

Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Context encoding enables machine learning-based quantitative photoacoustics

Thomas Kirchner
Janek Gröhl
Lena Maier-Hein

SPIE.

Thomas Kirchner, Janek Gröhl, Lena Maier-Hein, "Context encoding enables machine learning-based quantitative photoacoustics," *J. Biomed. Opt.* **23**(5), 056008 (2018), doi: 10.1117/1.JBO.23.5.056008.

Context encoding enables machine learning-based quantitative photoacoustics

Thomas Kirchner,^{a,b,*†} Janek Gröhl,^{a,c,†} and Lena Maier-Hein^{a,c,*}

^aGerman Cancer Research Center (DKFZ), Division of Computer Assisted Medical Interventions (CAMI), Heidelberg, Germany

^bHeidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany

^cHeidelberg University, Medical Faculty, Heidelberg, Germany

Abstract. Real-time monitoring of functional tissue parameters, such as local blood oxygenation, based on optical imaging could provide groundbreaking advances in the diagnosis and interventional therapy of various diseases. Although photoacoustic (PA) imaging is a modality with great potential to measure optical absorption deep inside tissue, quantification of the measurements remains a major challenge. We introduce the first machine learning-based approach to quantitative PA imaging (qPAI), which relies on learning the fluence in a voxel to deduce the corresponding optical absorption. The method encodes relevant information of the measured signal and the characteristics of the imaging system in voxel-based feature vectors, which allow the generation of thousands of training samples from a single simulated PA image. Comprehensive *in silico* experiments suggest that context encoding-qPAI enables highly accurate and robust quantification of the local fluence and thereby the optical absorption from PA images. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.23.5.056008](https://doi.org/10.1117/1.JBO.23.5.056008)]

Keywords: photoacoustics; quantification; multispectral imaging; machine learning.

Paper 170781R received Dec. 8, 2017; accepted for publication Apr. 25, 2018; published online May 18, 2018; corrected Apr. 7, 2020.

1 Introduction

Photoacoustic (PA) imaging is an imaging concept with a high potential for real-time monitoring of functional tissue parameters such as blood oxygenation deep inside tissue. It measures the acoustic waves arising from the stress-confined thermal response of optical absorption in tissue.¹ More specifically, a PA signal $S(v)$ in a location v is a pressure response to the locally absorbed energy $H(v)$, which, in turn, is a product of the absorption coefficient $\mu_a(v)$, the Grueneisen coefficient $\Gamma(v)$ and the light fluence $\phi(v)$

$$S(v) \propto H(v) = \mu_a(v) \cdot \Gamma(v) \cdot \phi(v). \quad (1)$$

Given that the local light fluence not only depends on the imaging setup but is also highly dependent on the optical properties of the surrounding tissue, quantification of optical absorption based on the measured PA signal is a major challenge.^{2,3} So far, the field of quantitative PA imaging (qPAI) has focused on model-based iterative optimization approaches to infer optical tissue parameters from measured signals (cf. e.g., Refs. 3–12). Although these methods are well suited for tomographic devices with high image quality (cf. e.g., Refs. 13–15) as used in small animal imaging, translational PA research with clinical ultrasound transducers or similar handheld devices (cf. e.g., Refs. 1 and 16–22) focuses on qualitative image analysis.

As an initial step toward clinical qPAI, we introduce a machine learning-based approach to quantifying PA

measurements. The approach features high robustness to noise while being computationally efficient. In contrast to all other approaches proposed to date, our method relies on learning the light fluence on a voxel level to deduce the corresponding optical absorption. Our core contribution is the development of a voxel-based context image (CI) that encodes relevant information of the measured signal voxel together with characteristics of the imaging system in a single feature vector. This enables us to tackle the challenge of fluence estimation as a machine learning problem that we can solve in a fast and robust manner. Comprehensive *in silico* experiments indicate high accuracy, speed, and robustness of the proposed context encoding (CE)-qPAI approach. This is demonstrated for estimation of (1) fluence and optical absorption from PA images, as well as (2) blood oxygen saturation as an example of functional imaging using multispectral PA images.

2 Materials and Methods

A common challenge when applying machine learning methods to biomedical imaging problems is the lack of labeled training data. In the context of PAI, a major issue is the strong dependence of the signal on the surrounding tissue. This renders separation of voxels from their context—as in surface optical imaging²³—impossible or highly inaccurate. Simulation of a sufficient number of training volumes covering a large range of tissue parameter variations, on the other hand, is computationally not feasible given the generally long runtime of Monte Carlo methods, which are currently the gold standard for the simulation of light transportation in tissue.¹¹

Inspired by an approach to shape matching, where the shape context is encoded in a so-called spin image specifically for each node in a mesh,²⁴ we encode the voxel-specific context in so-called CIs. This allows us to train machine learning algorithms

*Address all correspondence to: Thomas Kirchner, E-mail: t.kirchner@dkfz-heidelberg.de; Lena Maier-Hein, E-mail: l.maier-hein@dkfz-heidelberg.de

†These authors contributed equally.

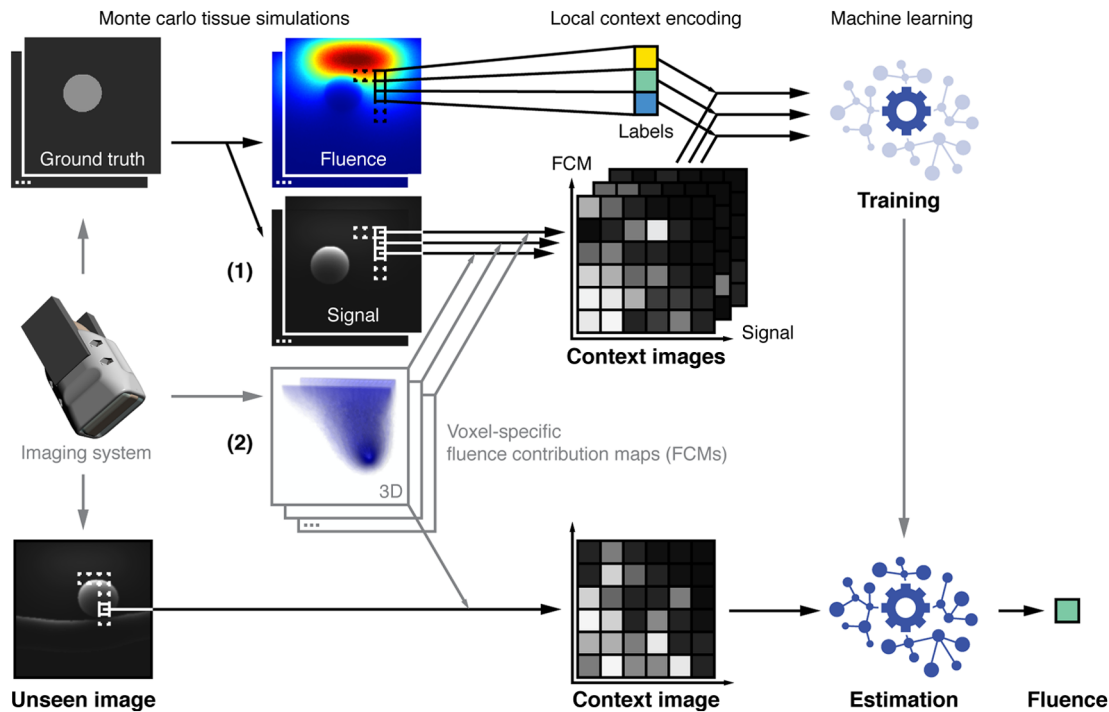


Fig. 1 Machine learning approach to fluence estimation with CIs. CIs are generated individually for each voxel and encode both (1) relevant information on the measured signal extracted from the PAI signal volume and (2) prior knowledge on the characteristics of the imaging system represented by FCMs. During algorithm training, a regressor is presented tuples of CIs and corresponding ground truth fluence values for each voxel in the training data. For estimation of optical absorption in voxels of a previously unseen image, the voxel-specific CI is generated and used to infer the local fluence using the trained regressor.

on a voxel level rather than image level and we thus require orders of magnitude fewer simulated training volumes. CIs encode relevant information of the measured signal as well as characteristics of the imaging system represented by so-called voxel-specific fluence contribution maps (FCMs). The CIs serve as a feature vector for said machine learning algorithm, which is trained to estimate fluence in a voxel. The entire quantification method is shown in Fig. 1, which serves as an overview with details given in the following sections.

2.1 Fluence Contribution Map

An important prerequisite for computing the CI for a voxel v is the computation of the corresponding FCM, referred to as $FCM[v]$. $FCM[v](v')$ represents a measure for the likelihood

that a photon arriving in voxel v has passed v' . In other words, an FCM reflects the impact of a PA signal in v' on the drop in fluence in voxel v . An illustration of an FCM corresponding to a typical handheld PA setup is shown in Fig. 2. The $FCM[v]$ is dependent on how the PA excitation light pulse propagates through homogeneous tissue to arrive in v given a chosen hardware setup. The $x \times y$ FCMs per imaging plane are generated once for each new hardware setup and each voxel in the imaging plane.

In this first implementation of the CE-qPAI concept, FCMs are simulated with the same resolution as the input data assuming a background absorption coefficient of 0.1 cm^{-1} and a constant reduced scattering coefficient of 1.5 cm^{-1} .²⁵ The number of photons is varied to achieve a consistent photon count in the

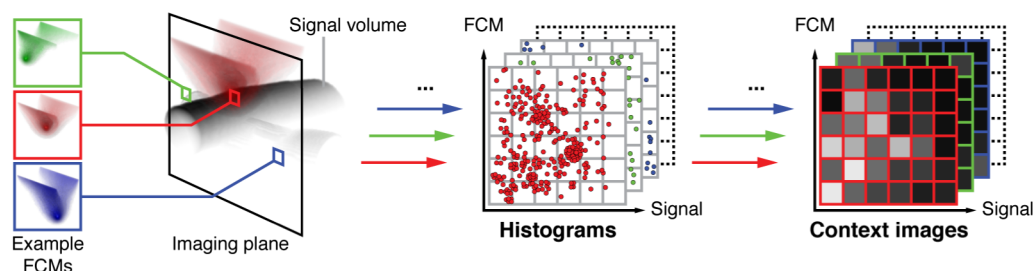


Fig. 2 Generation of CIs for three representative voxels based on their FCMs. The voxel-specific FCMs serve as a representation of how the PA excitation light pulse propagates through homogeneous tissue to arrive in a target voxel given a chosen hardware setup. For each voxel (here: green, red, and blue), tuples of measured signal and corresponding fluence contribution (for that voxel) are determined to generate the voxel-specific histograms from which the CI is generated.

target voxel. The FCMs are generated with the widely used Monte Carlo simulation tool mcxyz.²⁶ We integrated mcxyz into the open-source Medical Image Interaction Toolkit MITK²⁷ as mitkMcxyz and modified it to work in a multithreaded environment. Sample FCMs for three different voxels are shown in Fig. 2, which also shows the generation of CIs for those three example voxels.

2.2 Context Image

The CI for a voxel v in a PA volume is essentially a two-dimensional (2-D) histogram composed of (1) the measured PA signal S in the tissue surrounding v and (2) the corresponding FCM[v]. More specifically, it is constructed from the tuples $\{(S(v'), \text{FCM}[v](v')) | v' \in N(v)\}$ where $N(v)$ is defined as $N(v) = \{v' | \text{FCM}[v](v') > \epsilon\}$. This constraint is set to exclude voxels with a negligible contribution to the fluence in v . The tuples are arranged by magnitude of $S(v')$ and FCM[v](v') into a 2-D histogram and thereby encode the relevant context information in a compact form. In our prototype implementation of the CE-qPAI concept, the fluence contribution and signal axes of the histogram are discretized in 12 bins and scaled logarithmically to better represent the predominantly low signal and fluence contribution components. The ranges of the axes are set as $0 < \log(S) < \log(255)$ and $\log(\epsilon) < \log(\text{FCM}) < -1$. Signals and fluence contributions larger than the upper boundary are included in the highest bin, whereas smaller signals and fluence contributions are not. Figure 2 shows the generation of CIs from FCMs and PA signals. Labeled CIs are used for training a regressor that can later estimate fluence, which, in turn, is used to reconstruct absorption [Eq. (1)].

2.3 Machine Learning-Based Regression for Fluence Estimation

During the training phase, a regressor is presented tuples $[\text{CI}(v), \phi(v)]$ of $\text{CI}(v)$ and corresponding ground truth fluence values $\phi(v)$ for each voxel v in a set of PAI volumes. For estimation of optical absorption in a voxel v_u of a previously unseen image, the voxel-specific CI is generated and used to infer fluence $\phi(v_u)$ using the trained algorithm.

In our prototype implementation of the CE-qPAI method, we use a random forest regressor. A random forest regressor is an ensemble of decision trees, where the weighted vote of the individual trees is used as the estimation.²⁸ To train the random forest, all labeled CIs of the respective training set need to be evaluated at once. With voxel-based CIs, thousands of training samples can be extracted from a single slice of a simulated PA training volume. Ground truth training data generation is performed using a dedicated software plugin integrated into MITK and simulating the fluence with mitkMcxyz. It should be noted that the simulated images consist mainly of background voxels and not of vessel structures, which are our regions of interest (ROI). This leads to an imbalance in the training set. To avoid poor estimation for underrepresented classes,²⁹ we undersample background voxels in the training process to ensure a 1:1 ROI/background sample ratio. The parameters of the random forest are set to the defaults of sklearn 0.18 using python 2.7, except for the tree count which was set to $n_{\text{regressors}} = 100$. CIs are used as feature vectors and labeled with the optical property to be estimated (e.g., fluence or oxygenation). The parameters were chosen based on a grid search on a separate dataset not used in the experiments of this work.

2.4 Hardware Setup

We assume a typical linear probe hardware setup,³⁰ where the ultrasound detector array and the light source move together and the illumination geometry is the same for each image recorded. This is also the case for other typical tomographic devices.^{31,32} All simulations were performed on high-end CPUs (Intel i7-5960X).

3 Experiments and Results

In the following validation experiments, we quantify the fluence up to an imaging depth of 28 mm in unseen test images for each dataset. With our implementation and setup, all images comprise 3008 training samples, which results in an average simulation time of about 50 ms per training sample. This allows us to generate enough training samples in a feasible amount of time, to train a regressor that enables fluence estimation in a previously unseen image in near real time. The measured computational time for quantifying fluence in a single 64×47 voxel image slice is $0.9 \text{ s} \pm 0.1 \text{ s}$.

In the following, we present the experimental design and results of the validation of CE-qPAI. First, we will validate the estimation of absorption from PAI volumes acquired at a fixed wavelength and then estimate blood oxygenation from multispectral PAI volumes.

3.1 Monospectral Absorption Estimation

3.1.1 Experiment

To assess the performance of CE-qPAI in PA images of blood vessels, we designed six experimental datasets (DS) with varying complexities as listed in Table 1. With the exception of DS_{multi}, each of the six experimental DS is composed of 150 training items, 25 validation items, and 25 test items, where each item comprises a three-dimensional (3-D) simulated PA image of dimensions $64 \times 47 \times 62$ and 0.6-mm equal spacing as well as a corresponding (ground truth) fluence map.

As labels of the generated CIs, we used a fluence correction $\phi_c(v') = \phi(v')/\phi_h(v')$, where $\phi_h(v')$ is a fluence simulation based on a homogeneous background tissue assumption. We used five equidistant slices out of each volume, resulting in a generation of a total of 2,256,000; 376,000 and 376,000 CIs for each dataset—for training, parameter optimization, and

Table 1 The design parameters of the DS. All ranges denote sampling from uniform distributions within the given bounds.

Dataset	Vessel radius [mm]	Vessel absorption μ_a [cm^{-1}]	Vessel count	Background absorption μ_a [cm^{-1}]
DS _{base}	3	4.7	1	0.1
DS _{radius}	0.5 to 6	4.7	1	0.1
DS _{absorb}	3	1 to 12	1	0.1
DS _{vessel}	3	4.7	1 to 7	0.1
DS _{background}	3	4.7	1	10^{-4} to 0.2
DS _{multi}	0.5 to 6	1 to 12	1 to 7	10^{-4} to 0.2

testing, respectively. To account for the high complexity of DS_{multi} , we increased the number of training volumes for that set from 150 to 400. The baseline dataset DS_{base} represents simulations of a transcutaneously scanned simplified model of a blood vessel of constant radius (3 mm) and constant absorption (vessel: 4.73 cm^{-1} , background: 0.1 cm^{-1}) and reduced scattering coefficient (1.5 cm^{-1}). To approximate partial volume effects, the absorption coefficients in the ground truth images were Gaussian blurred with a sigma of 0.6 mm. Single slices were simulated using 2×10^6 photons for all training sets and 10^8 photons for the respective test and validation sets and then compounded in a fully scanned volume. Different shapes and poses of the vessel were generated by a random walk with steps r defined as

$$r_i = r_{i-1} + \eta \cdot a, \quad (2)$$

where η is a free parameter constant in each vessel with an inter-vessel variation within a uniform distribution ($0 < \eta < 0.2$) and a is varied for each of its components in each step within a uniform distribution ($-0.2 \text{ mm} < a_i < 0.2 \text{ mm}$). To investigate how variations in geometry and optical properties impact the performance of our method, we designed further experimental DS in which the number of vessels (DS_{vessel}), the radii of the vessels (DS_{radius}), the optical absorption coefficients within the vessels (DS_{absorb}), the absorption coefficient of the background ($DS_{\text{background}}$), as well as all of the above (DS_{multi}) were varied. We tested the robustness of CE-qPAI to this range of scenarios without retuning CI or random forest parameters.

Although most studies assess the performance of a method in the entire image (cf. e.g., Refs. 6, 33, and 34), it must be pointed out that the accuracy of signal quantification is often most relevant in a defined region of interest—such as in vessels or regions that provide a meaningful PA signal. These are typically also the regions, where quantification is particularly challenging due to the strongest signals originating from boundaries with discontinuous tissue properties. To address this important aspect we validated our method, not only on the entire image, but also in the ROI, which we define for our DS as voxels representing a vessel and at the same time having a contrast-to-noise ratio (CNR) of larger than 2, to only include significant signal in the ROI. We define CNR following Walvaert and Rosseel³⁵ in a voxel v as

$$\text{CNR} = \frac{S(v) - \text{avg}(b)}{\text{std}(b)}, \quad (3)$$

where the $\text{avg}(b)$ and $\text{std}(b)$ are the average and standard deviations of the background signal b over a simulated image slice with a background absorption coefficient of 0.1 cm^{-1} and no other structures. Using such an image without application of a noise model, we simulated an intrinsic background noise of (4.2 ± 2.8) a.u.

To investigate the robustness of CE-qPAI to noise, we added the following noise models to each dataset. The noise models consist of an additive Gaussian noise term applied on the signal volumes followed by a multiplicative white Gaussian noise term, similar to noise assumptions used in prior work.^{6,33} We examined three noise levels to compare against the simulation-intrinsic noise case:

1. 2% multiplicative and (0.125 ± 0.125) a.u. additive component
2. 10% multiplicative and (0.625 ± 0.625) a.u. additive component
3. 20% multiplicative and (1.25 ± 1.25) a.u. additive component

The additive and multiplicative noise components follow an estimation of noise components on a custom PA system.³⁰ For each experimental dataset introduced in Table 1 and each noise set, we applied the following validation procedure separately. Following common research practice, we used the training data subset for training of the random forest and the validation data subset to ensure the convergence of the training process, as well as to set suitable parameters for the random forest and ROI, whereas we only evaluated the test data subset to report the final results (as described in Ref. 36). As an error metric, we report the relative fluence estimation error e_r

$$e_r(v) = \frac{|\hat{\phi}(v) - \phi(v)|}{\phi(v)}, \quad (4)$$

rather than an absorption estimation error, to separate the error in estimating fluence with CE-qPAI from errors introduced through simulation-intrinsic or added noise on the signal, which will affect the quantification regardless of fluence estimation.

3.1.2 Results

Figures 3(a)–3(c) show representative examples of the previously unseen 125 simulated test images from the baseline dataset DS_{base} , with their corresponding fluence estimation results. The optical absorption is reconstructed using the fluence estimation. A histogram illustrating absorption estimation accuracy in ROI voxels of DS_{base} is shown in Fig. 3(d) and compared with a static fluence correction approach.

Table 2 summarizes the descriptive statistics of the relative fluence estimation errors e_r for the experiments on absorption estimation using single wavelength PA images. The relative fluence estimation error e_r does not follow a normal distribution due to large outliers especially in complex DS, which is why we report median e_r with interquartile ranges (IQR) for all DS. Even for the most complex dataset DS_{multi} with variations of multiple parameters, CE-qPAI yields a median overall relative fluence estimation error e_r below 4%. Errors are higher in the ROI, especially in DS with high variations of absorption.

Previously proposed qPAI approaches reveal high drops in estimation performance when dealing with noisy data (cf. e.g., Ref. 37). To remedy this, methods have been proposed to incorporate more accurate noise representations into model-based reconstruction algorithms.^{33,38} When validating the robustness of CE-qPAI to noise, it yields high accuracy even under unrealistically high noise levels of up to 20% (cf. Fig. 4). Regardless of the noise level applied, the highest median errors occur in the ROIs of DS that are characterized by high absorption and inhomogeneous tissue properties.

3.2 Multispectral Blood Oxygenation Estimation

The concept of CE cannot only be used to estimate fluence and absorption, but also derived functional parameters such as

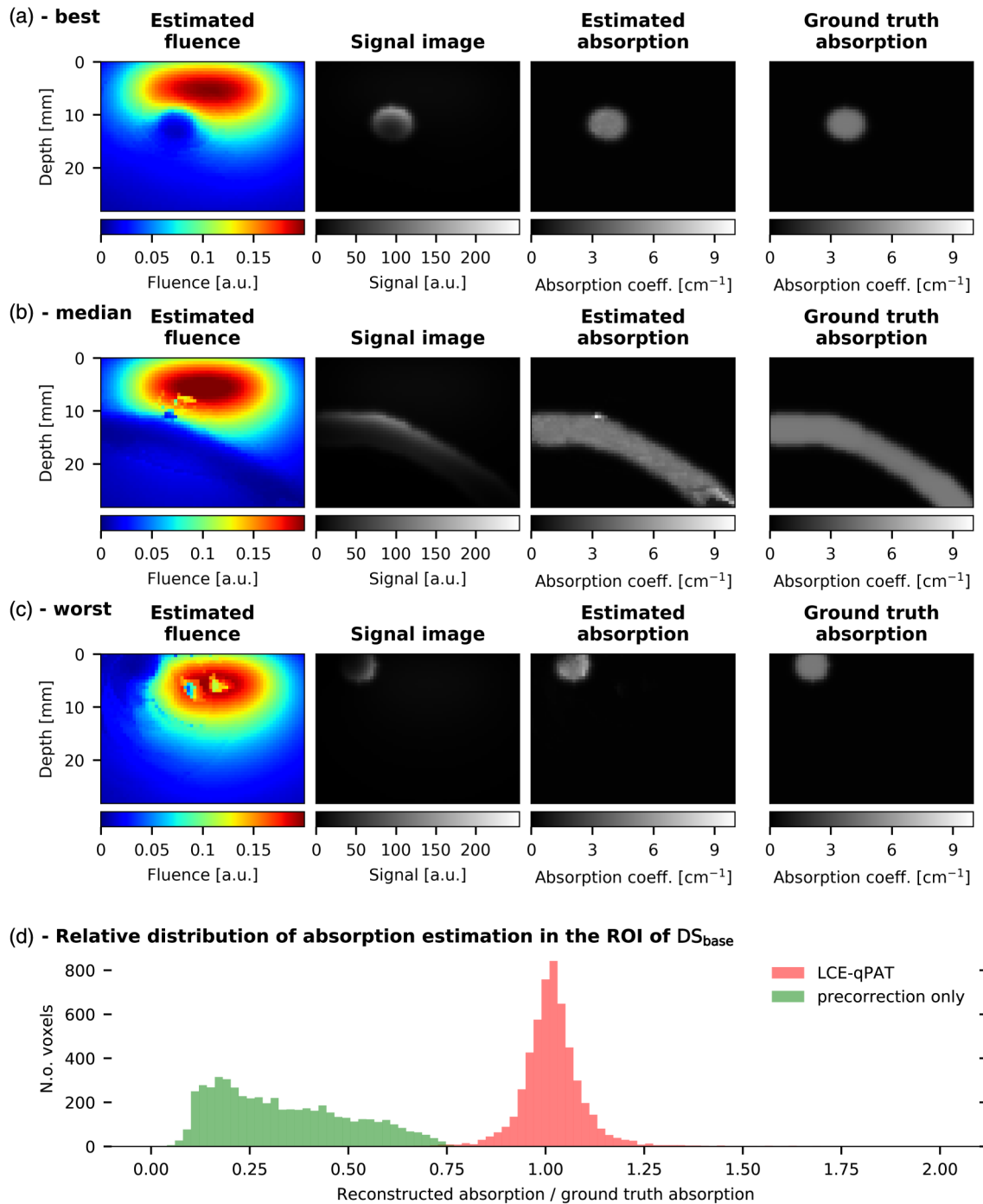


Fig. 3 Absorption reconstruction results after fluence estimation. For the slices with the (a) lowest, (b) median, and (c) highest median fluence estimation error e_f within the ROI of DS_{base} . We show (from left to right) the estimated fluence, the corresponding signal images, the resulting estimation of the absorption coefficient, and the ground truth optical absorption, for reference. (d) A histogram of the relative absorption estimation over all ROI voxels ($n = 5347$) in DS_{base} illustrating absorption estimation accuracy rather than fluence estimation accuracy measured by e_f . Precorrecting the signal with the fluence of a homogeneous tissue assumption underestimates the absorption and is considerably outperformed by CE-qPAI in the ROI. The CE-qPAI plot omits 5 outliers larger 2.

blood oxygenation. To this end, the estimated absorption in a voxel for multiple wavelengths can be applied to resolve oxygenation via linear spectral unmixing. Alternatively, a regressor can be trained using the CIs labeled with ground truth oxygenation.

3.2.1 Experiment

To investigate the performance of CE-qPAI for blood oxygenation (sO_2) estimation, we designed an additional multispectral simulated dataset DS_{oxy} using the wavelengths 750, 800, and 850 nm. It consists of 240 multispectral training volumes and

Table 2 Descriptive statistics of fluence estimation results. The median and IQR of the relative fluence estimation error e_r for the six validation DS used for the single wavelength experiments. The median error and IQR are provided (1) for all voxels in the respective test set as well as (2) for the voxels in the ROI only.

Dataset	Relative error e_r			
	All voxels		ROI	
	Median (%)	IQR (%)	Median (%)	IQR (%)
DS _{base}	1.0	(0.5, 1.9)	4.2	(1.9, 7.6)
DS _{radius}	1.4	(0.6, 3.3)	5.7	(2.4, 11.3)
DS _{absorb}	1.2	(0.5, 2.8)	14.7	(5.4, 32.2)
DS _{vessel}	1.8	(0.7, 6.2)	6.8	(3.0, 13.2)
DS _{background}	0.7	(0.3, 1.4)	4.1	(1.7, 7.3)
DS _{multi}	2.3	(0.7, 38.5)	15.7	(6.6, 40.0)

11 multispectral test volumes, each featuring homogeneous oxygenation and one vessel with a radius of 2.3 to 4 mm—modeled after a carotid artery.³⁹ For each image slice and at each wavelength, 10^7 photons were used for simulation. Oxygenation values for the training images were drawn randomly from a uniform sO_2 distribution $U(0\%, 100\%)$. For testing, we simulated 11 multispectral volumes at three wavelengths and 11 blood oxygenation levels ($sO_2 \in \{0\%, 10\%, 20\%, \dots, 100\%\}$). The optical absorption was adjusted by wavelength and oxygenation, as described by Jacques.²⁵ Hemoglobin concentration was assumed to be 150 g/L.²⁵ The blood volume fraction was set to 0.5% in the background tissue and to 100% in the blood vessels. The reduced scattering coefficient was again set to 1.5 cm^{-1} . We estimated the oxygenation using three methods:

1. *Linear spectral unmixing on the signal images as a baseline.*⁴⁰ For this, we applied a non-negative constrained least squares approach as also used in Ref. 15 that minimizes $\|Ax - \mathbf{b}\| = 0$, where A is the matrix containing the reference spectra, \mathbf{b} is the measurement vector, and x is the unmixing result.

Specifically, we used the python `scipy.optimize.minimize` function with the sequential least squares programming method and added a non-negativity inequality constraint. We evaluated the unmixing results of this method on all voxels in the ROI as well as exclusively on those voxels with the maximum intensity projection (MIP) along image x -axis at wavelength 800 nm to account for nonlinear fluence effects deep inside the vessels.

2. *Linear spectral unmixing of the signal after quantification of the three input images with CE-qPAI.* After correcting the raw signal images for nonlinear fluence effects using CE-qPAI, we applied the same method as described in (1) and evaluated on the same voxels that were used in (1) to ensure comparability of the results.
3. *Direct estimation of oxygenation using a functional adaptation of CE-qPAI.* For functional CE-qPAI (fCE-qPAI), triples of CIs for the three chosen wavelengths were concatenated into one feature vector and labeled with the ground truth oxygenation.

3.2.2 Results

Estimation of local blood oxygen saturation (sO_2) is one of the main qPAI applications and is only possible with multispectral measurements. As such, the presented approaches were validated together with the baseline method on the dataset DS_{oxy}. As shown in Fig. 5(a), the estimation results for both methods are in very close agreement with the ground truth. In fact, the median absolute oxygen estimation error was 3.1% with IQR (1.1% and 6.4%) for CE-qPAI and 0.8% with IQR (0.3% and 1.8%) for the fCE-qPAI adaptation. Furthermore, our methodology outperforms a baseline approach based on linear spectral unmixing of the raw signal (as also compared to in Ref. 15). By means of example Fig. 5(b) shows that the linear spectral unmixing of the ROI on the uncorrected signal fails deep inside the ROI, where the fluence varies strongly for different wavelengths. To compensate for this effect when comparing the approach to our method, we validate all methods only on the MIP along the depth axis (as also used in Ref. 41) in Fig. 5(a).

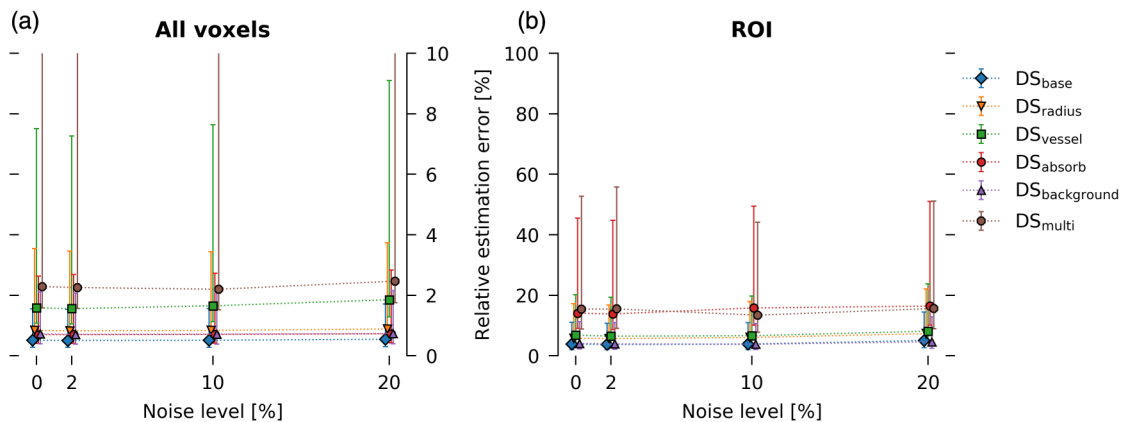


Fig. 4 Robustness of the fluence estimation against noise. Median relative fluence estimation errors e_r with IQR over all DS for, (a) all test voxels, and (b) in region of interest test voxels. The whiskers in this plot show the first and third quartile.

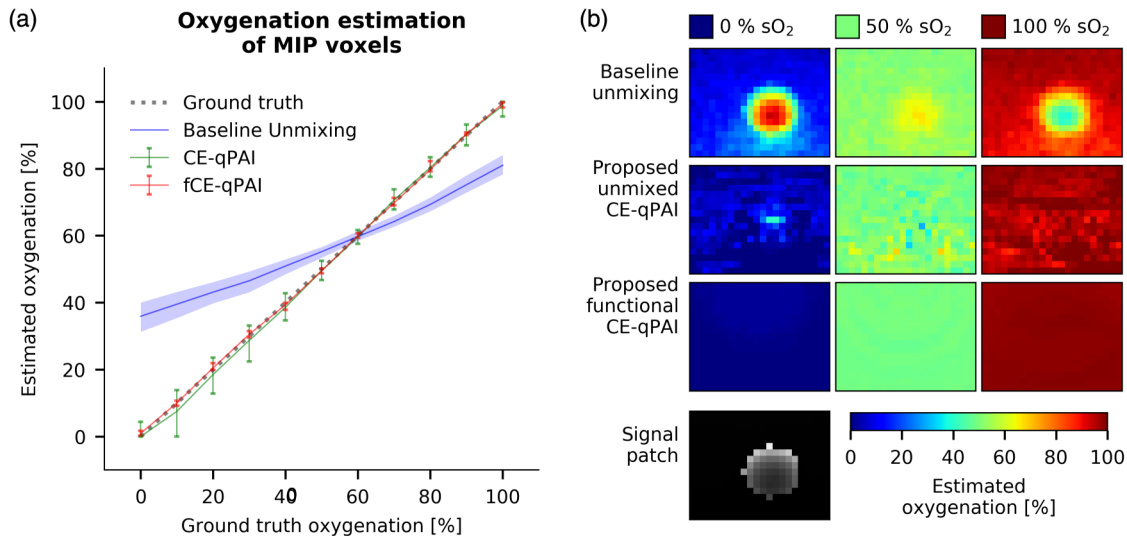


Fig. 5 Oxygenation estimation. (a) The median oxygen estimation with the IQR on the MIP voxels using linear spectral unmixing of (blue) the uncorrected signal, (green) the signal corrected by CE-qPAI, and (red) direct estimation by functional CE-qPAI (fCE-qPAI). (b) The oxygenation estimation for a representative patch of signal showing a vessel in 15-mm depth and with 3-mm radius. The signal for one of the measurement wavelengths is shown for reference together with the oxygen estimation results for 0%, 50%, and 100% ground truth homogeneous oxygenation and the three examined methods.

4 Discussion

This paper addresses one of the most important challenges related to PA imaging, namely the quantification of optical absorption based on the measured signal. In contrast to all other approaches proposed to qPAI to date (cf. e.g., Refs. 3–12), our method relies on learning the light fluence in a voxel to deduce the corresponding optical absorption. Comprehensive *in silico* experiments presented in this manuscript show the high potential of this approach to estimate optical absorption as well as derived functional properties, such as oxygenation, even in the presence of high noise.

Although machine learning methods have recently been applied to PAI related problems (cf. e.g., Refs. 42–44), these have mainly focused on image reconstruction but not signal quantification. We attribute this to the fact that *in vivo* training data generation for machine learning-based qPAI is not at all straightforward given the lack of reference methods for estimating optical absorption in depth. Despite recent developments related to hybrid diffusion approximation and Monte Carlo methods,⁴⁵ fast generation of *in silico* training data also remains an unsolved challenge. Note in this context that commonly applied methods of data augmentation (i.e., methods that may be used to automatically enlarge training data sets as discussed in Ref. 46) cannot be applied to PA images due to the interdependence of fluence and signal. With our contribution, we have addressed the challenge by introducing the concept of CIs, which allow us to generate one training case from each voxel rather than from each image.

As an important contribution with high potential impact, we adapted CE-qPAI to estimate functional tissue properties from multiwavelength data. Both variants—linear spectral unmixing of the fluence corrected signal, as well as direct estimation of oxygenation from multi wavelength CIs, yielded accurate results that outperformed a baseline approach based on linear spectral unmixing of the raw PA signal. It should be noted that linear spectral unmixing of the signal for sO₂ estimation is usually

performed on a wider range of wavelengths to increase accuracy. However, even this increase in the number of wavelengths cannot fully account for nonlinear fluence effects.³ Combined with the separately established robustness to noise, multiwavelength applications of CE-qPAI are very promising.

In our first prototype implementation of CE-qPAI, we used random forests regressors with standard parameters. It should be noted, however, that fluence estimation from the proposed CI can in principle be performed by any other machine learning method in a straightforward manner. Initial experiments suggest that even better performance can be achieved with convolutional neural networks.⁴⁷

By relating the measured signals $S(v')$ in the neighborhood of v to the corresponding fluence contributions $FCM[v](v')$ we relate the absorbed energy in v' , to the fluence contribution of v' to v . In this context, it has to be noted that the fluence contribution $FCM[v](v')$ is only an approximation of the true likelihood that a photon passing v has previously passed v' , because $FCM[v]$ is generated independently of the scene under observation assuming constant background absorption and scattering. Nevertheless due to the generally low variance of scattering in tissue, it serves as a reliable input for the proposed machine learning-based quantification.

A limitation of our study can be seen in the fact that we performed the validation *in silico*. To apply CE-qPAI *in vivo*, further research will have to be conducted in two main areas. First, we are working on accurately solving the acoustical inverse problem for specific scanners.⁴⁸ The method will be integrated into the quantification algorithm to enable quantification of images acquired with common PAI probes such as clinical linear transducers. Second, training data have to be generated as close to reality as possible—considering, for example, imaging artifacts.

In contrast to prior work (cf. e.g., Refs. 6, 7, 33, 49, and 34), our initial validation handles the whole range of near infrared absorption in whole blood at physiological hemoglobin concentrations and demonstrates high robustness to noise. The impact

of variations of scattering still needs investigation although these should be small in the near infrared.

Long-term goal of our work is the transfer of CE-qPAI to clinical data. In this context, run-time of the algorithm will play an important role. Although our current implementation can estimate absorption on single slices within a second, this might not be sufficient for interventional clinical estimation of whole tissue volumes and at higher resolutions. An efficient GPU implementation of the time intensive CI generation should enable real-time quantification.

In summary, CE-qPAI is the first machine learning-based approach to quantification of PA signals. The results of this work suggest that quantitative real-time functional PA imaging deep inside tissue is feasible.

Code and Data Availability

The code for the method as well as the experiments was written in C++ and python 2.7 and is partially open source and available at <https://phabricator.mitk.org/source/mitk.git>. Additional code and all raw and processed data generated in this work are available from the corresponding authors on reasonable request.

Disclosures

The authors have no relevant financial interests in this article and no potential conflicts of interest to disclose.

Acknowledgments

The authors would like to acknowledge support from the European Union through the ERC starting grant COMBIOSCOPY under the New Horizon Framework Programme grant agreement ERC-2015-StG-37960. We would like to thank the ITCF of the DKFZ for the provision of their computing cluster and C. Feldmann for her support with figure design.

References

- L. V. Wang and J. Yao, "A practical guide to photoacoustic tomography in the life sciences," *Nat. Methods* **13**(8), 627–638 (2016).
- L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**(6075), 1458–1462 (2012).
- B. T. Cox, J. G. Laufer, and P. C. Beard, "The challenges for quantitative photoacoustic imaging," *Proc. SPIE* **7177**, 717713 (2009).
- N. Iftimia and H. Jiang, "Quantitative optical image reconstruction of turbid media by use of direct-current measurements," *Appl. Opt.* **39**(28), 5256–5261 (2000).
- B. T. Cox et al., "Quantitative photoacoustic imaging: fitting a model of light transport to the initial pressure distribution," *Proc. SPIE* **5697**, 49–55 (2005).
- B. T. Cox et al., "Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method," *Appl. Opt.* **45**(8), 1866–1875 (2006).
- Z. Yuan and H. Jiang, "Quantitative photoacoustic tomography: recovery of optical absorption coefficient maps of heterogeneous media," *Appl. Phys. Lett.* **88**(23), 231101 (2006).
- J. Laufer et al., "Quantitative spatially resolved measurement of tissue chromophore concentrations using photoacoustic spectroscopy: application to the measurement of blood oxygenation and haemoglobin concentration," *Phys. Med. Biol.* **52**(1), 141–168 (2007).
- E. Malone, B. Cox, and S. Arridge, "Multispectral reconstruction methods for quantitative photoacoustic tomography," *Proc. SPIE* **9708**, 970827 (2016).
- M. Haltmeier, L. Neumann, and S. Rabanser, "Single-stage reconstruction algorithm for quantitative photoacoustic tomography," *Inverse Probl.* **31**(6), 065005 (2015).
- B. Cox et al., "Quantitative spectroscopic photoacoustic imaging: a review," *J. Biomed. Opt.* **17**(6), 061202 (2012).
- B. Banerjee et al., "Quantitative photoacoustic tomography from boundary pressure measurements: noniterative recovery of optical absorption coefficient from the reconstructed absorbed energy map," *J. Opt. Soc. Am. A* **25**(9), 2347–2356 (2008).
- L. V. Wang, "Multiscale photoacoustic microscopy and computed tomography," *Nat. Photonics* **3**(9), 503–509 (2009).
- J. Xia and L. V. Wang, "Small-animal whole-body photoacoustic tomography: a review," *IEEE Trans. Biomed. Eng.* **61**(5), 1380–1389 (2014).
- S. Tzoumas et al., "Eigenspectra photoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," *Nat. Commun.* **7**, 12121 (2016).
- J. J. Niederhauser et al., "Combined ultrasound and photoacoustic system for real-time high-contrast vascular imaging in vivo," *IEEE Trans. Med. Imaging* **24**(4), 436–440 (2005).
- S. Zackrisson, S. M. W. Y. van de Ven, and S. S. Gambhir, "Light in and sound out: emerging translational strategies for photoacoustic imaging," *Cancer Res.* **74**(4), 979–1004 (2014).
- P. K. Upputuri and M. Pramanik, "Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review," *J. Biomed. Opt.* **22**(4), 041006 (2017).
- J. Gamelin et al., "Curved array photoacoustic tomographic system for small animal imaging," *J. Biomed. Opt.* **13**(2), 024007 (2008).
- K. H. Song et al., "Noninvasive photoacoustic identification of sentinel lymph nodes containing methylene blue in vivo in a rat model," *J. Biomed. Opt.* **13**(5), 054033 (2008).
- C. Kim et al., "Handheld array-based photoacoustic probe for guiding needle biopsy of sentinel lymph nodes," *J. Biomed. Opt.* **15**(4), 046010 (2010).
- A. Garcia-Urbe et al., "Dual-modality photoacoustic and ultrasound imaging system for noninvasive sentinel lymph node detection in patients with breast cancer," *Sci. Rep.* **5**, 15748 (2015).
- S. J. Wirkert et al., "Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse Monte Carlo and random forest regression," *Int. J. Comput. Assist. Radiol. Surg.* **11**(6), 909–917 (2016).
- A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(5), 433–449 (1999).
- S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**(11), R37–R61 (2013).
- S. L. Jacques, "Coupling 3D Monte Carlo light transport in optically heterogeneous tissues to photoacoustic signal generation," *Photoacoustics* **2**(4), 137–142 (2014).
- I. Wolf et al., "The medical imaging interaction toolkit," *Med. Image Anal.* **9**(6), 594–604 (2005).
- L. Breiman, "Random forests," *Mach. Learn.* **45**(1), 5–32 (2001).
- A. Estabrooks, T. Jo, and N. Japkowicz, "A multiple resampling method for learning from imbalanced data sets," *Comput. Intell.* **20**(1), 18–36 (2004).
- T. Kirchner et al., "Freehand photoacoustic tomography for 3D angiography using local gradient information," *Proc. SPIE* **9708**, 97083G (2016).
- V. Neuschmelting et al., "Performance of a multispectral photoacoustic tomography (MSOT) system equipped with 2D vs. 3D handheld probes for potential clinical translation," *Photoacoustics* **4**(1), 1–10 (2016).
- A. Needles et al., "Development and initial application of a fully integrated photoacoustic micro-ultrasound system," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **60**(5), 888–897 (2013).
- T. Tarvainen et al., "Bayesian image reconstruction in quantitative photoacoustic tomography," *IEEE Trans. Med. Imaging* **32**(12), 2287–2298 (2013).
- R. J. Zemp, "Quantitative photoacoustic tomography with multiple optical sources," *Appl. Opt.* **49**(18), 3566–3572 (2010).
- M. Welvaert and Y. Rosseel, "On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data," *PLoS One* **8**(11), e77089 (2013).
- B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge (2007).

37. E. Beretta et al., "A variational method for quantitative photoacoustic tomography with piecewise constant coefficients," Chapter 6 in *Variational Methods*, M. Bergounioux et al., Eds., pp. 202–224, Walter de Gruyter (2016).
38. T. Tarvainen et al., "Image reconstruction with noise and error modeling in quantitative photoacoustic tomography," *Proc. SPIE* **9708**, 97083Q (2016).
39. J. Krejza et al., "Carotid artery diameter in men and women and the relation to body and neck size," *Stroke* **37**(4), 1103–1105 (2006).
40. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.* **19**(1), 44–57 (2002).
41. X. L. Deán-Ben, E. Bay, and D. Razansky, "Functional optoacoustic imaging of moving objects using microsecond-delay acquisition of multispectral three-dimensional tomographic data," *Sci. Rep.* **4**, 5878 (2014).
42. A. Reiter and M. A. L. Bell, "A machine learning approach to identifying point source locations in photoacoustic data," *Proc. SPIE* **10064**, 100643J (2017).
43. A. Hauptmann et al., "Model based learning for accelerated, limited-view 3D photoacoustic tomography," arXiv:1708.09832v1 (2017).
44. S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," arXiv:1704.04587v2 (2017).
45. C. Zhu and Q. Liu, "Hybrid method for fast Monte Carlo simulation of diffuse reflectance from a multilayered tissue model with tumor-like heterogeneities," *J. Biomed. Opt.* **17**(1), 010501 (2012).
46. A. Dosovitskiy et al., "Discriminative unsupervised feature learning with convolutional neural networks," in *Advances in Neural Information Processing Systems*, Vol. 27, pp. 766–774 (2014).
47. K. He et al., "Deep residual learning for image recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* (2016).
48. D. Waibel et al., "Reconstruction of initial pressure from limited view photoacoustic images using deep learning," *Proc. SPIE* **10494**, 104942S (2018).
49. W. Naetar and O. Scherzer, "Quantitative photoacoustic tomography with piecewise constant material parameters," *SIAM J. Imaging Sci.* **7**(3), 1755–1774 (2014).

Thomas Kirchner received his MSc degree in physics from the University of Heidelberg in 2015. He currently works on his PhD at the Division of Computer Assisted Medical Interventions, German Cancer Research Center (DKFZ), where he does research in computational biophotonics, focusing on real-time multispectral photoacoustics and signal quantification.

Janek Gröhl received his MSc degree in medical informatics from the University of Heidelberg and Heilbronn University of Applied Sciences in 2016. He currently works on his PhD at the Division of Computer Assisted Medical Interventions (CAMI), German Cancer Research Center (DKFZ) and does research in software engineering and computational biophotonics focusing on signal quantification in photoacoustic imaging.

Lena Maier-Hein received her PhD from the Karlsruhe Institute of Technology in 2009 and conducted her postdoctoral research in the Division of Medical and Biological Informatics, German Cancer Research Center (DKFZ), and in the Hamlyn Centre for Robotics Surgery, Imperial College London. She is leading the Division of Computer Assisted Medical Interventions (CAMI) at the DKFZ. Currently, she is working on multimodal image processing, surgical data science, and computational biophotonics.