



HAL
open science

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

Clément Linger, Yoann Atlas, Remy Winter, Marine Vandebrouck, Maxime
Faure, Théotim Lucas, S. Lori Bridal, Jérôme Gâteau

► **To cite this version:**

Clément Linger, Yoann Atlas, Remy Winter, Marine Vandebrouck, Maxime Faure, et al.. Volumetric and Simultaneous Photoacoustic and Ultrasound Imaging with a Conventional Linear Array in a Multiview Scanning Scheme. 2022. hal-03904339v1

HAL Id: hal-03904339

<https://hal.science/hal-03904339v1>

Preprint submitted on 16 Dec 2022 (v1), last revised 24 Apr 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Clément Linger, Yoann Atlas, Remy Winter, Marine Vandebrouck, Maxime Faure, Théotim Lucas, S. Lori Bridal, and Jérôme Gateau

Abstract— Volumetric dual photoacoustic (PA) / ultrasonic (US) imaging with precise spatial and temporal coregistration can provide valuable and complementary information for diagnosis and monitoring. Considerable research has sought to combine 3D PA/US imaging in configurations that can be transferred to clinical application but technical compromises currently result in poor image quality either for photoacoustic or ultrasonic modes. Simultaneous 3D PA/US tomography was implemented here by interlacing PA and US acquisitions during the rotate-translate scan of a 5-MHz linear array (12 angles and 30 mm translational range to image a cylindrical volume of 21 mm diameter and 19 mm length within 21 seconds). Volumetric image reconstruction was performed with synthetic aperture approaches. An original calibration method was developed to estimate 6 geometrical parameters and 1 temporal off-set providing sharpest and best superimposed reconstructions. Calibration thread phantom design and choice of metrics to build the cost function were based on analysis of a numerical phantom and the final selection demonstrates a high estimation accuracy of the 7 parameters. Experimental estimations validated the calibration repeatability. Experiments in an additional phantom showed a superposition distance between thread centers identified in the PA and US images to be smaller than 10% of the acoustic wavelength, and a spatial resolution on the order of the wavelength. Dual mode 3D imaging with high-quality co-registration and excellent, uniform spatial resolution was further demonstrated on phantoms with complementary contrasts, and should contribute to more sensitive and robust imaging to detect and follow biological changes or the accumulation of nanoagents in living systems.

Index Terms—Tomography, rotate-translate scan, volumetric imaging, simultaneous dual imaging, photoacoustic, ultrafast ultrasound imaging

I. INTRODUCTION

VOLUMETRIC and simultaneously co-registered multimodal imaging is increasingly developing in biomedical imaging due to the rich, multiplexed and complementary anatomical–functional information that can be precisely spatially and temporally correlated [1]. The combination of positron emission tomography (PET) and computed tomography (CT), introduced in the 1990s, was among the first volumetric and simultaneous bimodal imaging system to be clinically available [2]. In this dual-modality, CT provides the anatomical context needed to interpret the functional PET. Various combinations

of volumetric imaging have since been developed. For example, PET has been integrated with magnetic resonance imaging (MRI) [1], [2], photoacoustic imaging (PAI) and optical coherence tomography (OCT) have been superimposed [3] and PET has been combined with Doppler ultrasound imaging [4].

The advantages of simultaneously acquired, coregistered multimodal images are manifold. The information gained within one single imaging session is maximized which reduces the scan time and increases benefits for patients, researchers and clinical-management teams. Typically, complementary spatio-anatomical and functional information can be obtained through the combination of modalities for a more comprehensive characterization of the region under examination. Precisely correlated (both spatially and temporally) information is essential for a precise characterization. In addition, volumetric imaging provides a detailed view of regions under examination from various orientations for improved diagnosis, and facilitates the investigation and comparison for longitudinal studies or to monitor the progression or regression of pathology in response to therapy.

Ultrasound imaging (USI) and PAI are complementary imaging modalities since USI provides the anatomical correlation for the molecular information supplied by PAI. Indeed, USI is sensitive to differences in the mechanical microstructures of tissues and can be used to delineate organs and lesions. USI is used in clinical routine to obtain anatomical information. PAI is sensitive to optical absorption and can provide information on hemoglobin oxygenation and detect molecular and nanoparticulate contrast agents [5], [6]. Both USI and PAI are based on the detection of ultrasonic signals and can therefore be implemented with the same ultrasound detector for simultaneous co-registration. Recorded ultrasonic signals are generated in situ by the optical absorption of a laser excitation for PAI, while, for USI, they are created with the incident ultrasonic pulse and backscattered from structures with different acoustic impedance. In USI the transducer generally operates in both emission and reception modes. For 2D imaging, simultaneously co-registration of USI and PAI has already been demonstrated with a single ultrasound (US) detector array [7], [8] and with a dual array configuration [9]. However, PAI is intrinsically 3D in nature due to scattering of the excitation light by the biological tissue, which leads to out-of-plane artifacts in 2D imaging. Moreover, elongated

structures like blood vessels have a strong directionality in PAI and may not be visible in the limited view configuration of 2D PAI, when they emit outside of the limited angular aperture.

Spherical US detector matrices (2D arrays) with a large angular aperture have been developed specifically for 3D PAI, using transducers that could also emit ultrasonic waves [10], [11]. Given the limited number of elements used to cover the spherical surface, spherical US matrices can be considered sparse for ultrasonic waves. A large, sparse angular aperture can provide high quality photoacoustic (PA) images. However, US images obtained with such PAI spherical arrays present strong artifacts, that have been shown to be reduced using processing to extract the Doppler signal [12], [13] or the signal from sparse US contrast agents [12]. Indeed, pulse-echo US images are densely filled with echogenic structures and high-quality, 3D USI requires US matrices with a high spatial sampling and a large number of elements. Alternatively, planar US transducer matrices that were developed for 3D USI have been tested for 3D PAI [14], [15]. However, the limited angular aperture of such US transducer matrices and the poor sensitivity of the small elements lead to limited view artifacts, limited spatial resolution and poor sensitivity for 3D PAI. In short, simultaneous co-registration of volumetric PAI and USI with US matrices currently results in poor image quality for one of the two modalities.

A possible alternative to US matrices lies in the mechanical scanning of US linear arrays that have been designed for 2D imaging. Several systems based on a translational scan of a linear array have been developed recently and their ability to provide simultaneously co-registered 3D USI / PAI has been demonstrated [16]–[19]. In this configuration, linear arrays operating in simultaneous 2D USI / PAI are scanned in the elevational direction, perpendicularly to the imaging plane. The 2D images are then stacked to obtain a volume. This approach is easy to implement and to transfer to a clinical environment. However, the angular aperture in the elevational direction of a linear US array is very limited, which induces strong limited view artifacts in PAI. This issue is not addressed with the translational scan, even when synthetic aperture focusing is used [20], [21]. Moreover, the spatial resolution in the translational direction is strongly degraded compared to the in-plane resolution both in PAI and USI. 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 compared to a translational scan [22], [23]. The rotate-translate approach has been demonstrated independently on two different imaging systems for 3D PAI [22] and for 3D USI [23], with a significant decrease of the scan duration between the two demonstrations. In this paper, we demonstrate simultaneous co-registration of 3D PAI and 3D USI with a single rotate-translate scan of a linear US array. After a calibration step that enables us to finely determine the position of the array in the scanning system, we validate the co-registration performance *in vitro* with an imaging phantom presenting both US and PA contrast. We also illustrate the capabilities of the system and the complementarity of volumetric USI and PAI by imaging phantoms that have mixed structures with either US or a PA contrast.

II. MATERIALS AND METHODS

A. Experimental set-up

The experimental setup is presented in Fig. 1. It can be divided in four main parts: (1) the optical excitation comprising a nanosecond laser and an optical fiber bundle, (2) the US acquisition system consisting of an US linear array driven by a programmable US platform, (3) the scanning system comprising two motorized stages and their motion controller, and (4) the synchronization system piloted with a programmable trigger generator. The acquisition process was fully automated.

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 was used to generate the optical excitation at 700 nm. A bifurcated fiber bundle (CeramOptec GmbH, Bonn, Germany) guided the light toward the imaged volume to obtain a 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 in the laser.

The US transducer array was a 128-element linear array (L7–4, ATL) driven by a programmable, 64-channel 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. Between two laser pulses, transmit events produced “plane wave” ultrasound emissions (a beam that is unfocused in the lateral direction of the array) [24]. 6.25 ms after each laser pulse, five tilted, plane waves were emitted with steered angles of -4° , -2° , 0° (the 64 elements in the transducer array were fired at the same time), 2° and 4° at a PRF of 160 Hz. Transmitted US pulses were 1 cycle long at 5.2 MHz. Pulse-echo US signals and PA signals were recorded at a sampling frequency of 20 MS/s and 62.5 MS/s, respectively. A constant gain was adjusted for the PA and US signals to ensure a good amplitude digitalization of received signals without saturation (no time gain compensation was applied).

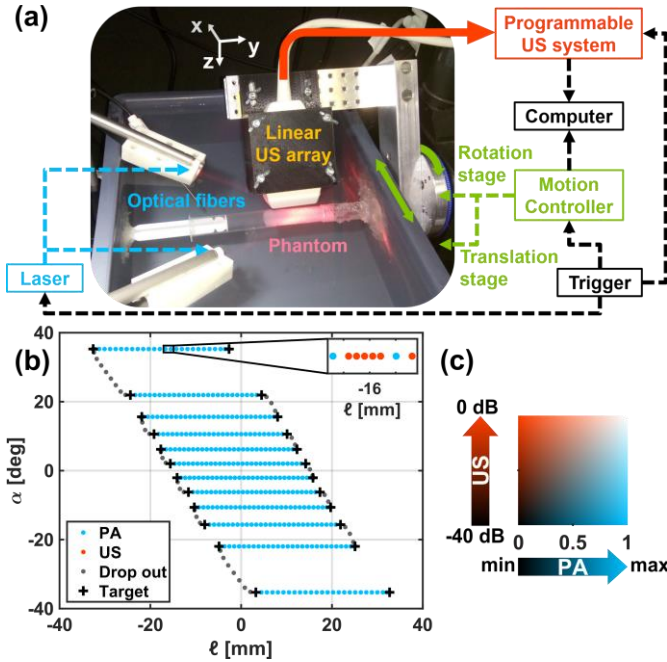


Fig. 1. (a) Annotated picture of the experimental setup. (b) Motor positions for the PA acquisitions over one scan. For a better readability, positions corresponding to the US and PA acquisitions are indicated only for the inset in the upper right corner of the graph. Positions shown in grey (drop out) were not used for the reconstruction, but are 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 from -40 to 0 dB.

The US array was mounted on a rotate-translate scanning system described in detail in ref [23]. In brief, a rotation stage was mounted on a translation stage (Physik instrumente, Karlsruhe, Germany). 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 period of $\Delta\alpha = 4^\circ$. For each angle, the imaged volume was scanned with a translation range $L = 30$ mm and a linear sampling period $\Delta\ell = 1$ mm. Relative to parameters used in ref [23], the scan parameters, in particular the translation velocity, had to be adjusted to adapt for the PRF of the Laser. The motion was continuous and, for each laser emission and each US plane wave emission, the motor positions were recorded and stored in the motor controller. The scan was automated and initiated with an external trigger. The motor positions are presented in Fig. 1(b).

To synchronize the instruments and to coordinate the acquisition sequence, a trigger generator (BNC Model 577, Berkeley Nucleonics, San Rafael, CA, USA) was used to send external triggers simultaneously to the Laser, the programmable US system and the motion controller.

The imaged sample, the array elements and the outputs of the fiber bundle were immersed in a water bath filled with tap water. The speed of sound in the bath was derived from the water temperature [25], measured with a thermometer (HI98509, Hanna instruments, Lingolsheim, France).

For a volumetric acquisition, a total of 411 laser pulses were fired for a total duration of 21 s. Video 1 presents a video of the acquisition and illustrates the rotate-translate scan of the array.

At the end of the acquisition, the radiofrequency data of the PA and US events, the motor positions and the pyrometer values were transferred to a computer for signal processing and image reconstruction.

B. Image reconstruction

1) Image Reconstruction Algorithm

The 3D 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 so that x_O and z_O are the coordinates of the rotation axis, when the translation stage is at its center position ($\ell=0$), and y_O is the coordinate of the center of the array. In this grid, the voxel dimensions were chosen equal to $p_x \times p_y \times p_z = 71 \mu\text{m} \times 143 \mu\text{m} \times 71 \mu\text{m}$, with an anisotropy reflecting the best-expected resolutions. The image volume is defined by a diamond-shaped, cross-sectional area (DSCA, Fig. 2) 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. PA signals were bandpass filtered between 2 MHz and 10 MHz (Butterworth, order 3).

Image reconstruction was performed with simple delay-and-sum beamforming algorithms. The one-way (PA) and two-way (US) travel times between the US transducer element positions (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. Then, the value of each voxel was computed by summing the signal values assessed at the voxel-associated travel time, over all the elements of the array and all the different tomographic positions. The US image reconstruction algorithm was detailed in ref [23]. The same apodizations were used for PA and US image reconstruction.

Three-dimensional, envelope-detected images were obtained. The 3-D images were displayed using maximum amplitude projection (MAP) images along the axes of the coordinate system. Rotating MAP images around the z axis were obtained with the 3D project option of ImageJ [26]. The colorscale used for the images is displayed in Fig. 1(c).

2) Time delays for the reconstruction

The recording of the ultrasound signals was set to start at the same time as the Laser emission for the PA acquisition or at the same time as the ultrasound emission for the US acquisition. However, due to the lens effects [27] and additional time delays induced by the acquisition hardware, we found that two time parameters $t_{0\text{PA}}$ and $t_{0\text{US}}$ needed to be determined to convert the voxel-associated travel times into the time indexes of the recorded signals. $t_{0\text{PA}}$ and $t_{0\text{US}}$ correspond to effective offsets of the recorded signals.

$t_{0\text{PA}}$ 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

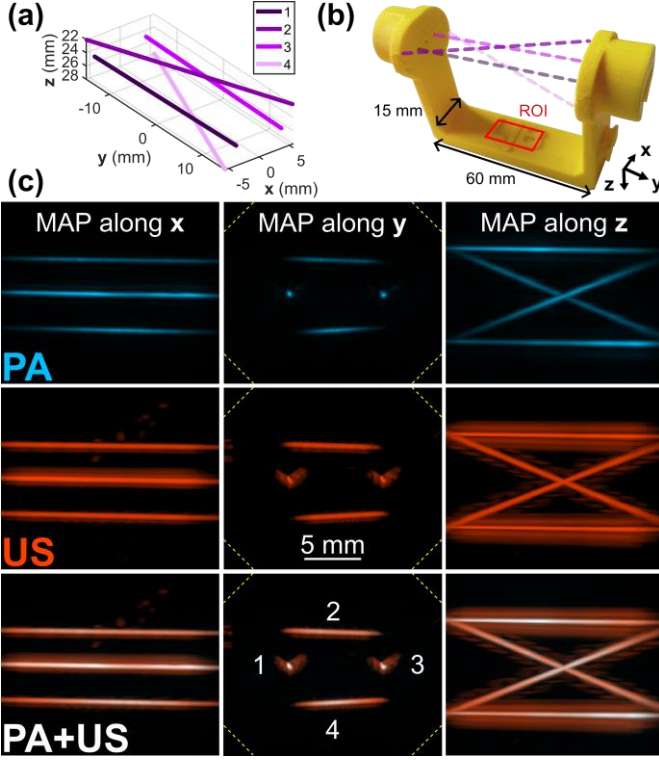


Fig. 2. (a) Imaged region (also called region of interest (ROI)) of the calibration phantom in a 3D coordinate system with dimensions in mm. (b) Picture of the calibration phantom. Four 20- μm nylon threads are mounted on a yellow 3D-printed frame. To ease the readability, the threads have been highlighted 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: photoacoustic image; second row: ultrasound image; Third row: Combined PA/US image. The colorscale is presented in Fig. 1 (c). Each image is a maximum amplitude projection (MAP) image. The visible part of the diamond-shaped, cross-sectional area (DSCA) is shown with yellow dashed lines in the MAP images along y .

with regards to the wavelength at the center frequency of the array ($\Lambda_c \approx 300 \mu\text{m}$), the threads can be assumed to be point absorbers in 2D. 2D images ($p_y \times p_z = 71 \mu\text{m} \times 71 \mu\text{m}$) were beamformed for different values of $t_{0 \text{ PA}}$, with the speed of sound determined with the temperature of the water bath. The one-dimension Brenner's gradient of the images was computed along the lateral dimension of the array:

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

where $f(y, z)$ represents the gray level intensity in the 2D image. The Brenner's gradient provides a quantitative measure of image sharpness and was shown to be an efficient metric for the speed of sound calibration in PA [28]. It was found to be maximized for $t_{0 \text{ PA}} = -1.3 \mu\text{s}$. In the recorded signals, $|t_{0 \text{ PA}}|$ corresponded also 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 gives the motor positions in their own coordinate systems (integrated position sensors), while the reconstruction algorithm requires the position and orientation of the array in the fixed coordinate system ($O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$). Therefore, the spatial transformation matrix from the mobile Cartesian coordinate system ($O_a, \mathbf{u}, \mathbf{v}, \mathbf{w}$) attached to the transducer array to ($O, \mathbf{e}_x,$

$\mathbf{e}_y, \mathbf{e}_z$) 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 origin O_a is the center of the active transducer aperture. O_a is located on the interface between the transducer array and the water.

A 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 equal zero ($\alpha=0$) when the radial axis is vertical. The position of the center of the element number n of the array ($n \in \llbracket 1, N \rrbracket$ with $N=64$) in ($O, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$) can be decomposed as:

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

With p the interelement spacing of the array. Here, $p = 298 \mu\text{m}$. Moreover, the transformation matrices are:

$$\begin{pmatrix} u_x(\alpha) & v_x(\alpha) & w_x(\alpha) \\ u_y(\alpha) & v_y(\alpha) & w_y(\alpha) \\ u_z(\alpha) & v_z(\alpha) & w_z(\alpha) \end{pmatrix} = \mathcal{R}_\alpha \cdot \mathcal{R}_{\text{Pitch}} \cdot \mathcal{R}_{\text{Yaw}} \cdot \mathcal{R}_{\text{Roll}} \quad (3)$$

$$\begin{pmatrix} x_{O_a}(\ell, \alpha) \\ y_{O_a}(\ell, \alpha) \\ z_{O_a}(\ell, \alpha) \end{pmatrix} = \mathcal{R}_\alpha \cdot \begin{pmatrix} \Delta x \\ 0 \\ \Delta z \end{pmatrix} + \ell \cdot \begin{pmatrix} \cos(\theta) \cdot \cos(\varphi) \\ \sin(\theta) \\ \cos(\theta) \cdot \sin(\varphi) \end{pmatrix} \quad (4)$$

With \mathcal{R}_α the rotation matrix around the y -axis due to the motion of the rotation motor. $\mathcal{R}_{\text{Roll}}$, $\mathcal{R}_{\text{Pitch}}$ and \mathcal{R}_{Yaw} are rotation matrices around the 1st-axis, the 2nd-axis and the 3rd-axis, respectively. Δx and Δz are the coordinates of O_a when $\alpha=0$ and $\ell=0$. The product of these rotation matrices and the offsets models the misalignment of the array axis \mathbf{v} compared to the rotation axis of the stage \mathbf{e}_y . The angles θ and φ account for misalignment of the translation axis with \mathbf{e}_x . Therefore, a total of 7 geometrical parameters independent of (ℓ, α) need to be determined: Roll, Pitch, Yaw, Δx , Δz , θ , φ .

C. Calibration method

The inadequate estimation of the reconstruction parameters results in degraded image quality in terms of sharpness and a misalignment of the PA and US images. Therefore, we developed a calibration method for the determination of the required parameters. The calibration method needs to be applied once and the parameters can be used to reconstruct subsequent acquisitions as long as the setup is not mechanically misaligned.

The calibration method combines a calibration phantom and an optimization algorithm to estimate the 7 parameters: $t_{0 \text{ US}}$, Roll, Pitch, Yaw, Δx , Δz , θ . The parameter φ was fixed equal to zero because the perpendicularity between the translation axis and \mathbf{e}_z was precisely ensured by the mechanical design, and because of equivalent solutions for our optimization algorithm and for different sets of (Pitch, φ) values.

1) Calibration Phantom

The required specifications for the calibration phantom were that it be simple to build, easy to use and does not require an absolute and tedious positioning procedure. We developed a wire phantom, inspired by phantoms used for the calibration of

freehand 3D ultrasound systems such as Z-fiducial phantoms [29]. Wires or threads have several advantages. First, black threads have a dual contrast PA and US compared to water, and are therefore expected to be 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, the orientation can be varied. Finally, with a few non-crossing and well-separated threads, the segmentation of the 3D image allows local assessments of the image quality, in particular, the image sharpness in slices intersecting the threads.

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 $\pm 22^\circ$ placed in parallel planes below and above, respectively (Thread 2 and 4). The angle was chosen to provide a 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 points 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, Vetsuture, France) mounted on a 3D-printed frame (Fig. 2 (b)). A numerical phantom was also designed in order to validate the calibration method with a known set of parameters. For this numerical phantom, each thread was discretized into point sources (PA) or point scatterers (US) spaced by $\Lambda_c/5$ where Λ_c is the ultrasound wavelength at the center frequency of the transducer ($\Lambda_c \approx 300\mu\text{m}$). The numerical simulation is based on the sum over all the points of time signals delayed by the travel times of the ultrasound waves. A one cycle sinusoidal signal at 5 MHz with a gaussian envelop was used for the US simulation, and the derivative of a gaussian pulse (standard deviation 17 ns) for the PA simulation. The finite size (width and height) and cylindrical focusing of the array elements were modeled with a cylindrical transducer discretized with point transducers spaced by $\Lambda_c/5$. Additionally, the 7 geometrical parameters and the two time-offsets were considered in the simulation.

2) Calibration algorithm

The calibration algorithm is based on an optimization algorithm that minimizes a cost function.

a) Cost function

For one set of the 7 parameters and one volumetric acquisition comprising PA and US (5 steered angles) data, the cost function was assessed using the images of 5 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 $p_y \times p_z = 71 \mu\text{m} \times 71 \mu\text{m}$ using the 3D reconstruction algorithm. 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 is expected to correspond to one thread. The cost was set to zero if less than four connected components were counted. For each region, the centroid position (x_j, z_j) with weights based on the grayscale image intensity value was computed. The distance between the centroids determined in the US image and in the PA image in slice i was:

$$d_{\text{US-PA}}(i, j) = \sqrt{(x_j^{\text{US}} - x_j^{\text{PA}})^2 + (z_j^{\text{US}} - z_j^{\text{PA}})^2} \quad (4)$$

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

For Threads 2 and 4, several quantities were computed. 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 over the two threads of the coefficient of determination gave R^2_{US} and R^2_{PA} for the US images and the PA images, respectively. R^2_{US} and R^2_{PA} equals 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 $d_{\text{PA-US}}$, named $D_{\text{PA-US}}$, was computed over the two threads and the five slices. $D_{\text{PA-US}}$ evaluated the superposition of the PA and US images. $D_{\text{PA-US}}$ is expected to be equal to zero for the correct set of parameters. $D_{\text{PA-US}}$ was found to be sensitive to errors in the parameter $t_{0\text{US}}$. $D_{\text{PA-US}}$ is expressed in mm and the cost was set equal to zero for $D_{\text{PA-US}} > 1$ mm. 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 [28]. The mean of the local normalized variance over the two threads and over the five slices was named N_{VUS} . N_{VUS} is expected to be maximal for the correct set of parameters. N_{VUS} was found to be sensitive to errors in the parameters: Pitch, Yaw, Δx , Δz , θ .

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

$$S_{N2D} = \frac{1}{\sum_{x,z} f(x,z)} \left(\sum_{x,z} (f(x+1, z) - f(x, z))^2 + \sum_{x,z} (f(x, z+1) - f(x, z))^2 \right) \quad (5)$$

where $f(x, z)$ represents the gray level intensity in the 2D ultrasound image. The mean of the normalized squared gradient over the five slices was named S_{NUS} . S_{NUS} was found to be sensitive to errors in the parameters: Roll, Pitch, Yaw, Δx , Δz .

The metrics R^2_{US} , R^2_{PA} , $D_{\text{PA-US}}$, N_{VUS} and S_{NUS} were selected based on numerical simulations and observations of the variations in the metric values induced by each parameter individually. Among the diversity of tested metrics, we kept the most sensitive and we made sure that all the parameters were covered by at least one metric. We found 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 the metrics with a product and the cost was then defined by:

$$Cost = \begin{cases} 0 & \text{for } D_{PA-US} \geq 1 \text{ mm} \\ -R_{US}^2 \cdot R_{PA}^2 \cdot (1 - D_{PA-US})^2 \cdot N_{VUS} \cdot S_{NUS} & \end{cases} \quad (6)$$

b) Optimization algorithm

The calibration algorithm relies on the cost function. An initial combination of 7 parameters was given as an 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 non-zero cost. The combination with the smallest cost was used for the second step: the application of a Particle Swarm optimization algorithm.

With the numerical calibration phantom, we found a variability of the output combination after the calibration algorithm and of the associated cost. This indicates local minima of the cost function. To mitigate this variability, we chose to run the optimization algorithm 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 5 lowest costs.

3) 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_{0US} , 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 pitch of the array is equal to Λ_c , the acceptability range for angles was set to $\sin^{-1} \frac{1}{64} = 15.6 \text{ mrad}$.

b) Variability quantifications

To assess the variability induced by the optimization algorithm, the numerical calibration phantom was used. The simulated combination reflected the experimental data (mean over 10 experiments). Absolute difference between obtained and expected parameters were computed and divided by the acceptability threshold to be expressed as a percentage. This metric evaluated the accuracy of the calibration method and therefore is named A_c .

To assess inter-acquisition variability on experimental data, parameters were obtained for 10 acquisitions. These parameters were compared to the mean parameter over the 10 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 R_p .

4) 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 USI and in PAI. As for the image processing used in the calculation of D_{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 5 slices located at $y = i \times 3 \text{ mm}$ with $i \in \llbracket -2; 2 \rrbracket$. Each slice was reconstructed with pixel sizes $p_y \times p_z = 71 \mu m \times 71 \mu m$. The mean distance of d_{US-PA} (see equation (4)) 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 2D-Gaussian model of equation:

$$f(x, z) = A \exp\left(-\left(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(z-z_0)^2}{2\sigma_z^2}\right)\right) \quad (7)$$

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 $FWHM_i = 2\sqrt{2 \ln 2} \times \sigma_i$. As nylon threads can be considered small compared to Λ_c , the FWHM is an estimate of the width of the Line Spread Function (LSF).

D. Imaging phantoms

The first phantom (Ph1) was a wire phantom designed to evaluate the co-registration capabilities using a phantom with a geometry different than the calibration phantom and in addition with a different material. Three 20- μm diameter black nylon threads (NYL02DS, Vetsuture, France) and three black polyester threads (Coat Epic 150) were mounted on a 3D-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 prepared with agar powder 2% w/v (A1296, Sigma Aldrich, St. Louis, MO, USA) and cellulose powder 1% w/v (Sigmacell cellulose Type 20, Sigma Aldrich) in water. Cellulose particles (20 μm) act as ultrasound scatterers to mimic the scattering properties of biological tissues for US imaging. The gel was molded in a cylindrical mold (diameter of 20 mm) with three cylindrical solid inclusions of 5 mm in diameter and of the same length as the mold. 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. 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 crossed 20- μm diameter black nylon threads were embedded in

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.

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 the 7 parameters are fully within the acceptability ranges, for both Ac and Rp.

For the accuracy, the calibration outputs were compared to the ground truth thanks to a numerical simulation. The mean Ac over the 7 parameters was found equal to 26% and Ac had a maximum of 55% for the Roll parameter. We can then consider that the developed calibration method enables to accurately determine the set of parameters.

To evaluate the repeatability, ten acquisitions performed on various days (distributed in three imaging sessions over one week) were used. 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 7 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 parameters, 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’ refer to nylon and polyester threads, respectively. For instance, the thread 1n is symmetrical to 1p, with respect to the

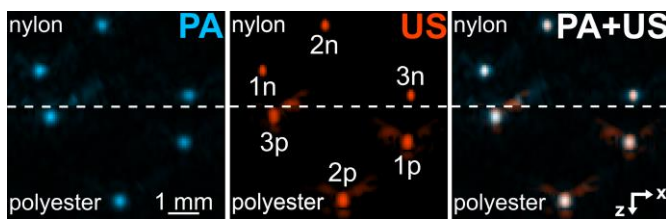


Fig. 3. Image reconstruction of the central plane of Ph1 ($y=0$). The horizontal white dashed line represents the separation between nylon and polyester zones. Threads with the same number are symmetrical and n stands for nylon and p for polyester. PA image (left) is presented on a linear scale while the US image (center) is presented with a threshold at -30dB.

TABLE II
RESOLUTION AND SUPERPOSITION ASSESSMENT

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

center of the image. The nylon threads were grouped on the top part of the phantom while the polyester threads are grouped on 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 times 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. However, the superposition of the images can be observed for the two materials and for all the threads regardless of their orientation or position in space (Fig. 3 and Video2). The blue lateral halo around the white spots in the combined image (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 thread with different orientations than in the calibration phantom, and it validates that the calibration ensures the superpositions for objects with a dual contrast but in another material. The standard deviation computed over the three threads and five imaging planes is small as well 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 provides 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 approach. For comparison, the FWHM_x was on the order of 1 mm [23] 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 USI [23] and PAI [22] is then confirmed for the simultaneously co-registered imaging.

In each direction, the US resolution is slightly better than the PA resolution (70-100% of Λ_c for US vs 130% of Λ_c for PA).

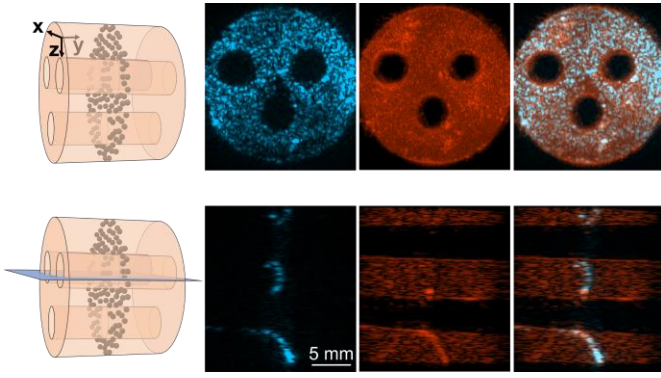


Fig. 4. Image reconstruction of Ph2. Schematic drawing of Ph2 are displayed in the left column. The first row displays MAP images along y while the second row shows a slice perpendicular to the z -axis. PA images are presented in linear color scale. However, because a few bright spots (probably clusters of microspheres) were dominating the color scale and hiding the rest, voxel values were saturated at 40% of their maximum before normalization. US images were thresholded at -40 dB.

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 USI, 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 USI compared to PAI. This sampling factor has been shown to influence the resolution along x [23]. 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 the dual modality imaging, phantoms with complementary contrasts were designed and produced. The images of Ph2 and Ph3 are displayed in Fig. 4 and 5, respectively.

For the ultrasound contrast, Ph2 is a homogeneously

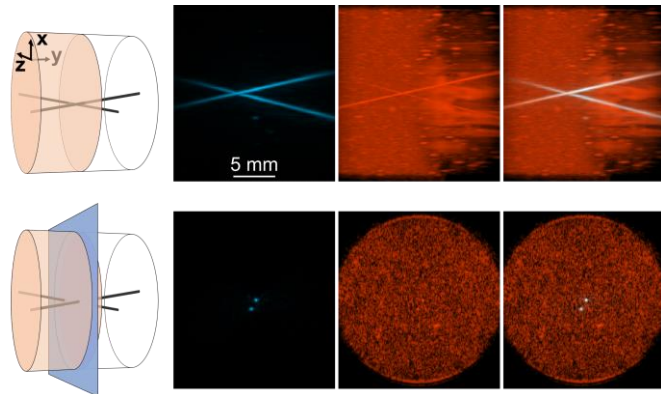


Fig. 5. Image reconstruction of Ph3. Schematic drawing of Ph3 are displayed in the left column. The first row displays MAP images along z while the second row shows a slice perpendicular to the y -axis. US images were thresholded at -40 dB. Along y -axis, left part is agar with cellulose and right part is agar alone.

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 (Fig. 4). The optical absorption of the agar gel and the water are negligible. In the first row of the Fig. 4, a MAP image along y is shown for PAI, USI, 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 PAI and USI. 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 USI image along the y -axis can be observed. As expected, the three holes appear with a negative contrast for both PAI and USI independently of their spatial position. The holes allow to further validate the superimposition of the two modalities. They are concentric in USI and in PAI. We can notice that the outlines of the holes are blurrier and the diameter of holes seems smaller on the US images. This effect can be attributed to the stronger side lobes in the US images induced 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). The microspheres were hardly visible in the agar matrix in USI, or only when numerous spheres are gathered together, while they appear with a 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 is homogeneously echoic for instance. The US images 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 ultrasound 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. However, the threads 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 co-registered PAI and USI. 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 for the subsequent acquisitions of images without the need of fiducial markers on every 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. 7 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 motion of the array occurs. For this study, no significant variation could be noticed over a period of one week. The calibration phantom was easy to build and was comprised of four well-separated threads for a facilitated identification and measurements of local properties. Our cost function used only five imaging planes to avoid the reconstruction of the entire volume and the associated long computation time. Our selection of metrics with the sharpest variations with regards to the parameters implied that the cost was dominated by the properties of the US image, but with an influence of 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 to determine a solution close to the ground truth than a downhill simplex method. The refinement of the optimization algorithm was beyond the scope of this study as we focused on the effectivity of the method, but it will be considered in a future study.

This calibration procedure was found 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 PA imaging and in US imaging, independently, leading to various calibration procedures. For US imaging, coherent compounding of images acquired with two US transducer arrays have been developed in 2D [30] and in 3D [31]. Dedicated calibration phantoms were also used that consisted in isotropic scatterers (5 wires in 2D and 3 spheres in 3D). The phantoms were combined with a cost function linked to the coherence of the echoes from the isotropic scatterers on the different elements of the array to determine geometrical parameters (4 parameters in 2D and 6 in 3D). 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 long computation time of volumetric images during the calibration process. Incoherent compounding of US images was also investigated for 2D [32]

and 3D [33] imaging. The spatial transformation matrix between the positions of the arrays were 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 PA imaging, multiperspective or multiview imaging is almost inherent to the modality as PA tomography with spherical and hemispherical scans of an US detector has been investigated in the early PA scanners [34]. However, planar detection geometries able to acquire 3D PA images have been angulated to enhance the image view with a synthetic aperture approach [35], [36]. In ref [35], 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, segmenting the images of the threads and determining the rigid transformation between arrays. In ref [36], fiducial markers were incorporated in the imaged region. Multiview imaging was also performed by stitching volumetric images [37], which is similar to incoherent compounding and relies on the features obtained in partially overlapping images. In brief, no standard calibration exists either in USI or PAI, but only dedicated calibration phantoms were shown to allow synthetic aperture without fiducial markers. Additionally, while for a single imaging modality, some parameters could compensate each other, which put less constraints on the calibration method, in volumetric and simultaneous PAI-USI, the superposition and sharpness of the dual images could only be obtained with accurately determined parameters, in particular because of the different travel times in PAI and in USI between the US transducers and the voxels. To the best of your knowledge, we presented here the first dual PA-US 3D calibration method.

The co-registered volumetric maps of PAI and USI allow to obtain complementary information (Fig. 4 and 5). In the frame of longitudinal studies with different imaging sessions, the US image is expected to give the anatomical reference to better understand and to co-register images acquired at different time points, while PA images could reveal molecular or functional phenomena with a slow kinetic, such as the accumulation of a nanoparticle contrast agent with a long circulation time.

For other simultaneously co-registered multimodal imaging such as PET/CT, the anatomical imaging modality contains information that can be used to improve the image quality of the functional imaging modality. For PAI/USI, recent work on the reconstruction of 2D PA images showed an improvement of the PA image quality of blood vessels using structural information from ultrasound images [38]. Additionally, light fluence distribution could be modeled using US images and used to improve PA image quality [39].

The study presented here was performed at a single optical wavelength. Multispectral approaches [40] will be investigated with the developed scanner to enable discrimination between different chromophores and contrast agents and to evaluate physiological parameters such as the oxygen saturation. On the US side, the system presented here operated at an US center frequency of 5 MHz. However, the rotate-translate scan can be scaled to other ultrasound frequencies to gain sensibility and information on other spatial scale [22], [41]. The spatial

resolution increases with the US frequency of the array for both PAI and USI, 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 [16] will be investigated. For the translation in clinics, the bulky motorized stages will be replaced by a scanning system with a smaller foot print such as those developed for the translation scans [16], [17].

REFERENCES

- [1] K. Heinzmann, L. M. Carter, J. S. Lewis, and E. O. Aboagye, "Multiplexed imaging for diagnosis and therapy," *Nat. 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," *Semin. Nucl. Med.*, vol. 39, no. 5, pp. 348–353, 2009, doi: 10.1053/j.semnuclmed.2009.03.001.
- [3] Z. Hosseinaee, J. A. Tummon Simmons, and P. H. Reza, "Dual-Modal Photoacoustic Imaging and Optical Coherence Tomography [Review]," *Front. Phys.*, vol. 8, no. January, 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," *Nat. Biomed. Eng.*, vol. 2, no. 2, pp. 85–94, 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, no. July, p. 100144, 2019, 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. 639–50, Jul. 2016, doi: 10.1038/nmeth.3929.
- [7] A. Needles *et al.*, "Development and initial application of a fully integrated photoacoustic micro-ultrasound system," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 60, no. 5, pp. 888–897, May 2013, doi: 10.1109/tuffc.2013.2646.
- [8] M. Kuniyil Ajith 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 (2015)*, vol. 3, M. Olivo and U. S. Dinish, Eds. Singapore: Springer Singapore, 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, 2017, doi: 10.1063/1.4983462.
- [10] 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, 2020, doi: 10.1364/optica.400517.
- [11] X. L. Deán-Ben and D. Razansky, "Adding fifth dimension to optoacoustic imaging: Volumetric time-resolved spectrally enriched tomography," *Light Sci. Appl.*, vol. 3, no. May 2013, pp. 1–5, 2014, doi: 10.1038/lsa.2014.18.
- [12] 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. Imaging*, vol. 41, no. 4, pp. 846–856, 2022, doi: 10.1109/TMI.2021.3125398.
- [13] G. Godefroy, B. Arnal, S. Vilov, and E. Bossy, "3D photoacoustic fluctuation imaging provides visibility artefacts removal and enhanced contrast. Simultaneous implementation with ultrasound doppler imaging," in *Opto-Acoustic Methods and Applications in Biophotonics V*, Dec. 2021, vol. 2021, p. 11, doi: 10.1117/12.2615874.
- [14] 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, 2020, doi: 10.1109/TBME.2019.2963464.
- [15] 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, p. 061208, 2012, doi: 10.1117/1.jbo.17.6.061208.
- [16] C. Lee, W. Choi, J. Kim, and C. Kim, "Three-dimensional clinical handheld photoacoustic/ultrasound scanner," *Photoacoustics*, vol. 18, no. March, p. 100173, 2020, doi: 10.1016/j.pacs.2020.100173.
- [17] 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, 2022, doi: 10.1364/ol.458274.
- [18] 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, 2020, doi: 10.1038/s41598-020-64966-6.
- [19] 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, p. 034902, Jul. 2022, doi: 10.1063/5.0093619.
- [20] 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, 2022, doi: 10.1109/TBME.2022.3154754.
- [21] 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. Imaging*, vol. 33, no. 2, pp. 535–545, 2014, doi: 10.1109/tmi.2013.2289930.
- [22] J. Gateau, M. Gesnik, J.-M. 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. 5, p. 056004, 2015, doi: 10.1117/1.JBO.20.5.056004.
- [23] 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.
- [24] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 61, no. 1, pp. 102–119, Jan. 2014, doi: 10.1109/tuffc.2014.6689779.
- [25] W. Marczak, "Water as a standard in the measurements of speed of sound in liquids," *J. Acoust. Soc. Am.*, vol. 102, no. 5, pp. 2776–2779, 1997, doi: 10.1121/1.420332.
- [26] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nat. Methods*, vol. 9, no. 7, pp. 671–675, Jul. 2012, doi: 10.1038/nmeth.2089.
- [27] 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. Am.*, vol. 146, no. 1, pp. EL78–EL84, Jul. 2019, doi: 10.1121/1.5118238.
- [28] 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, p. 090501, 2011, doi: 10.1117/1.3619139.
- [29] 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, 2005, doi: 10.1016/j.ultrasmedbio.2004.11.015.
- [30] L. Peralta, A. Gomez, Y. Luan, B. 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.

- [31] L. Peralta, D. Mazierli, A. Gomez, J. V Hajnal, P. Tortoli, and A. Ramalli, "3-D Coherent Multi-Transducer Ultrasound Imaging with Sparse Spiral Arrays," pp. 1–10, Nov. 2022, [Online]. Available: <http://arxiv.org/abs/2211.03189>.
- [32] 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. Imaging*, vol. 39, no. 11, pp. 3714–3724, 2020, doi: 10.1109/TMI.2020.3003430.
- [33] 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, 2022, doi: 10.1016/j.ultrasmedbio.2022.09.005.
- [34] 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.
- [35] 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, p. 041009, 2016, doi: 10.1117/1.jbo.22.4.041009.
- [36] 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.
- [37] 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.
- [38] H. Yang *et al.*, "Soft ultrasound priors in optoacoustic reconstruction: Improving clinical vascular imaging," *Photoacoustics*, vol. 19, no. March, p. 100172, 2020, doi: 10.1016/j.pacs.2020.100172.
- [39] 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, p. 100222, 2021, doi: 10.1016/j.pacs.2020.100222.
- [40] V. Ntziachristos and D. Razansky, "Molecular Imaging by Means of Multispectral Optoacoustic Tomography (MSOT)," *Chem. Rev.*, vol. 110, no. 5, pp. 2783–2794, 2010, doi: 10.1021/Cr9002566.
- [41] 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.