± std)	296 ± 24	379 ± 21		

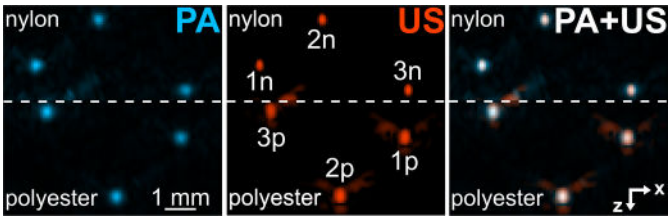


Fig. 3. Image reconstruction of the central plane of Ph1 ($y = 0$). All images are 7-mm wide (x-axis) × 6.7-mm high (z-axis) and are centered at ($x = 2$ mm and $z = 25$ mm). The horizontal white dashed line represents the separation between nylon and polyester zones. Threads with the same number are symmetrical and n stands for nylon and p for polyester. PA image (left) is presented on a linear scale while the US image (center) is presented in dB with a threshold at −30 dB.

crossed 20-μm diameter black nylon threads were embedded in the gel.

III. RESULTS

A. Accuracy and Repeatability of the Calibration

The accuracy and the repeatability of the calibration outputs are presented in Table I, in percentage of the acceptability range. First, it can be noticed that all seven parameters are fully within the acceptability ranges for both Ac and Rp.

For accuracy, the calibration outputs were compared to the ground truth thanks to the numerical simulation. The mean Ac over the seven parameters was found to be equal to 26%, and Ac had a maximum of 55% for the roll parameter. We can thus consider that the developed calibration method enables to accurately determine the set of parameters.

To evaluate the repeatability, ten acquisitions were performed on various days (distributed in three imaging sessions over one week). For each acquisition, the optical fibers and the phantom were repositioned to avoid any bias. Rp estimates the mean absolute deviation of each parameter. The mean Rp over the seven parameters is around 16%, and a maximum of 39% was reached for Δz. Therefore, the calibration is repeatable and stable over time.

Despite slight variations in the evaluation of each parameter, the accuracy and the repeatability of the calibration method are highly satisfying. We can therefore consider the developed calibration method to be reliable and robust.

B. Superimposition Quality and Spatial Resolution

Fig. 3 presents PA/US images of the slice $y = 0$ (center of the linear array) of Ph1. Ph1 aims to test if the experimental calibration remains valid for a phantom different than the calibration phantom in the spatial arrangement of the threads but also in the thread material. The phantom Ph1 is comprised of three nylon threads and three polyester threads arranged so that each thread has a different orientation. Only the threads numbered 2 were set parallel to the y-axis. For an easier comparison between the two materials, Ph1 was built so that each of the three nylon threads had a symmetrical polyester thread with respect to the center of the frame. In Fig. 3, symmetrical threads have the same number, and the suffixes "n" and "p" refers to nylon and polyester threads, respectively. For instance, the thread 1n is symmetrical to 1p, with respect to the center of the image. The nylon threads were grouped on the top part of the phantom, while the polyester threads are grouped in the bottom part. Video 2 displays rotating MAP images around z of Ph1 and therefore shows the spatial arrangement of the threads. As nylon threads are thinner than polyester ones, both US and PA signals were weaker for nylon. The amplitude in the reconstructed image was 3× smaller in PA. To facilitate the visualization in Fig. 3, the upper and the lower part of the volumetric images were normalized by their local maximum and not the global one. The separation between the two parts is illustrated by the horizontal dashed white line in Fig. 3. We can visually see that polyester threads appear larger than the nylon ones both on the PA and US images. This is expected, given their larger diameter. The superposition of the images can, however, be observed for the two materials and for all the threads regardless of their orientation or position in space (Fig. 3 and Video 2). The blue lateral halo around the white spots in the combined image [see Fig. 3(left)] indicates that the lateral resolution is wider for PA images.

For a more quantitative description, Table II presents the superposition distance and the FWHM calculations. The superposition distance between the center of a thread in PA and in US is small compared to the US wavelength Λ_c (less than 10% of Λ_c) and smaller than the spot obtained for each object in the image. The superposition distance is not significantly different for the two materials. This result indicates that the calibration enables the superposition of nylon threads with different orientations than in the calibration phantom, and it validates that the calibration ensures the superposition for objects with a dual contrast but in another material. The standard deviation computed over the three threads and five imaging planes is also small compared to Λ_c. This result shows the low dispersion of the superimposition distance both with the thread orientation and with the spatial position in the imaged volume.

The FWHMs calculated for the nylon threads provide an estimate of the LSF because the thread diameter is much smaller than Λ_c. We first notice that for both US and PA and for both x- and z-directions, the FWHM is on the order of magnitude of Λ_c. Along the x-direction, the resolution is limited by the diffraction. The high resolution in the x-direction results from the large angular aperture provided by the rotation scan and the synthetic aperture

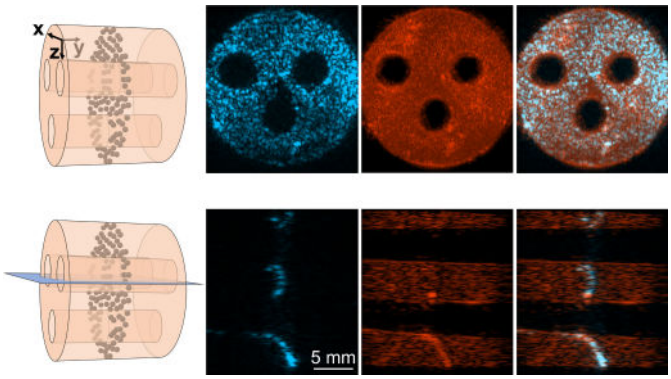


Fig. 4. Image reconstruction of Ph2. Schematic drawing of Ph2 are displayed in the left. The first row displays MAP images along y while the second row shows a slice perpendicular to the z-axis. PA images are presented in linear color scale. Because a few bright spots (probably clusters of microspheres) were dominating the color scale and hiding the rest, voxel values were, however, saturated at 40% of their maximum before normalization. US images were thresholded at -40 dB. First row images are 19-mm wide (x-axis) $\times$ 21-mm high (z-axis) and centered at ($x = -0.6$ mm and $z = 24.3$ mm); second row images are 19-mm wide (x-axis) $\times$ 21-mm high (y-axis) and centered at ($x = 0$ mm and $y = 0$ mm).

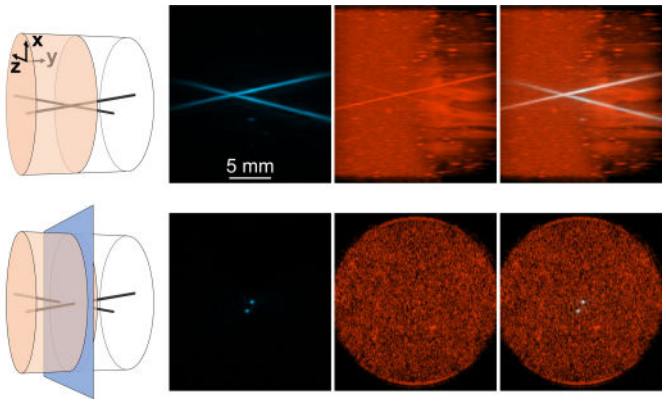


Fig. 5. Image reconstruction of Ph3. Schematic drawing of Ph3 are displayed in the left column. The first row displays MAP images along z that are 19 mm wide (y-axis) $\times$ 21 mm high (x-axis) while the second row shows a slice perpendicular to the y-axis that is 19 mm wide (x-axis) $\times$ 21-mm high (z-axis) and centered at $x = 0$ mm, $z = 25.4$ mm. US images were thresholded at -40 dB. Along y-axis, left part is agar with cellulose and right part is agar alone.

approach. For comparison, the FWHM_x was on the order of 1 mm [27] for a translation-only scan with the same US array. The low standard deviation of FWHM_x indicates that the resolution is independent of the position of the object in the volume and its orientation. The spatial homogeneity associated with the rotate-translation scheme and previously observed independently in US imaging [27] and PAI [26] is then confirmed for the simultaneously coregistered imaging.

In each direction, the US resolution is slightly better than the PA resolution (70%–100% of Λ_c for US versus 130% of Λ_c for PA). Along x, two main factors can explain the FWHM_x differences between US and PA. First, PA images rely on US signals produced by the illuminated object and not on backscattered US signals generated by the US array. For small objects, the US frequency spectrum recorded by the array is then usually broader in PAI than in US imaging and especially contains low frequencies which may decrease the diffraction-limited resolution. Second, for one laser excitation, five tilted plane waves are emitted, which increases the number of independent views and the spatial sampling in US imaging compared to PAI. This sampling factor has been shown to influence the resolution along x [27]. Along z, which corresponds to the axial direction, the LSF is mainly influenced by the pulse duration. The pulse-echo mechanism can explain the better resolution of the US images compared to PA.

C. Complementary Distributions of US and PA Contrasts

To further demonstrate the advantages of dual-modality imaging, phantoms with complementary contrasts were designed and produced. The images of Ph2 and Ph3 are displayed in Figs. 4 and 5, respectively.

For the US contrast, Ph2 is a homogeneously scattering medium (agar with cellulose) with three cylindrical and anechoic holes filled with water. For the PA contrast, numerous black-dyed micro-spheres are arranged in a central plane (see Fig. 4). The optical absorption of the agar gel and the water is negligible. In the first row of Fig. 4, a MAP image along y is shown for PA imaging, US imaging and the superposition of both. Rotating MAP images round the z-axis are presented in Video 3. The 20-mm diameter cylindrical shape of the phantom could be retrieved both in PA and US imaging. One can note the homogeneous image quality in the xz plane for both modalities. In the second row of Fig. 4, images of a slice perpendicular to the z-axis are presented. The slice was chosen to cut two of the three holes. The homogeneity of the US image along the y-axis can be observed. As expected, the three holes appear with a negative contrast for both PA and US imaging, independently of their spatial position. The holes allow to further validate the superimposition of the two modalities. They are concentric in US imaging and in PAI. We can notice that the outlines of the holes are blurrier and the diameter of holes seems smaller in the US images. This effect can be attributed to the stronger side lobes in the US images induced by the log compression (display in dB) of the color scale compared to the linear scale used in PAI. Such side lobes are also visible in Fig. 2(c). In US imaging, the microspheres have a contrast relative to the agar matrix only when they are numerous and clumped (diagonal line at the bottom of Fig. 4-second row) or when they are distributed along a surface that visually integrates their contribution in MAP rendering (Video 3). In Fig. 4-second row, however, microspheres are hardly visible in the middle and top parts of the US image, while they appear with strong contrast in PAI. With the superposition, both the agar and the microspheres are visible. This is of interest to assess the distribution of PA contrast agents in an organ, which, for instance, is homogenously echoic. The US image gives the contour of the phantom, which is similar to the anatomical context, while the PA image gives the distribution of the marked spheres, which are analogous to a contrast agent. In a clinical application, the PA contrast agents could be therapeutic nanoagents accumulating locally. Holes are mimicking bodies with a negative contrast such as cysts.

For the second phantom Ph 3, media with two different US contrasts were used: agar with cellulose is more echogenic than agar alone. Two black nylon threads bring the optical absorption contrast. Images of Ph3 are presented in Fig. 5. The US image in the first row (MAP along z) and the rotating

MAP images in Video 4 clearly show the contrast between the two blocks of gel. The nylon threads are visible in the MAP US images, but mainly due to their elongated shape as they are barely differentiable from the surrounding speckle in the slice perpendicular to the y-axis and intersecting the two threads in the agar with cellulose part (second row of Fig. 5). Agar and agar with cellulose do not have any contrast in PA. The threads, however, show a homogenous and high contrast in PAI over the two agar blocks. This phantom could typically mimic a blood vessel (black thread), highly visible in PAI, perfusing two different organs with different echogenicity.

IV. Discussion

We demonstrated high-quality, volumetric, and simultaneously coregistered PA and US imaging. The simultaneous dual imaging was made possible by the use of a linear US transducer array, and the high quality (resolution, contrast, visibility) of the volumetric images resulted from the large synthetic angular aperture enabled by the rotate-translate scan geometry. Images showed a homogenous quality over a large imaged volume (cylinder of diameter of 21 mm and length of 19 mm). The effective synthetic aperture for both imaging modalities and the superposition of the PA and US images for features having a dual contrast required an accurate determination of the positions of the US array. To this end, we developed and validated a calibration method, which was determined to be both accurate and repeatable. This initial calibration process allowed the reconstruction of images from subsequent acquisitions without the need of fiducial markers on every imaged volume under conditions for which the speed of sound can be considered constant in the imaged volume. We demonstrated the superposition of PA and US images with phantoms having a dual contrast and the complementarity of the mapped information with phantoms having complementary spatial distributions of US and PA contrast agents.

The calibration was based on the combination of a dedicated calibration phantom, a cost function, and an optimization process. Seven parameters were determined for our scan geometry to obtain superimposed and sharp PA and US images. The mechanical mount and the position sensors of the stages ensure that the determined parameters remain valid for subsequent scans as long as no deformation or accidental misalignment of the array occurs. For this study, no significant loss of calibration was observed over a period of one week. The calibration phantom was easy to build and was comprised of four well-separated threads to facilitate the identification and measurement of local image properties. The cost function used only five imaging planes to avoid the reconstruction of the entire volume and the associated long computation time. Because metrics with the sharpest variations with regard to the parameters were selected for the cost function, it was dominated by the properties of the US image but was also influenced by the properties of PA images. For the optimization algorithm, we found that the selection of the initial guess was a crucial step due to the presence of local minima in the cost function. Additionally, the Particle Swarm optimization was found to be more effective in determining a solution close to the ground truth than a downhill simplex method. The refinement of the optimization algorithm is beyond the scope of this study since we focused on obtaining an effective method that is simple to apply. Further refinement will be considered in a future study.

This calibration procedure was found to be effective for our system, although it requires the initial imaging of a dedicated phantom. In the past decade, several multiperspective or multiview imaging systems with US transducer arrays have been developed both in PAI and in US imaging, independently, leading to various calibration procedures. For US imaging, coherent compounding of images acquired with two US transducer arrays has been developed in 2-D [37] and 3-D [38]. Dedicated calibration phantoms were also used that consisted isotropic scatterers (five wires in 2-D and three spheres in 3-D). The phantoms were combined with a cost function linked to the coherence of the echoes from the isotropic scatterers received by the different elements of the array to determine geometrical parameters (four parameters in 2-D and six in 3-D). A simplex search method was used. The calibration showed an improvement in the contrast and in the resolution of the images, but the use of isotropic scatterers resulted in a long computation time for volumetric images during the calibration process. Incoherent compounding of US images was also investigated for 2-D [39] and 3-D [40] imaging. The spatial transformation matrix between the positions of the arrays was assessed directly using the images acquired on the sample of interest. This approach could enable free hand scanning and imaging of moving organs but could not achieve a synthetic aperture approach which, therefore, limits the final image quality. In PAI, multiperspective or multiview imaging is almost inherent to the modality since PA tomography with spherical and hemispherical scans of a US detector has been investigated in the early PA scanners [41]. Planar detection geometries able to acquire 3-D PA images have, however, been angulated to enhance the image view with a synthetic aperture approach [42], [43]. In [42], two planar arrays were assembled in a rigid configuration. An initial calibration was performed by imaging a dedicated phantom comprised of three threads with different orientations with each array independently. Thread position and orientation were then identified in each segmented volume to determine the rigid transformation between arrays. In [43], fiducial markers were incorporated in the imaged region. Multiview imaging was also performed by stitching together volumetric images [44], which is similar to incoherent compounding and relies on the features obtained in partially overlapping images. In brief, no standard calibration exists either in US or PAI. To date, only dedicated calibration phantoms have been shown to allow synthetic aperture reconstruction without fiducial markers. Additionally, when using a single imaging modality, some reconstruction parameters may compensate for errors in other reconstruction parameters. For volumetric and simultaneous PA and US imaging, the superposition and sharpness of the dual-mode tomography can only be obtained if each parameter is determined accurately. In particular, this accuracy is necessary because the propagation times between transducer elements and the voxels are different for PAI and US imaging.

Coregistered volumetric PA and US maps provide complementary information (see Figs. 4 and 5). In the framework of longitudinal studies with repeated imaging sessions, the US image is expected to give the anatomical reference so that the

PA response can be better positioned within the anatomy and so that robust coregistration can be performed between data acquired at different time points. PA response can be applied to reveal molecular or functional phenomena with a slow kinetic, such as the accumulation of nanoparticular contrast agents with a long circulation time.

For other simultaneously coregistered multimodal imaging, such as PET/CT, it has been shown that the anatomical imaging modality contains information that can be used to improve the image quality of the functional imaging modality. Recent work on the reconstruction of 2-D PA images has shown that the image quality of blood vessels in PA images can be improved using structural information from US images [45]. Additionally, light fluence distribution could be modeled using US images to further improve PA image quality [46].

Our image reconstruction algorithms adapted conventional delay-and-sum beamforming approaches to the tomographic system. While these algorithms led to high-quality images for the tested phantoms, more complex image reconstruction algorithms will be investigated in future studies to incorporate the acoustic properties of the sample or of the transducer. Indeed, heterogeneous speed of sound or aberrating layers in the sample would result in degraded image quality with our algorithm but may be considered with iterative reconstruction approaches [47], [48]. Modeling the spatial impulse response of the transducer in the reconstruction algorithm could allow a further improvement of the image quality and a reduction of the number of scan positions, and consequently, of the scan duration [49]. The motion of the sample during the scan is expected to induce blurring artifacts. Because of the interlaced PA and US events, motion artifacts are, however, expected to be similar for the two imaging modalities. These artifacts may be reduced by gating the acquisition with physiological signals to select phases for which the imaged volume is still or has returned to the same position.

The study presented here was performed at a single optical wavelength. Multispectral approaches [50] will be investigated with the developed scanner to enable discrimination between different chromophores and contrast agents and to evaluate physiological parameters such as oxygen saturation. On the US side, the system presented here operated at a US center frequency of 5 MHz. The rotate-translate scan can, however, be scaled to other US frequencies to gain sensibility and information on other spatial scales [26], [51]. The spatial resolution increases with the US frequency of the array for both PA and US imaging, but higher frequencies will also improve the sensitivity to small absorbing regions in PAI. Linear US arrays are commercially available in a wide range of center frequencies.

Building on the in vitro proof of concept presented here, we are currently adapting the scanner for in vivo imaging with the addition of an acoustic coupling system to remove the large water tank. A small water tank with an acoustically transparent membrane sealing the bottom [20] will be investigated. For the translation toward clinical applications, the bulky motorized stages will be replaced by a scanning system with a smaller footprint, such as those that have been previously developed for translational scans [20], [21].

The imaged volume (cylinder of a diameter of 21 mm and length of 19 mm centered at a depth of 25 mm), submillimeter spatial resolution, and volumetric imaging rate (~20 s per volume) can usefully offer dual-mode contrast for preclinical investigations of murine models [44], [52]. Potential clinical applications that could benefit from the combined, volumetric PA and US imaging offered by this technique include the evaluation of inflammation (finger and hand joints [53], for instance) and cancerous lesions (such as thyroid cancer [54]) in relatively superficial zones of the body.

Conclusion

In this study, the application of rotate-translate scan tomography of a linear US array to bimodal US and PA imaging allows the acquisition of high quality simultaneous volumetric images. In order to ensure optimized superposition of the two modalities, a dual 3-D calibration method is proposed for the first time. The resulting superposition distance between US and PA volumes is better than 10% of the acoustic wavelength. This work paves the way to sensitive and robust monitoring of slow-kinetic molecular and functional phenomena in living systems in a dual mode using photoacoustics to monitor nanoagent accumulation and US to localize this accumulation within anatomical structures.

Acknowledgment

Clément Linger is with the Laboratoire d'Imagerie Biomédicale (LIB), CNRS, INSERM, Sorbonne Université, 75006 Paris, France, and also with the Institut Galien Paris-Saclay (IGPS), CNRS, Université Paris-Saclay, 91400 Orsay, France (e-mail: clement.linger@sorbonne-universite.fr).

Yoann Atlas, Marine Vandebrouck, S. Lori Bridal, and Jérôme Gateau are with the Laboratoire d'Imagerie Biomédicale (LIB), CNRS, INSERM, Sorbonne Université, 75006 Paris, France (e-mail: lori.bridal@sorbonne-universite.fr; jerome.gateau@sorbonne-universite.fr).

Rémy Winter was with the Laboratoire d'Imagerie Biomédicale (LIB), CNRS, INSERM, Sorbonne Université, 75006 Paris, France. He is now with the Unité de Recherche Clinique Côte d'Azur (UR2CA), Université Côte d'Azur, 06000 Nice, France (e-mail: remy.winter@etu.univ-cotedazur.fr).

Maxime Faure was with the Laboratoire d'Imagerie Biomédicale (LIB), CNRS, INSERM, Sorbonne Université, 75006 Paris, France. He is now with the Laboratoire de Traitement de l'Information Médicale (LaTIM), Inserm, Université de Bretagne Occidentale, 29220 Brest, France (e-mail: maxime.faure@univ-brest.fr).

Théotim Lucas is with the Laboratoire d'Imagerie Biomédicale (LIB), CNRS, INSERM, Sorbonne Université, 75006 Paris, France, and also with the Matière et Systèmes Complexes (MSC), CNRS, Université Paris-Cité, 75006 Paris, France (e-mail: theotim.lucas@sorbonne-universite.fr).

References

[1] K. Heinzmann, L. M. Carter, J. S. Lewis, and E. O. Aboagye, "Multiplexed imaging for diagnosis and therapy," *Nature Biomed. Eng.*, vol. 1, no. 9, pp. 697–713, Sep. 2017, doi: 10.1038/s41551-017-0131-8.

[2] S. R. Cherry, "Multimodality imaging: Beyond PET/CT and SPECT/CT," *Seminars Nucl. Med.*, vol. 39, no. 5, pp. 348–353, Sep. 2009, doi: 10.1053/j.semnuclmed.2009.03.001.

[3] Z. Hosseinaee, J. A. T. Simmons, and P. H. Reza, "Dual-modal photoacoustic imaging and optical coherence tomography [review]," *Frontiers Phys.*, vol. 8, pp. 1–19, Jan. 2021, doi: 10.3389/fphy.2020.616618.

[4] J. Provost et al., "Simultaneous positron emission tomography and ultrafast ultrasound for hybrid molecular, anatomical and functional imaging," *Nature Biomed. Eng.*, vol. 2, no. 2, pp. 85–94, Feb. 2018, doi: 10.1038/s41551-018-0188-z.

[5] A. B. E. Attia et al., "A review of clinical photoacoustic imaging: Current and future trends," *Photoacoustics*, vol. 16, Dec. 2019, Art. no. 100144, doi: 10.1016/j.pacs.2019.100144.

[6] J. Weber, P. C. Beard, and S. E. Bohndiek, "Contrast agents for molecular photoacoustic imaging," *Nat. Methods*, vol. 13, no. 8, pp. 639–650, Jul. 2016, doi: 10.1038/nmeth.3929.

[7] A. Needles et al., "Development and initial application of a fully integrated photoacoustic micro-ultrasound system," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 60, no. 5, pp. 888–897, May 2013, doi: 10.1109/tuffc.2013.2646.

[8] M. K. A. Singh, W. Steenbergen, and S. Manohar, "Handheld probe-based dual mode ultrasound/photoacoustics for biomedical imaging," in *Frontiers in Biophotonics for Translational Medicine: In the Celebration of Year of Light*, vol. 3, M. Olivo and U. S. Dinish, Eds. Singapore: Springer, 2016, pp. 209–247.

[9] X. L. Deán-Ben, E. Merčep, and D. Razansky, "Hybrid-array-based optoacoustic and ultrasound (OPUS) imaging of biological tissues," *Appl. Phys. Lett.*, vol. 110, no. 20, May 2017, Art. no. 203703, doi: 10.1063/1.4983462.

[10] V. Ntziachristos, "Going deeper than microscopy: The optical imaging frontier in biology," *Nature Methods*, vol. 7, no. 8, pp. 603–614, Aug. 2010, doi: 10.1038/nmeth.1483.

[11] S. Vilov, G. Godefroy, B. Arnal, and E. Bossy, "Photoacoustic fluctuation imaging: Theory and application to blood flow imaging," *Optica*, vol. 7, no. 11, p. 1495, Nov. 2020, doi: 10.1364/optica.400517.

[12] X. L. Deán-Ben and D. Razansky, "Adding fifth dimension to optoacoustic imaging:, Volumetric time-resolved spectrally enriched tomography," *Light Sci. Appl.*, vol. 3, pp. 1–5, May 2014, doi: 10.1038/lsa.2014.18.

[13] 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, doi: 10.1148/radiol.2019181325.

[14] X. L. Deán-Ben and D. Razansky, "Functional optoacoustic human angiography with handheld video rate three dimensional scanner," *Photoacoustics*, vol. 1, nos. 3–4, pp. 68–73, Dec. 2013, doi: 10.1016/j.pacs.2013.10.002.

[15] J. Robin, A. Ozbek, M. Reiss, X. L. Dean-Ben, and D. Razansky, "Dual-mode volumetric optoacoustic and contrast enhanced ultrasound imaging with spherical matrix arrays," *IEEE Trans. Med. Imag.*, vol. 41, no. 4, pp. 846–856, Apr. 2022, doi: 10.1109/TMI.2021.3125398.

[16] G. Godefroy, B. Arnal, and E. Bossy, "Full-visibility 3D imaging of oxygenation and blood flow by simultaneous multispectral photoacoustic fluctuation imaging (MS-PAFI) and ultrasound Doppler," *Sci. Rep.*, vol. 13, no. 1, p. 2961, Feb. 2023, doi: 10.1038/s41598-023-29177-9.

[17] S. Liu, W. Song, X. Liao, T. T.-H. Kim, and Y. Zheng, "Development of a handheld volumetric photoacoustic imaging system with a central-holed 2D matrix aperture," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 9, pp. 2482–2489, Sep. 2020, doi: 10.1109/TBME.2019.2963464.

[18] Y. Wang et al., "In vivo three-dimensional photoacoustic imaging based on a clinical matrix array ultrasound probe," *J. Biomed. Opt.*, vol. 17, no. 6, 2012, Art. no. 061208, doi: 10.1117/1.jbo.17.6.061208.

[19] J. Provost et al., "3D ultrafast ultrasound imaging in vivo," *Phys. Med. Biol.*, vol. 59, no. 19, pp. L1–L13, Oct. 2014, doi: 10.1088/0031-9155/59/19/L1.

[20] C. Lee, W. Choi, J. Kim, and C. Kim, "Three-dimensional clinical handheld photoacoustic/ultrasound scanner," *Photoacoustics*, vol. 18, Jun. 2020, Art. no. 100173, doi: 10.1016/j.pacs.2020.100173.

[21] Z. Pang, Y. Wang, W. Qin, W. Qi, and L. Xi, "Handheld volumetric photoacoustic/ultrasound imaging using an internal scanning mechanism," *Opt. Lett.*, vol. 47, no. 10, p. 2418, May 2022, doi: 10.1364/ol.458274.

[22] M. Yang et al., "Quantitative analysis of breast tumours aided by three-dimensional photoacoustic/ultrasound functional imaging," *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, May 2020, doi: 10.1038/s41598-020-64966-6.

[23] E. Zheng, H. Zhang, W. Hu, M. M. Doyley, and J. Xia, "Volumetric tri-modal imaging with combined photoacoustic, ultrasound, and shear wave elastography," *J. Appl. Phys.*, vol. 132, no. 3, Jul. 2022, Art. no. 034902, doi: 10.1063/5.0093619.

[24] C. Yoon, J. Kang, T.-K. Song, and J. H. Chang, "Elevational synthetic aperture focusing for three-dimensional photoacoustic imaging using a clinical one-dimensional array transducer," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 9, pp. 2817–2825, Sep. 2022, doi: 10.1109/TBME.2022.3154754.

[25] L. Vionnet, J. Gateau, M. Schwarz, A. Buehler, V. Ermolayev, and V. Ntziachristos, "24-MHz scanner for optoacoustic imaging of skin and burn," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 535–545, Feb. 2014, doi: 10.1109/tmi.2013.2289930.

[26] J. Gateau, M. Gesnik, J.-M. Chassot, and E. Bossy, "Single-side access, isotropic resolution, and multispectral three-dimensional photoacoustic imaging with rotate-translate scanning of ultrasonic detector array," *J. Biomed. Opt.*, vol. 20, no. 05, p. 1, May 2015, doi: 10.1117/1.JBO.20.5.056004.

[27] T. Lucas, I. Quidu, S. L. Bridal, and J. Gateau, "High-contrast and resolution 3-D ultrasonography with a clinical linear transducer array scanned in a rotate-translate geometry," *Appl. Sci.*, vol. 11, no. 2, p. 493, Jan. 2021, doi: 10.3390/app11020493.

[28] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 1, pp. 102–119, Jan. 2014, doi: 10.1109/tuffc.2014.6689779.

[29] J. Gateau, A. Chekkoury, and V. Ntziachristos, "High-resolution optoacoustic mesoscopy with a 24 MHz multidetector translate-rotate scanner," *J. Biomed. Opt.*, vol. 18, no. 10, Oct. 2013, Art. no. 106005, doi: 10.1117/1.JBO.18.10.106005.

[30] W. Marczak, "Water as a standard in the measurements of speed of sound in liquids," *J. Acoust. Soc. Amer.*, vol. 102, no. 5, pp. 2776–2779, Nov. 1997, doi: 10.1121/1.420332.

[31] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, Jul. 2012, doi: 10.1038/nmeth.2089.

[32] M. D. Gray and C. C. Coussios, "Compensation of array lens effects for improved co-registration of passive acoustic mapping and B-mode images for cavitation monitoring," *J. Acoust. Soc. Amer.*, vol. 146, no. 1, pp. EL78–EL84, Jul. 2019, doi: 10.1121/1.5118238.

[33] B. E. Treeby, T. K. Varslot, E. Z. Zhang, J. G. Laufer, and P. C. Beard, "Automatic sound speed selection in photoacoustic image reconstruction using an autofocus approach," *J. Biomed. Opt.*, vol. 16, no. 9, 2011, Art. no. 090501, doi: 10.1117/1.3619139.

[34] L. Mercier, T. Langø, F. Lindseth, and D. L. Collins, "A review of calibration techniques for freehand 3-D ultrasound systems," *Ultrasound Med. Biol.*, vol. 31, no. 4, pp. 449–471, Apr. 2005, doi: 10.1016/j.ultrasmedbio.2004.11.015.

[35] M. D. Verweij, B. E. Treeby, K. W. A. van Dongen, and L. Demi, "Simulation of ultrasound fields," in *Comprehensive Biomedical Physics*, vol. 2, Amsterdam, The Netherlands: Elsevier, 2014, pp. 465–500.

[36] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 56, no. 3, pp. 489–506, Mar. 2009, doi: 10.1109/TUFFC.2009.1067.

[37] L. Peralta, A. Gomez, Y. Luan, B.-H. Kim, J. V. Hajnal, and R. J. Eckersley, "Coherent multi-transducer ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 8, pp. 1316–1330, Aug. 2019, doi: 10.1109/TUFFC.2019.2921103.

[38] L. Peralta, D. Mazierli, A. Gomez, J. V. Hajnal, P. Tortoli, and A. Ramalli, "3-D coherent multitransducer ultrasound imaging with sparse spiral arrays," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 70, no. 3, pp. 197–206, Mar. 2023, doi: 10.1109/TUFFC.2023.3241774.

[39] H. de Hoop, N. J. Petterson, F. N. van de Vosse, M. R. H. M. van Sambeek, H.-M. Schwab, and R. G. P. Lopata, "Multiperspective ultrasound strain imaging of the abdominal aorta," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3714–3724, Nov. 2020, doi: 10.1109/TMI.2020.3003430.

[40] M. Sjoerdsma, S. C. F. P. M. Verstraeten, E. J. Maas, F. N. van de Vosse, M. R. H. M. van Sambeek, and R. G. P. Lopata, "Spatiotemporal registration of 3-D multi-perspective ultrasound images of abdominal aortic aneurysms," *Ultrasound Med. Biol.*, vol. 49, no. 1, pp. 318–332, Jan. 2023, doi: 10.1016/j.ultrasmedbio.2022.09.005.

[41] M. Xu and L. V. Wang, "Photoacoustic imaging in biomedicine," *Rev. Sci. Instrum.*, vol. 77, no. 4, pp. 1–22, 2006, doi: 10.1063/1.2195024.

[42] R. Ellwood, O. Ogunlade, E. Zhang, P. Beard, and B. Cox, "Photoacoustic tomography using orthogonal Fabry–Pérot sensors," *J. Biomed. Opt.*, vol. 22, no. 4, Dec. 2016, Art. no. 041009, doi: 10.1117/1.jbo.22.4.041009.

[43] M. Omar et al., "Optical imaging of post-embryonic zebrafish using multi orientation raster scan optoacoustic mesoscopy," *Light, Sci. Appl.*, vol. 6, no. 1, pp. 1–6, 2017, doi: 10.1038/lsa.2016.186.

[44] 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, pp. 1–8, 2017, doi: 10.1038/lsa.2016.247.

[45] H. Yang et al., "Soft ultrasound priors in optoacoustic reconstruction: Improving clinical vascular imaging," *Photoacoustics*, vol. 19, Sep. 2020, Art. no. 100172, doi: 10.1016/j.pacs.2020.100172.

[46] T. Han, M. Yang, F. Yang, L. Zhao, Y. Jiang, and C. Li, "A three-dimensional modeling method for quantitative photoacoustic breast imaging with handheld probe," *Photoacoustics*, vol. 21, Mar. 2021, Art. no. 100222, doi: 10.1016/j.pacs.2020.100222.

[47] J. Poudel, Y. Lou, and M. A. Anastasio, "A survey of computational frameworks for solving the acoustic inverse problem in three-dimensional photoacoustic computed tomography," *Phys. Med. Biol.*, vol. 64, no. 14, Jul. 2019, Art. no. 14TR01, doi: 10.1088/1361-6560/ab2017.

[48] R. Rau, D. Schweizer, V. Vishnevskiy, and O. Goksel, "Ultrasound aberration correction based on local speed-of-sound map estimation," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2019, pp. 2003–2006, doi: 10.1109/ULTSYM.2019.8926297.

[49] M. A. A. Caballero, J. Gateau, X.-L. Dean-Ben, and V. Ntziachristos, "Model-based optoacoustic image reconstruction of large three-dimensional tomographic datasets acquired with an array of directional detectors," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 433–443, Feb. 2014, doi: 10.1109/tmi.2013.2286546.

[50] V. Ntziachristos and D. Razansky, "Molecular imaging by means of multispectral optoacoustic tomography (MSOT)," *Chem. Rev.*, vol. 110, no. 5, pp. 2783–2794, May 2010, doi: 10.1021/Cr9002566.

[51] J. Gateau, A. Chekkoury, and V. Ntziachristos, "Ultra-wideband three-dimensional optoacoustic tomography," *Opt. Lett.*, vol. 38, no. 22, pp. 4671–4674, Nov. 2013, doi: 10.1364/OL.38.004671.

[52] W. R. Thompson et al., "Characterizing a photoacoustic and fluorescence imaging platform for preclinical murine longitudinal studies," *J. Biomed. Opt.*, vol. 28, no. 3, pp. 1–20, Mar. 2023, doi: 10.1117/1.JBO.28.3.036001.

[53] J. Jo et al., "Photoacoustic tomography for human musculoskeletal imaging and inflammatory arthritis detection," *Photoacoustics*, vol. 12, pp. 82–89, Dec. 2018, doi: 10.1016/j.pacs.2018.07.004.

[54] J. Kim et al., "Multiparametric photoacoustic analysis of human thyroid cancers in vivo," *Cancer Res.*, vol. 81, no. 18, pp. 4849–4860, Sep. 2021, doi: 10.1158/0008-5472.CAN-20-3334.

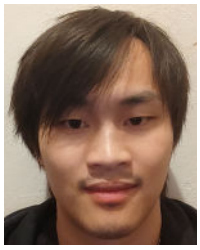
Clément Linger (Member, IEEE) was born in Suresnes, France, in 1998. He received the Engineer degree from the Ecole Nationale Supérieure de Chimie de Paris (ENSCP), Paris Sciences et Lettres University, Paris, France, and the master's degree from the University of Paris Sciences et Lettres in 2021. He is currently pursuing the Ph.D. degree with the Laboratory of Biomedical Imaging, Sorbonne University, Paris, and the Institut Galien Paris-Saclay, Paris-Saclay University, Orsay, France, working on photoacoustic imaging and labeled drug nanovectors formulation.

Yoann Atlas, photograph and biography not available at the time of publication.

Rémy Winter was born in Metz, France, in 1998. He received the Mechanical Engineer degree from the Ecole Nationale d'Ingénieurs de Metz, University of Lorraine, Metz, and the master's degree in biomechanics from the University of Lorraine, in 2021. He is currently pursuing the Ph.D. degree with the Côte d'Azur Clinical Research Unit (UR2CA), Université Côte d'Azur, Nice, France, working on the development of a 3-D wrist model for surgery planning.

Marine Vandebrouck, photograph and biography not available at the time of publication.

Maxime Faure was born in Meulan-en-Yvelines, France, in 1997. He received the Engineer degree from the Université Paris-Est Créteil (UPEC), Créteil, France, and the École Supérieure d'Ingénieurs en Électrotechnique et Électronique (ESIEE Paris), Noisy-le-Grand, France, in 2020, and the master's degree in image processing from Sorbonne University, Paris, France, in 2021. He is currently pursuing the Ph.D. degree in image-guided cataract surgery with the Laboratoire de Traitement de l'Information Médicale (LaTIM), Brest, France.



Théotim Lucas was born in Hô-Chi-Minh-Ville, Vietnam, in 1995. He received the Engineer degree from Polytech Lyon, Villeurbanne, France, the master's degree in biomedical imaging from Université Claude Bernard Lyon 1, Villeurbanne, in 2019, and the Ph.D. degree from the Laboratory of Biomedical Imaging and the Laboratoire Matière Système Complexe (MSC), Paris, France, in 2023, working on the photoacoustic characterization of agents for photothermal therapy.

S. Lori Bridal (Senior Member, IEEE) received the B.S. degree in physics from Oklahoma State University, Stillwater, OK, USA, in 1988, as a National Merit Scholar, and the M.S. and Ph.D. degrees from Washington University in Saint Louis, St. Louis, MO, USA, in 1990 and 1994, respectively.

She is currently a Research Director with the French National Research Center (CNRS) and the Director of the Laboratory of Biomedical Imaging (Sorbonne University, Inserm, and CNRS), Paris, France, developing biomedical imaging research to address the major public health of age-related bone modifications, cancer, cardiovascular, and neurodegenerative disease. She began research at the Parametric Imaging Laboratory, Paris, during her post-doctoral research supported by the French Chateaubriand Fellowship and became a CNRS Researcher in 1997.

Dr. Bridal received the Bronze Medal from CNRS in 2004, a national recognition of research significance. In 2021, she received the IEEE Ultrasonics, Ferroelectrics and Frequency Control Society (UFFC-S) Award for Distinguished Service and is currently with VP Publications within the society.

Jérôme Gateau was born in Cholet, France, in 1983. He received the Engineering degree from the Ecole Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI), Paris, France, in 2007, and the M.S. and Ph.D. degrees in physics (acoustics) from University Paris Diderot, Paris, for his work on ultrafast ultrasound imaging of cavitation events for applications to ultrasonic therapy and detection mapping, in 2007 and 2011, respectively.

He worked as a Postdoctoral Researcher in photoacoustic imaging with the Institute of Biological and Medical Imaging of Helmholtz Munich, Oberschleißheim, Germany, and the Institut Langevin, Paris. He is currently a Researcher with the French National Research Center (CNRS), Laboratory of Biomedical Imaging, Paris. His current research interests include photoacoustic and ultrasound tomography for biomedical applications and the quantitative characterization of therapeutic nanoagents with photoacoustic or photothermal properties.