

Washington University School of Medicine

Digital Commons@Becker

Open Access Publications

2021

Ultrasound segmentation-guided edge artifact reduction in diffuse optical tomography using connected component analysis

Shuying Li

Menghao Zhang

Quing Zhu

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs



Ultrasound segmentation-guided edge artifact reduction in diffuse optical tomography using connected component analysis

SHUYING LI,¹ MENGHAO ZHANG,² AND QUING ZHU^{1,3,*} 

¹Department of Biomedical Engineering, Washington University in St. Louis, 1 Brookings Dr, St. Louis 63130, USA

²Department of Electrical & Systems Engineering, Washington University in St. Louis, 1 Brookings Dr, St. Louis 63130, USA

³Department of Radiology, Washington University School of Medicine, 660 S Euclid Ave, St. Louis 63110, USA

*zhu.q@wustl.edu

Abstract: Ultrasound (US)-guided diffuse optical tomography (DOT) has demonstrated potential value for breast cancer diagnosis and treatment response assessment. However, in clinical use, the chest wall, poor probe-tissue contact, and tissue heterogeneity can all cause image artifacts. These image artifacts, appearing commonly as hot spots in the non-lesion regions (edge artifacts), can decrease the reconstruction accuracy and cause misinterpretation of lesion images. Here we introduce an iterative, connected component analysis-based image artifact reduction algorithm. A convolutional neural network (CNN) is used to segment co-registered US images to extract the lesion location and size to guide the artifact reduction. We demonstrate its performance using Monte Carlo simulations on VICTRE digital breast phantoms and breast patient images. In simulated tissue mismatch models, this algorithm successfully reduces edge artifacts without significantly changing the reconstructed target absorption coefficients. With clinical data it improves the optical contrast between malignant and benign groups, from 1.55 without artifact reduction to 1.91 with artifact reduction. The proposed algorithm has a broad range of applications in other modality-guided DOT imaging.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Near-infrared diffuse optical tomography (DOT) has demonstrated its great potential in breast cancer diagnosis and treatment response assessment [1–6]. However, due to the intense light scattering, DOT reconstruction is ill-posed and ill-conditioned. To improve DOT reconstruction, other imaging modalities, such as magnetic resonance imaging (MRI) [2,7], X-ray CT or mammography [8,9], and ultrasound (US) [10–16], have been employed. These modalities provide prior information to DOT imaging reconstruction algorithms to improve the reconstruction accuracy. We have developed a US-guided DOT technique for diagnosing and characterizing US-visible breast lesions, with the goal of reducing benign breast biopsies. For US-guided DOT to be successful, a high optical contrast between malignant and benign breast lesions is crucial.

In clinical studies, the measured diffuse reflectance from the lesion side breast is typically normalized to the contralateral normal breast (the reference side) to produce perturbation for imaging reconstruction [17]. However, perturbation is sensitive to probe positions due to tissue heterogeneity differences (tissue compositions, and chest wall positions) underneath the probe, and to poor probe-tissue contact, which can generate image artifacts, appearing commonly as hot spots in the non-lesion regions (edge artifacts). Reducing these artifacts can more accurately quantify breast lesions and improve diagnosis. Our group has previously reported on the effects of chest wall mismatch and optode-tissue coupling mismatch and their correction [17–20]. The correction approaches are targeted at a single error source. Elsewhere in the literature, optical

properties and optode-tissue coupling coefficients are optimized simultaneously to compensate for optode coupling errors [21–24]. This compensation method assumes a constant coupling coefficient for each optode among all measurements involving that optode. Hence, it may not be applicable to reducing image artifacts caused by a larger scale tissue component, such as the chest wall or tissue heterogeneity. Based on the assumption that absorption maps of different wavelengths are similar, Uddin et al. proposed a correction algorithm to remove measurements that have large model errors [25]. However, by reducing the number of measurements, this method may lose information. Thus, a correction algorithm is needed that can deal with multiple error sources without reducing useful measurement information. Connected component analysis, a method for detecting connected regions in binary images, has been applied in medical imaging segmentation [26,27], and we adopt it here to remove edge artifacts in DOT images.

By modeling breast tissue as a random matrix with uniformly distributed entries, Ntziachristos et al. showed that tissue heterogeneity can cause image artifacts and decrease lesion reconstruction accuracy [28]. A more realistic heterogeneous tissue model could better demonstrate tissue heterogeneity caused image artifacts and guide the development of correction algorithms. Loue et al. developed anatomically realistic numerical breast phantoms from clinical contrast-enhanced MRI data [29]. Deng et al. extracted realistic digital breast phantoms from dual-energy X-ray mammographic imaging of human breasts and used them in DOT in-silico tests [30,31]. The numbers of these digital phantoms are limited because the limited availability of real patient measurements, and they cannot be customized to represent various types of breasts. Recently, the FDA has released an open-source software model – the VICTRE Breast Phantom [32]. The software can flexibly customize digital breast phantoms to resemble human breasts with different sizes, shapes, densities, and so on. It has been applied to X-ray imaging and photoacoustic tomography [32–35]. Here, we will use VICTRE Breast Phantom to create image edge artifacts arising from tissue background mismatch.

In this study, we propose a fully automated, connected components analysis-based algorithm to reduce edge artifacts. We adopt VICTRE breast phantoms in DOT for the first time to account for tissue heterogeneity, and we demonstrate the effectiveness of the artifact reduction algorithm. The algorithm, combined with an interactive convolutional neural network (CNN) that segments coregistered US images to extract the lesions, is also evaluated with clinical data that includes 14 patients of seven benign and seven malignant lesions.

2. Methods

2.1. US-guided DOT system

Our compact, frequency domain US-guided DOT system was previously described in Ref. [15]. Briefly, four laser diodes, with wavelengths of 740, 780, 808, and 830 nm, are modulated at 140 MHz and sequentially delivered light to nine source positions on the probe. The light reflected from the tissue is detected by ten parallel photomultiplier tube (PMT) detectors. The source-detector separations range from 0.9 cm to 7.4 cm, and the arrangements are shown in Fig. 1. The detected signals are demodulated to 20 kHz by a local oscillator of 140.02 MHz and then digitized and sampled by two 8-channel A/D data acquisition cards. A US probe in the middle of the DOT probe provides the lesion shape and location.

2.2. Monte Carlo simulation on VICTRE breast phantom

The VICTRE software was used to generate digital breast phantoms with a voxel size of $0.5 \times 0.5 \times 0.5 \text{ mm}^3$. The phantoms measured 14 cm in the x- and y-direction and 8 cm in the z-direction. Breast phantoms with different densities were created by varying the fat fraction from 20% to 80%, in 20% steps. The breast phantoms consisted of five tissue types – fat, glandular,

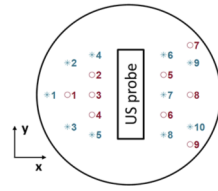


Fig. 1. US-guided DOT probe geometry.

skin, muscle (chest wall), and blood vessels, and cross-sections of representative digital breasts are shown in Fig. 2.

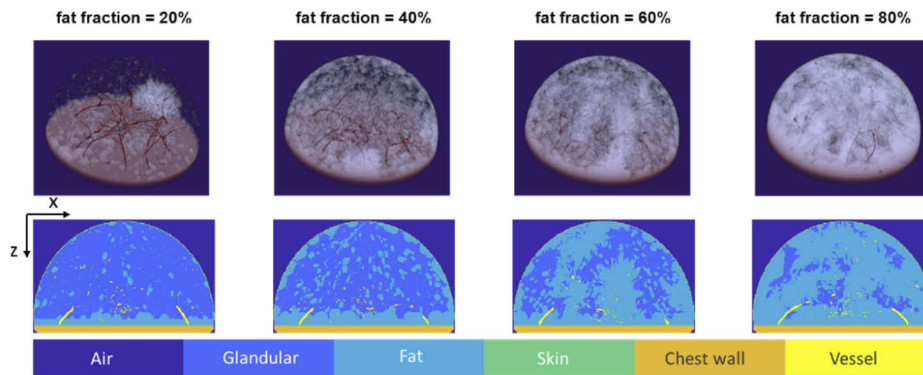


Fig. 2. Three-dimensional visualizations (top row) and cross-sections (bottom row) of selected digital breast phantoms with different densities.

The digital phantoms were numerically compressed uniformly to 5 cm in the z-direction to simulate compression by our hand-held reflection-mode DOT probe. After compression, the volume was down-sampled using nearest neighbor interpolation to achieve a voxel size of $2.5 \times 2.5 \times 2.5 \text{ mm}^3$ for Monte Carlo simulation [36]. Figure 3 illustrates the pre-processing of the digital phantoms.

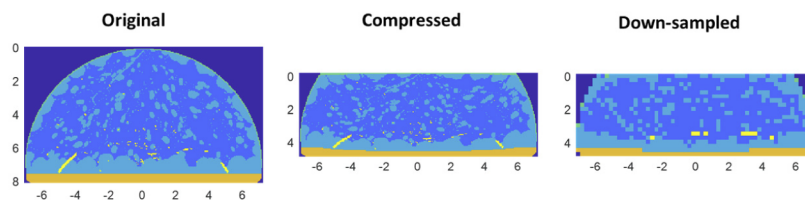


Fig. 3. Pre-processing of digital breast phantoms for Monte Carlo simulation. The digital breast has a fat fraction of 40%.

The values of the optical absorption coefficient (μ_a) and reduced scattering coefficient (μ_s') assigned to each tissue component were calculated and are shown in Table 1. Spherical lesions of both high contrast (with oxygenated hemoglobin (HbO) and de-oxygenated hemoglobin (HbR) equal to the average values from the malignant group [15]) and low contrast (with HbO and HbR equal to the average values from the benign group [15]) were inserted in the lesion side breast. To calculate the normalized perturbation, diffuse reflectance was measured from both the lesion-side and reference-side breasts. The volume surrounding the compressed digital breasts was filled with air, which leaves space for phantom translation and rotation, achieving a total volume of 63

$\times 63 \times 20$ grids. One hundred million photons were used for simulation, and the time required to generate one set of forward data (9×10 measurements) was less than two minutes using GPU-accelerated code [36] on a NVIDIA Quadro P2000. To better simulate tissue heterogeneity mismatch due to probe positioning in clinical settings, tissue components of the reference side and the lesion side were assumed to be the same, but the reference side was translated by 0.5 cm in the +x, -x, +y, or -y-direction, or was rotated 10° around the x, y, or z-axis, achieving a total of seven background tissue mismatch conditions. For each target contrast (high contrast with target tHb = $98.5 \mu\text{M}$, and low contrast with target tHb = $63.2 \mu\text{M}$), we tested three target radii (0.5, 0.75, and 1 cm), three target depths (1.5, 2.0, and 2.5 cm), four breast densities, and seven different background mismatch conditions, resulting in a total of 252 cases.

Table 1. Tissue optical properties used in Monte Carlo simulation

Tissue [refs]	HbO (μM)	HbR (μM)	tHb (μM)	μ_a (cm^{-1})				μ_s' (cm^{-1})			
				at 740 nm	at 780 nm	at 808 nm	at 830 nm	at 740 nm	at 780 nm	at 808 nm	at 830 nm
Fat [30,37]	13.8	4.8	18.7	0.028	0.036	0.037	0.041	7.96	7.56	7.31	7.13
Glandular tissue [30,37]	19.0	6.5	25.4	0.036	0.047	0.048	0.053	9.00	8.56	8.28	8.07
Skin [38]				0.45	0.44	0.43	0.40	17.23	17.03	16.90	16.80
Muscle [39]				0.15	0.17	0.18	0.12	9.98	9.16	8.65	8.28
Blood vessel [37]	204.8	5.2	210.0	3.47	5.39	6.39	7.25	0.00	0.00	0.00	0.00
Low contrast lesion [15,37]	38.0	25.2	63.2	0.10	0.12	0.12	0.13	11.85	11.28	10.92	10.65
High contrast lesion [15,37]	65.0	33.5	98.5	0.15	0.19	0.18	0.20	11.85	11.28	10.92	10.65

2.3. Optode coupling mismatch and chest wall mismatch model

To further demonstrate the generalizability of the edge artifact reduction algorithm, an optode coupling and a chest wall mismatch model were also used to generate edge artifacts. As shown in Fig. 4(a), to create optode coupling mismatch, a homogeneous (fat) digital breast was used to minimize the effects of tissue heterogeneity and focus on the effects of optode coupling errors. We assumed that one side of the breast was in good contact, but one detector located at $(-4 \text{ cm}, 0, 0)$ on the other side had an airgap [20]. To simulate breasts with chest walls, two-layer digital breasts were used, with the top layer consisting of fat and the bottom layer muscle. The chest wall was at 2.5 cm deep. To create chest wall angle mismatch, the chest wall on one side was tilted 10° and the chest wall on the other side was kept flat (Fig. 4(b)). High contrast targets were inserted in the lesion-side breasts. The target had a 0.5 cm radius and a 1 cm depth. The optical properties assigned to the fat and muscle can be found in Table 1.

2.4. Clinical data and ultrasound segmentation using CNN

Patient data from seven benign and seven malignant cases were used to evaluate the artifact reduction algorithm. The study protocol was approved by the local Institutional Review Board and was compliant with the Health Insurance Portability and Accountability Act. The data used in this manuscript were deidentified. This group of fourteen patients had an average age of 47.5 years (± 18.2 years). The means and standard deviations of the lesion radii in the x-direction, the lesion radii in the z-direction, and the depths measured with the US segmentation method described below are $0.76 \pm 0.39 \text{ cm}$, $0.51 \pm 0.19 \text{ cm}$, and $1.32 \pm 0.50 \text{ cm}$, respectively.

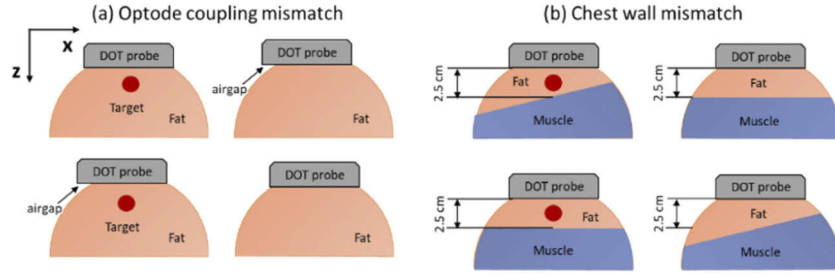


Fig. 4. Illustration of (a) optode coupling mismatch and (b) chest wall mismatch model.

We used a semi-automated CNN to segment breast lesions from the coregistered US images of clinical data and guide the edge artifact reduction. This CNN was developed by Maninis et al. by fine-tuning a CNN model pre-trained on ImageNet, with PASCAL, a well-known standard dataset [40], and it has been successfully applied to breast lesion segmentation in US images to improve DOT reconstruction [16]. After we selected four boundary markers, two in the lateral dimension and two in the depth dimension, the CNN automatically segmented the 2D shape of the lesion. After the lesion was extracted, its centroid in the x - z plane (x_c, z_c) was calculated. Because our US transducer was located in the middle of our DOT probe, we assumed the lesion center to be $(x_c, 0, z_c)$. We then approximated the lesion shape as an ellipsoid centered at $(x_c, 0, z_c)$, with axes in the x -, y -, and z -directions equal to the maximum dimensions in the x -, x - (approximately as y -), and z -directions of the segmented lesion. Then we used an ellipsoid with the same center and long and short axes of three times as long to define the outer boundary of the lesion area. Outside the larger ellipsoid, we can confidently assume any hot spots were edge artifacts. The larger ellipsoid will be used to determine the $\max(\delta\mu_{a,\text{tar}})$ in **Algorithm 1** and quantify the effectiveness of the algorithm applied on background tissue mismatched simulations. The algorithm defined a larger boundary than that of the US-identified lesion for four reasons: 1) DOT has much lower spatial resolution than US; 2) Edges of the breast lesions seen by US are often not clear due to poor contrast; 3) Co-registered US B-Scan images provide lesion size in one spatial dimension, and the size in another dimension is based on an estimate that the breast lesion in general is ellipsoid; and 4) Lesion US images are based on mechanical contrast and optical distribution is based on absorption contrast. Thus, it is not always possible to co-locate both US and optical images. As an example, in Fig. 5(b), the red shaded area is the CNN segmented lesion, with green crosses indicating four markers selected by hand. A yellow diamond marks the centroid of the lesion, the blue line shows the approximated lesion shape, and the white dashed line marks the boundary of the lesion area.

2.5. Image reconstruction

The normalized perturbation U_{sc} was used to reconstruct the unknown optical properties, and it is calculated as the frequency domain measurements from the lesion-side breast normalized to the contralateral reference-side breast:

$$U_{sc}(i) = \frac{U_l(i) - U_r(i)}{U_r(i)}, \quad (1)$$

where $U_l(i)$ and $U_r(i)$ are the lesion-side and the reference-side measurements, respectively. A conjugate gradient algorithm and a dual mesh scheme were employed to solve the DOT inverse problem, which was linearized as the following regularized optimization problem [41]:

$$f(x) = \arg \min_{\delta\mu_a} \left(\|U_{sc} - W\delta\mu_a\|^2 + \frac{\lambda}{2} \|\delta\mu_a\|^2 \right) \quad (2)$$

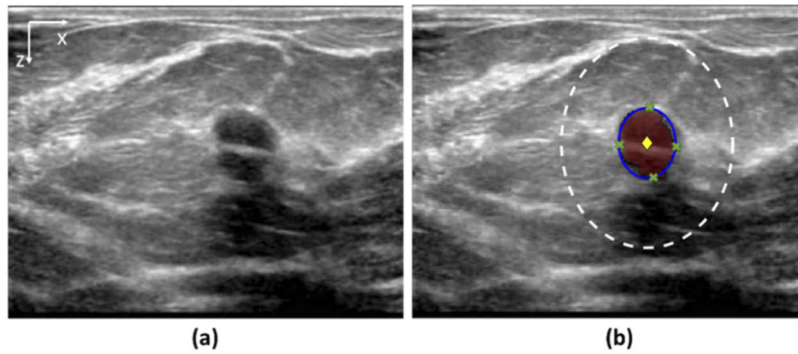


Fig. 5. Illustration of US segmentation. (a) Coregistered US lesion image. (b) US segmentation and information extraction. The red shaded area is the lesion, segmented using CNN. Green crosses are four markers selected by hand. The lesion's centroid is marked with a yellow diamond, the blue line shows the approximate lesion shape, and the white dashed line marks the boundary of the lesion area, an ellipsoid with the same center but three times as large as the lesion.

Here, $\delta\mu_a$ represents the unknown changes in μ_a compared to the reference side, and W is a sensitivity matrix of the semi-infinite homogenous medium. In both simulation and patient experiments, W was calculated using Born approximation based on the optical properties μ_{a0} and μ_{s0} , obtained from the reference side measurements via linear fitting of the log amplitude vs. source-detector distance and phase vs. source-detector distance [42]. λ is a regularization parameter. The μ_a map on the lesion side was obtained by adding the reconstructed $\delta\mu_a$ to the background absorption coefficient, μ_{a0} . A total of seven layers in the depth- or z -direction are reconstructed, with a 0.5-cm spacing between adjacent layers. For better visualization, we show only the top layer of the reconstructed image containing the target.

2.6. Edge artifact reduction and quantification

When there are edge artifacts outside the lesion area, the algorithm removes the portion of the perturbation that contributes to the artifacts ($U_{sc,artifacts}$) from U_{sc} by zeroing out the absorption coefficient of the artifacts ($\delta\mu_{a,artifacts}$) in the original reconstructed absorption map, as formulated in Eq. (3).

$$U_{sc,new} = U_{sc} - U_{sc,artifacts} = U_{sc} - W\delta\mu_{a,artifacts}. \quad (3)$$

Thus, the algorithm finds all the connected components in the reconstructed $\delta\mu_a$ maps and iteratively removes all other components on the edge (artifacts), until there is only one component (the lesion) left, whose centroid is closest to the ultrasound-identified lesion centroid. For details, see **Algorithm 1** and Fig. 6. After edge artifacts were reduced at each wavelength, the reconstructed absorption maps of all four wavelengths were linearly weighted using extinction coefficients obtained from Ref. [39] to obtain the total hemoglobin (tHb) distribution. The Wilcoxon rank sum test, a non-parametric test which is appropriate when comparing two small groups of samples without making a normal distribution assumption, was used to evaluate the statistical significance. A p -value of less than 0.05 was considered statistically significant. All image reconstruction, artifact reduction, and statistical tests were conducted using MATLAB.

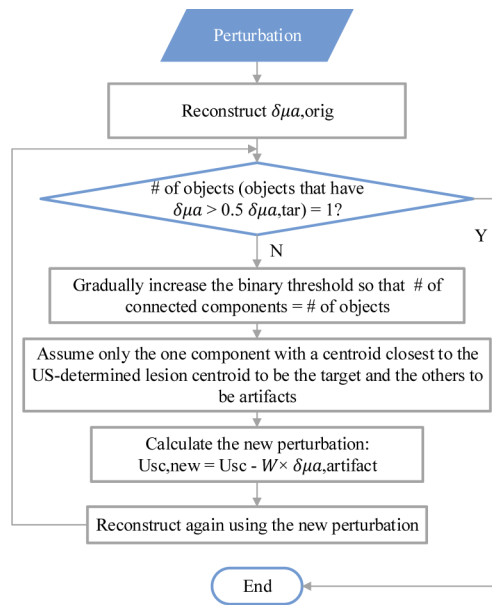


Fig. 6. Flow chart for the edge artifact reduction algorithm.

3. Results

3.1. Simulation results

3.1.1. Background tissue mismatch

Figure 7 illustrates the edge artifact reduction process at 780 nm applied to a high contrast target with a 1 cm radius at a 2 cm depth. In this example, the algorithm identified two connected components in the original reconstructed image, and successfully reduced the one edge artifact on the top right.

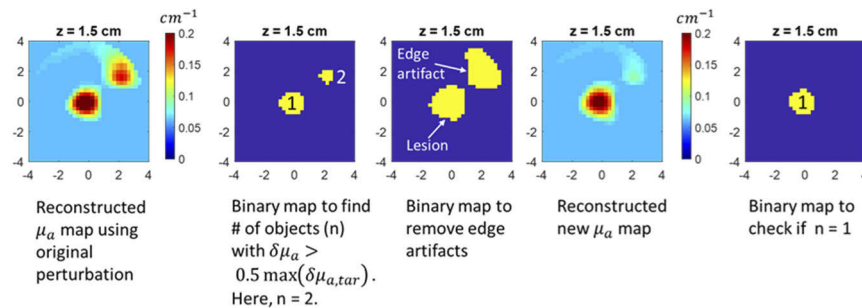


Fig. 7. Illustration of the edge artifact reduction process (780 nm). The breast phantom had a fat fraction of 40%, and the reference side was translated 0.5 cm in the -x direction. The high contrast target, located at depth of 2 cm, had a radius of 1 cm.

Algorithm 1. US segmentation-guided edge artifact reduction algorithm

1. Reconstruct the original absorption map $\delta\mu_{a,orig}$, using U_{sc} according to equation (2).
2. Obtain the centroid of the lesion ($x_c, 0, z_c$) from the coregistered US image for patient data. For simulation, the centroid of the lesion is (0, 0, depth).
3. Find the number of objects n in the reconstructed absorption map:
 - 3.1 Generate a binary image from $\delta\mu_{a,orig}$ using a binary threshold of $0.5\max(\delta\mu_{a,tar})$, where $\max(\delta\mu_{a,tar})$ is the maximum reconstructed $\delta\mu_a$ within the lesion area, shown by the white dashed line in Fig. 5. Here, we deal only with image artifacts that are higher than $0.5\max(\delta\mu_{a,tar})$, since these will be more likely to cause misinterpretation of the reconstructed lesion images.
 - 3.2 Obtain the number of connected objects (n) using connected component analysis, which is the total number of targets and artifacts. The steps of connected component analysis are:
 - 1) Search for the next unlabeled pixel, p .
 - 2) Starting from p , perform a 26-neighbor (faces, edges, or corners) flood-fill to find all pixels connected to p by a path, and give these pixels a new label. This group of pixels forms one connected component.
 - 3) Repeat steps 1) and 2) until all the pixels are labeled.
 The number of objects, n , is the number of unique labels. If $n > 1$, perform the following steps; otherwise end the loop.
4. Reduce edge artifacts:
 - 4.1 Gradually increase the binary threshold from $0.1\max(\delta\mu_{a,tar})$ with a small increment of $1e-4\text{ cm}^{-1}$, until the number of connected components = n in the binary map.
 - 4.2 Find the centroid of each connected component; keep the component with centroid closest to ($x_c, 0, z_c$) and set all other connected component to be 0.
 - 4.3 Calculate the new perturbation using equation (3), and reconstruct a new $\delta\mu_a$.
 - 4.4 Repeat steps 3-4.

Figure 8 shows four examples of reconstructed tHb maps from VICTRE digital breast simulations, before and after artifact reduction. For all breast densities, when the same background tissue is used for the reference side and the lesion side, tissue heterogeneity has less impact on the reconstructed images (Fig. 8(b)). On the other hand, when there is a mismatch in the background tissue between the two sides, it can cause significant image artifacts, appearing as hot spots on the edges of the fine-mesh area of the dual-mesh scheme (Fig. 8(c), enclosed in black circles). The artifact reduction algorithm successfully reduces these edge artifacts and restores the reconstructed target image to be almost comparable to the images without background mismatch. As shown in Table 2, after the artifact reduction algorithm is applied, the maximum reconstructed tHb for all four examples slightly changes (absolute changes $< 3\%$) from its value before artifact reduction.

Table 2. Maximum reconstructed tHb, with and without background tissue mismatch

Fat fraction	No background mismatch, μM	Background mismatch, μM		
		Before artifact reduction	After artifact reduction	Change after artifact reduction
20%	83.8	85.7	87.7	+2.36%
40%	90.8	89.3	87.2	-2.41%
60%	83.4	81.5	82.5	+1.25%
80%	94.6	74.1	75.6	+2.07%

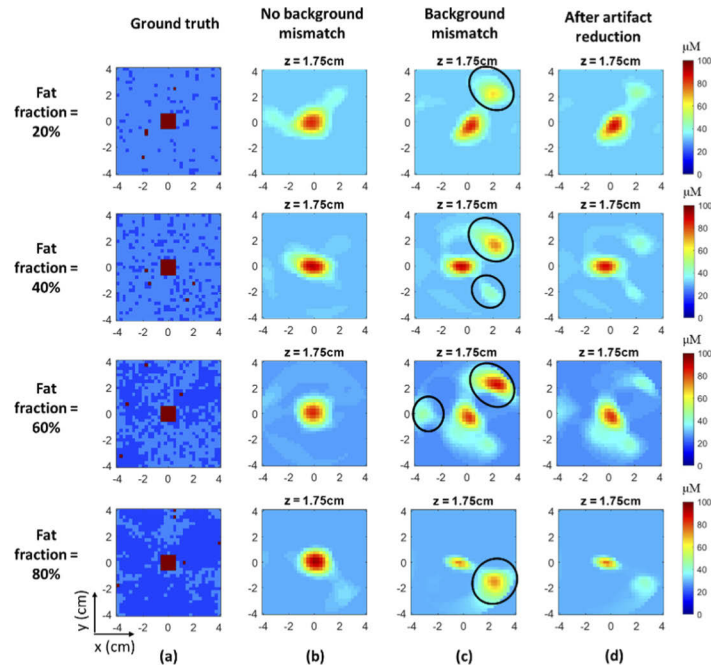


Fig. 8. Total hemoglobin maps of four digital breasts with different densities. (a) Ground truth. (b) Reconstructed tHb maps without background mismatch. (c) and (d), Reconstructed tHb maps with background mismatch, before and after artifact reduction. The target tHb = 98.5 μM , radius = 0.75 cm, and depth = 2 cm. Each image slice measures 8 cm by 8 cm and remains the same for subsequent tHb maps. Image artifacts that have $\delta\mu_a$ larger than $0.5\max(\delta\mu_{a,tar})$ for at least one wavelength are enclosed in black ellipses.

To better demonstrate the performance of the artifact reduction algorithm, Fig. 9 shows box plots of the maximum reconstructed tHb inside and outside the target area for high contrast targets of different sizes and depths, before and after applying the algorithm. As described in Section 2.2, each target comprised four breast densities and seven different background mismatch conditions, making a set of 28 images in all. Each box plot provides statistical information about the images with edge artifacts among this set of 28 images. The target area, defined as a sphere centered on the target, has a radius three times as large as the target's radius. For all images with edge artifacts, the maximum reconstructed tHb inside the target area is in general lower and more spread out for smaller and deeper targets, and the averaged maximum reconstructed tHb outside the target area is roughly the same for all target sizes and depths. The Wilcoxon rank sum test showed that the artifact reduction algorithm successfully reduced the image artifacts outside the target area for all cases ($p < 0.05$), except for targets with a 0.5-cm radius and a 2.5-cm depth. The maximum reconstructed target tHb was not significantly changed for all cases ($p > 0.05$).

Table 3 gives the numbers of images with edge artifacts for each target radius and depth, the ratio of the maximum reconstructed tHb inside the target area over that outside the target area of these images, and statistical test results for both high and low contrast targets. For smaller and deeper targets, more images have edge artifacts, and the ratios of maximum reconstructed tHb inside the target area over that outside the target area are in general smaller, indicating that smaller and deeper targets are more likely to be misinterpreted due to background tissue mismatch. For all cases, the artifact reduction algorithm improves this ratio. The statistical test for low contrast targets had the same conclusions as the high contrast targets: the artifact reduction algorithm significantly reduced the image artifacts outside the target area for all cases

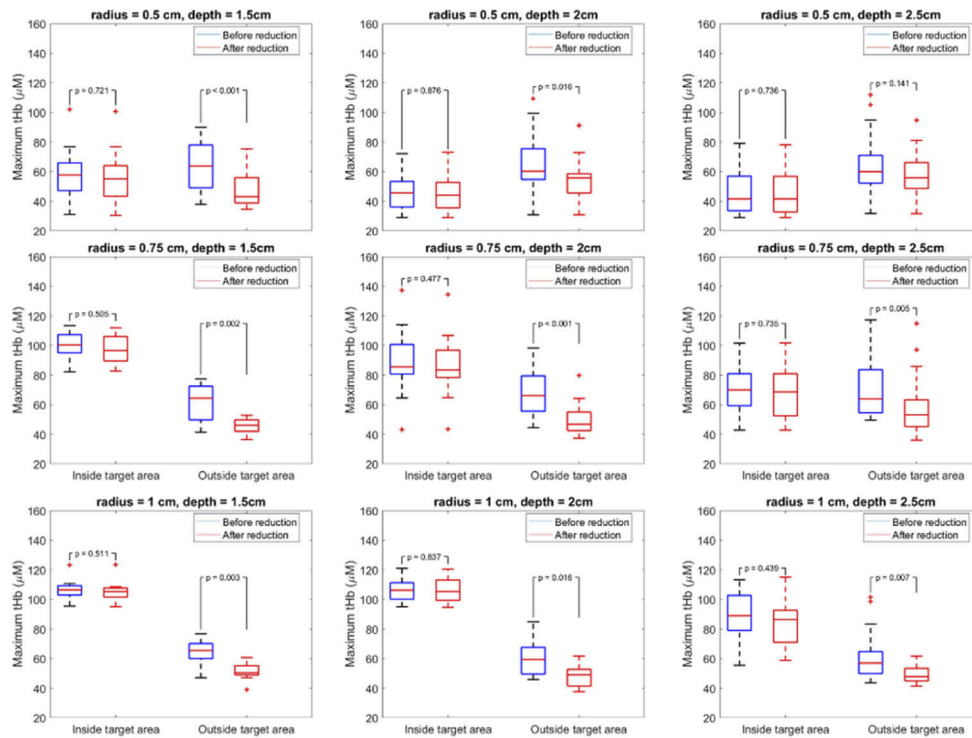


Fig. 9. Box plots of maximum reconstructed tHb inside and outside the target area, from simulations of high contrast targets with different radii and depths, four breast densities, and seven background tissue mismatch conditions, before and after artifact reduction.

except for targets with a 0.5-cm radius and a 2.5-cm depth, and it did not significantly change the maximum reconstructed target tHb for all cases.

3.1.2. Optode coupling and chest wall mismatch

Figure 10 shows reconstructed images with optode coupling mismatch and chest wall mismatch, before and after artifact reduction. In (a) top, one detector on the reference side has an air gap, creating an edge artifact on the left, close to the detector with the air gap. In (a) bottom, one detector on the target side has an air gap, creating two hot spots on the left, close to the two detectors close to the boundary. In (b), the chest wall on the target side or the reference side is tilted 10°, and both conditions induce edge artifacts on the reconstructed image. The algorithm successfully reduced the edge artifacts in all cases, without significantly changing the target tHb (absolute change < 4%). The restored images are comparable to the results using the optode coupling correction [20], which removes measurements corresponding to the optode(s) with large modeling errors, or to the chest wall correction method [17], which scales the original perturbation using simulations based on the chest wall geometry and the fitted optical properties of the tissue and the chest wall.

3.2. Clinical examples

Figure 11 shows representative tHb maps from two patients, one with a benign lesion (a) and another with a malignant tumor (b), before and after the artifact reduction algorithm was applied. In Fig. 11(a) left, the coregistered US image of a 40-year-old patient with a benign fibroadenoma, the 1.1-cm deep lesion measures 1.5 cm and 0.7 cm in the x- and z-directions, respectively.

Table 3. Number of images with artifacts (at least one wavelength), ratio of maximum reconstructed tHb inside the target area over that outside the target area, and statistical test results for digital breast simulations

	Target radius (cm)	Target depth (cm)	Number of images with artifacts / total number of images	Target maximum tHb / artifact maximum tHb (mean \pm std, in μ M)		p-value: before vs. after	
				Before	After	Target	Artifact
High contrast	0.5	1.5	26 / 28	0.907 \pm 0.253	1.204 \pm 0.338	0.721	< 0.001
	0.5	2	27 / 28	0.729 \pm 0.174	0.838 \pm 0.188	0.876	0.016
	0.5	2.5	27 / 28	0.763 \pm 0.218	0.823 \pm 0.213	0.736	0.141
	0.75	1.5	14 / 28	1.658 \pm 0.305	2.130 \pm 0.238	0.505	0.002
	0.75	2	24 / 28	1.360 \pm 0.404	1.799 \pm 0.513	0.477	< 0.001
	0.75	2.5	26 / 28	1.056 \pm 0.245	1.259 \pm 0.338	0.735	0.005
	1	1.5	11 / 28	1.670 \pm 0.276	2.071 \pm 0.280	0.511	0.003
	1	2	13 / 28	1.835 \pm 0.324	2.256 \pm 0.308	0.837	0.016
	1	2.5	24 / 28	1.529 \pm 0.431	1.738 \pm 0.427	0.439	0.007
Low contrast	0.5	1.5	26 / 28	0.780 \pm 0.207	1.005 \pm 0.286	0.641	< 0.001
	0.5	2	26 / 28	0.681 \pm 0.157	0.806 \pm 0.184	0.749	0.008
	0.5	2.5	27 / 28	0.730 \pm 0.196	0.785 \pm 0.186	0.829	0.210
	0.75	1.5	23 / 28	1.462 \pm 0.321	1.872 \pm 0.346	0.429	< 0.001
	0.75	2	25 / 28	1.082 \pm 0.276	1.376 \pm 0.428	0.388	< 0.001
	0.75	2.5	25 / 28	0.966 \pm 0.168	1.076 \pm 0.230	0.308	0.033
	1	1.5	17 / 28	1.559 \pm 0.272	1.965 \pm 0.235	0.389	< 0.001
	1	2	23 / 28	1.473 \pm 0.261	1.908 \pm 0.323	0.410	< 0.001
	1	2.5	25 / 28	1.228 \pm 0.246	1.399 \pm 0.326	0.244	0.002

The proposed algorithm successfully reduced two hot spots on the edge, and the resulting tHb map correctly shows the lesion in the corresponding region in the US image. The maximum reconstructed tHb decreased by 45.1%. In (b), a 51-year-old patient with an invasive mammary ductal carcinoma, the lesion has a depth of 1.2 cm, and it measures 1.2 cm and 0.8 cm in the x- and z-directions, respectively. The reconstructed hemoglobin map has one hot spot outside the lesion area. After the algorithm has processed the data, the edge artifacts outside the tumor region shown in the US image are reduced, and the maximum reconstructed tHb has decreased by 14.9%.

Figure 12 shows box plots of the maximum reconstructed tHb for seven benign and seven malignant cases, before and after artifact reduction. The algorithm decreased the maximum reconstructed tHb for both malignant and benign lesions. However, the ratio of the averaged maximum reconstructed tHb of the malignant lesions ($n=7$) to that of benign lesions ($n=7$) was improved (from 1.55 to 1.91), indicating an increase in optical contrast distinguishing between malignant and benign lesions. The Wilcoxon rank sum test showed that after artifact reduction, the malignant group and benign group were significantly different in terms of the maximum reconstructed tHb ($p < 0.005$), while before artifact reduction they are not ($p > 0.05$).

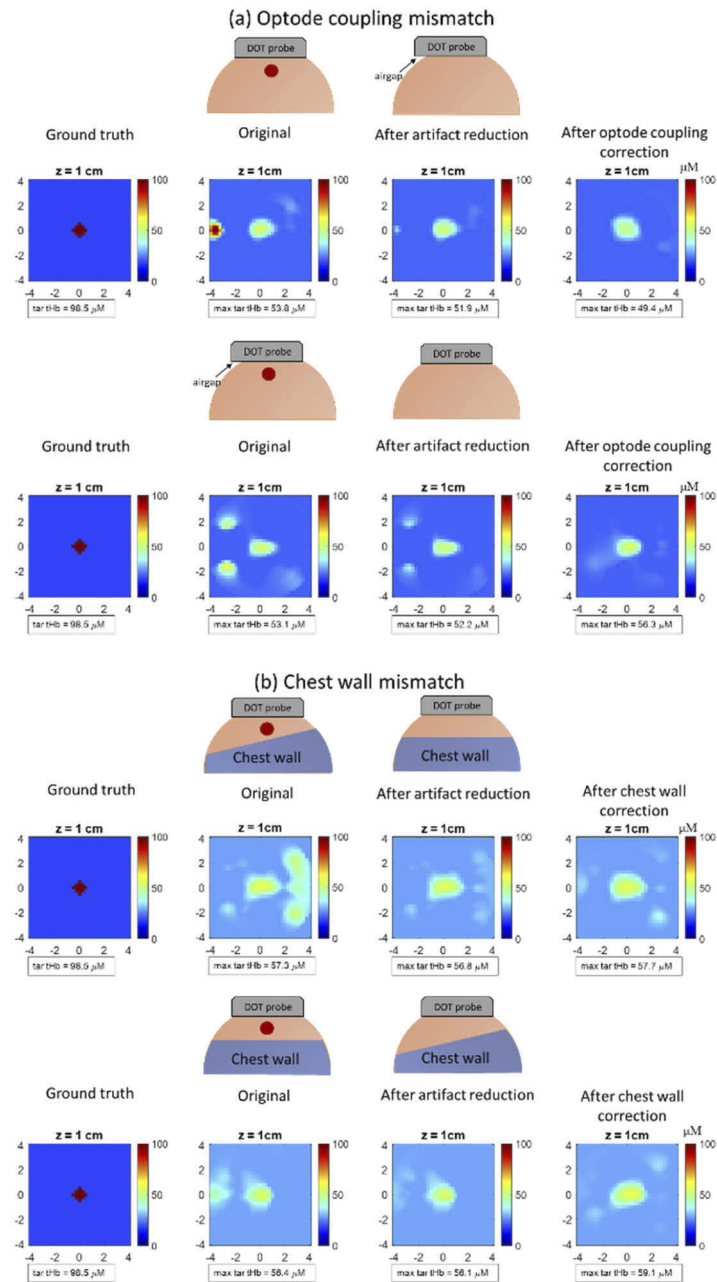


Fig. 10. (a) Ground truth tHb maps and reconstructed tHb maps with optode coupling mismatch, before and after artifact reduction and after optode coupling correction [20]. One detector (-4 cm, 0 , 0) had an airgap. (b) Ground truth tHb maps and reconstructed tHb maps with chest wall mismatch, before and after artifact reduction, and after chest wall correction [17]. The chest wall has a depth of 2.5 cm and is tilted 10° . The high contrast lesion has a radius of 0.5 cm and a depth of 1 cm.

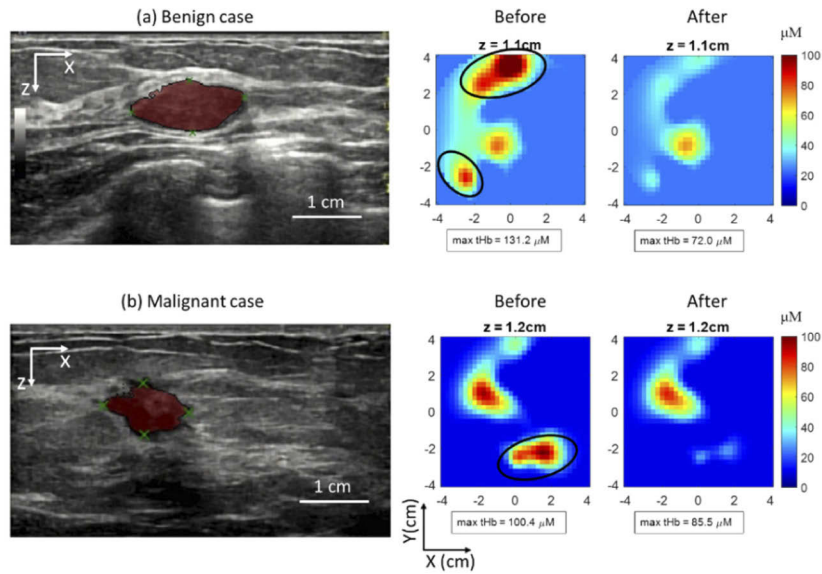


Fig. 11. Coregistered US images, original reconstructed tHb maps, and tHb maps after applying the artifact reduction algorithm for (a) a benign fibroadenoma and (b) a malignant tumor. Color bars indicate the tHb level, in units of μM . Image artifacts that have $\delta\mu_a$ larger than $0.5\max(\delta\mu_{a,tar})$ for at least one wavelength are enclosed in black ellipses.

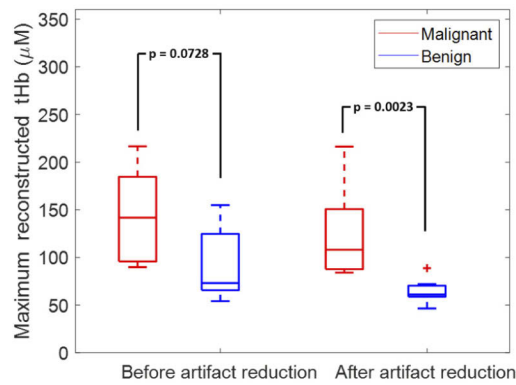


Fig. 12. Box plots of maximum reconstructed tHb for seven benign and seven malignant cases, before and after artifact reduction.

4. Discussion and summary

In this work, we introduced an iterative edge artifact reduction algorithm based on connected component analysis. The proposed artifact-reduction algorithm uses the structural prior and modifies the data to produce images that agree with the structural prior. Unlike previous correction algorithms that deal with optode coupling or the chest wall [17, 24], it does not make assumptions about the sources that caused the image artifacts. The algorithm successfully reduced hot spots appearing in non-lesion regions (edge artifacts), which, in patient experiments, can arise from mismatches of the chest wall position underneath the breast tissue [18], poor probe-tissue contact [20], and tissue heterogeneity. It may also reduce the high optical contrast at the edges of healthy tissue, such as dense glands, biopsy sites, or scar due to lumpectomy, improving the lesion-to-background tissue contrast for more accurate lesion diagnosis and characterization. We first demonstrated the performance of the algorithm using Monte Carlo simulations on VICTRE digital breast phantoms. Realistic and flexible VICTRE breast phantoms of different fat fractions were used in DOT simulations for the first time to account for tissue heterogeneity. When the same background tissue was used for the reference side and the lesion side, fewer edge artifacts were observed in the reconstructed images, because our “normalized perturbation” approach reduced the impact of tissue heterogeneity. However, when there was a tissue mismatch between the two sides, edge artifacts were often generated, but they were successfully reduced by our algorithm. The proposed algorithm can also reduce edge artifacts caused by optode coupling mismatch or chest wall mismatch, and it produces images comparable to those using the previous correction algorithms developed by our group that deals with a single mismatch error [17, 20]. Compared to the optode coupling correction method in [20], the proposed algorithm does not decrease the number of the measurements. Compared to the chest wall correction method in [17], there is no need for accurate estimation of the chest wall geometry, time-consuming fitting of the optical properties of breast tissue and the chest wall, or additional forward simulations.

We also tested the algorithm using breast patient data with image artifacts outside the lesion areas. Interestingly, compared to all the patient data we had, we noticed from US B-scan images that this group of patients with image artifacts was more likely to have dense breasts. Denser breasts make it harder for the DOT probe to make good contact and harder to match the positions of the two sides. A CNN-based US segmentation algorithm was incorporated to extract lesion size and location information and guide the artifact reduction. Other segmentation approaches, such as a clustering-based approach [43] or a watershed-based approach [44], can also be used to obtain lesion information from breast lesion US images. Combined with the US segmentation algorithm, this edge artifact reduction algorithm showed success in removing edge artifacts and improving tHb contrast between malignant and benign lesions for clinical data. Although the edge artifacts in this study were generated by probe positioning differences between the reference and lesion side measurements, unwanted artifacts outside the target region can also occur if optode coupling errors exist in a single DOT measurement [21–24]. Our approach may also be applicable to such cases.

Our studies have several limitations. First, although different densities of digital breast phantoms were included in this study, breast types were still limited in terms of size, shape, chest wall depth and angle, and many other variables. Second, due to modest available computational power, the Monte Carlo grids used in this study were relatively large ($2.5 \times 2.5 \times 2.5 \text{ mm}^3$), which reduced the level of detail in the digital breasts and decreased the reality of the simulations. For example, after numerically compressing and down-sampling the digital breast, the skin layer, with only a 1 mm thickness, was no longer present on the top plane of the breast. In the Monte Carlo simulation, photon noise also increased inaccuracy in the forward calculation. More accurate and realistic simulations can be achieved with greater computational resources. Further, we assumed the background tissue had the same component distribution on the lesion side and the

reference side, and that the mismatch was generated from probe positioning. In patients, the tissue components of the two side breasts can be different.

The algorithm demonstrated its effectiveness in reducing artifacts on the edges. It slightly changes the reconstructed lesion μ_a and does not alter the lesion position or shape. When a strong tissue background mismatch causes a large shift in the target position or causes the target and the artifact to merge into one component (conditions that are more frequently observed with smaller and deeper lesions), this algorithm does not work well. Also, in rare cases, there may be multiple lesions within the field of view of our DOT probe, in which case co-registered US can identify the number of lesions and the lesion locations. The proposed algorithm can be modified accordingly to keep the number of desired connected components as the number of the lesions. In addition, this algorithm is not applicable to large lesions for which the defined lesion area (three times the lesion size) covers the whole reconstructed image volume. Also, cancers larger than 4 to 5 cm may contain necrotic cores due to the aggressive growth of the cancer cells. In comparison with the tHb map of a single-connected blob, the tHb value of this necrotic core is lower than the peripheral tHb value, and the algorithm does not apply to such cases. However, as shown in Fig. 10 and Table 3, the ratios of the averaged reconstructed μ_a inside the target area over that outside the target area are larger for larger targets, indicating they are less influenced by background tissue mismatch. This algorithm also increases the reconstruction time, and the additional time mainly comes from the additional reconstruction process in each iteration. The median number of iterations needed to remove the image artifacts was two, adding approximately twice the original reconstruction time.

In summary, to reduce edge artifacts in US-DOT, we proposed an edge artifact reduction algorithm based on connected component analysis and combined with a US-segmentation CNN. We demonstrated its effectiveness using simulations on VICTRE digital breast phantoms and clinical data. Our approaches have significant implications for reducing image edge artifacts in other modality-guided DOT imaging in general, where the lesion count, size, and position can be reliably obtained through other modalities.

Funding. National Cancer Institute (CA228047).

Acknowledgments. We thank Drs. Mark Anastasio and Seonyeong Park for introducing digital breast phantoms, and Professor James Ballard for reviewing and editing the manuscript.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. B. J. Tromberg, A. Cerussi, N. Shah, M. Compton, A. Durkin, D. Hsiang, J. Butler, and R. Mehta, "Diffuse optics in breast cancer: detecting tumors in pre-menopausal women and monitoring neoadjuvant chemotherapy," *Breast Cancer Res* **7**(6), 279–285 (2005).
2. R. Choe, A. Corlu, K. Lee, T. Durduran, S. D. Konecky, M. Grosicka-Koptyra, S. R. Arridge, B. J. Czerniecki, D. L. Fraker, A. Demichele, B. Chance, M. A. Rosen, and A. G. Yodh, "Diffuse optical tomography of breast cancer during neoadjuvant chemotherapy: A case study with comparison to MRI," *Med. Phys.* **32**(4), 1128–1139 (2005).
3. E. Heffer Pera, V. Schutz, O. Siebold, and H. Fantini, "Near-infrared imaging of the human breast: complementing hemoglobin concentration maps with oxygenation images," *J. Biomed. Opt.* **9**(6), 1152–1160 (2004).
4. Q. Zhu and S. Poplack, "A review of optical breast imaging: Multi-modality systems for breast cancer diagnosis," *Eur. J. Radiol.* **129**(01), 109067 (2020).
5. B. Chance, S. Nioka, J. Zhang, E. F. Conant, E. Hwang, S. Briest, S. G. Orel, M. D. Schnall, and B. J. Czerniecki, "Breast cancer detection based on incremental biochemical and physiological properties of breast cancers: A six-year, two-site study," *Acad. Radiol.* **12**(8), 925–933 (2005).
6. D. R. Leff, O. J. Warren, L. C. Enfield, A. Gibson, T. Athanasiou, D. K. Patten, J. Hebden, G. Z. Yang, and A. Darzi, "Diffuse optical imaging of the healthy and diseased breast: A systematic review," *Breast Cancer Res. Treat.* **108**(1), 9–22 (2008).
7. B. Brooksby, S. Jiang, H. Dehghani, B. W. Pogue, K. D. Paulsen, C. Kogel, M. Doyley, J. B. Weaver, and S. P. Poplack, "Magnetic resonance-guided near-infrared tomography of the breast," *Rev. Sci. Instrum.* **75**(12), 5262–5270 (2004).

8. T. A. R. Hoda Badr, C. L. Carmack, D. A. Kashy, and M. Cristofanilli, "Combined optical imaging and mammography of the healthy breast: optical contrast derived from breast structure and compression," *IEEE Trans. Med. Imaging* **23**(1), 30–42 (2005).
9. Q. Zhang, T. J. Brukilacchio, A. Li, J. J. Stott, T. Chaves, E. Hillman, T. Wu, M. Chorlton, E. Rafferty, R. H. Moore, D. B. Kopans, and D. A. Boas, "Coregistered tomographic x-ray and optical breast imaging: initial results," *J. Biomed. Opt.* **10**(2), 024033 (2005).
10. L. Wang, N. Schwartz, J. M. Cochran, T. Ko, W. Baker, L. He, K. Abramson, and A. Yodh, "Noninvasive measurement of placental hemodynamics with ultrasound-guided fd-diffuse optical spectroscopy," in *Optical Tomography and Spectroscopy* (Optical Society of America, 2020), pp. SW4D-5.
11. T. Ko, L. He, V. C. Kavuri, A. Parthasarathy, V. C. Morano, J. Cochran, W. B. Baker, J. Seigel, A. G. Yodh, D. J. Licht, and N. Schwartz, "Quantifying placental oxygenation using ultrasound-guided frequency-domain near-infrared spectroscopy (FD-NIRS)," *Am. J. Obstet. Gynecol.* **216**(1), S172 (2017).
12. L. Cortese, G. Lo Presti, M. Zanoletti, M. Buttafava, M. Renna, A. D. Mora, A. Pifferi, P. Taroni, A. Tosi, G. Aranda, S. Ruiz, M. Squarcia, F. Hanzu, M. M. Porta, S. Wojtkiewicz, H. Dehghani, M. Weigel, S. De Fraguier, B. Rosinski, and T. Durduran, "LUCA Device : a multi-wavelength time-resolved spectroscopy and diffuse correlation spectroscopy device with an integrated clinical ultrasound module / probe," in *Optical Tomography and Spectroscopy* (Optical Society of America, 2020), pp. SW4D-7.
13. C. Lindner, M. Mora, P. Farzam, M. Squarcia, J. Johansson, U. M. Weigel, I. Halperin, F. A. Hanzu, and T. Durduran, "Diffuse optical characterization of the healthy human thyroid tissue and two pathological case studies," *PLoS One* **11**(1), e0147851 (2016).
14. Q. Zhu, E. B. Cronin, A. A. Currier, H. S. Vine, M. Huang, N. Chen, and C. Xu, "Benign versus malignant breast masses: optical differentiation with us-guided optical imaging reconstruction," *Breast Dis.* **237**(1), 57–66 (2005).
15. Q. Zhu, A. Ricci, P. Hegde, M. Kane, E. Cronin, A. Merkulov, Y. Xu, B. Tavakoli, and S. Tannenbaum, "Assessment of functional differences in malignant and benign breast lesions and improvement of diagnostic accuracy by using us-guided diffuse optical tomography in conjunction with conventional us1," *Radiology* **280**(2), 387–397 (2016).
16. M. Zhang, K. M. S. Uddin, S. Li, and Q. Zhu, "Target depth-regularized reconstruction in diffuse optical tomography using ultrasound segmentation as prior information," *Biomed. Opt. Express* **11**(6), 3331 (2020).
17. Y. Ardehshirpour, M. Huang, and Q. Zhu, "An optical tomography method that accounts for a tilted chest-wall in breast imaging," in *Proc. SPIE, Optical Tomography and Spectroscopy of Tissue VIII (2009)*, 7174(February 2009), p. 71740U.
18. Y. Ardehshirpour, M. Huang, and Q. Zhu, "Effect of the chest wall on breast lesion reconstruction," *J. Biomed. Opt.* **14**(4), 044005 (2009).
19. Y. Ardehshirpour, N. Biswal, A. Aguirre, and Q. Zhu, "Artifact reduction method in ultrasound-guided diffuse optical tomography using exogenous contrast agents," *J. Biomed. Opt.* **16**(4), 046015 (2011).
20. S. Li, K. Huang, M. Zhang, K. M. S. Uddin, and Q. Zhu, "Effect and correction of optode coupling errors in breast imaging using diffuse optical tomography," *Biomed. Opt. Express* **12**(2), 689–704 (2021).
21. D. Boas, T. Gaudette, and S. Arridge, "Simultaneous imaging and optode calibration with diffuse optical tomography," *Opt. Express* **8**(5), 263–270 (2001).
22. M. Schweiger, I. Nissila, D. A. Boas, and S. R. Arridge, "Image reconstruction in optical tomography in the presence of coupling errors," *Appl. Opt.* **46**(14), 2743–2756 (2007).
23. R. Fukuzawa, S. Okawa, S. Matsuhashi, T. Kusaka, Y. Tanikawa, Y. Hoshi, F. Gao, and Y. Yamada, "Reduction of image artifacts induced by change in the optode coupling in time-resolved diffuse optical tomography," *J. Biomed. Opt.* **16**(11), 116022 (2011).
24. M. Mozumder, T. Tarvainen, S. R. Arridge, J. Kaipio, and V. Kolehmainen, "Compensation of optode position and sensitivity errors in diffuse optical tomography," *Biomed. Opt. Express* **1**(10), BM3A.76 (2014).
25. K. M. S. Uddin and Q. Zhu, "Reducing image artifact in diffuse optical tomography by iterative perturbation correction based on multiwavelength measurements," *J. Biomed. Opt.* **24**(05), 1 (2019).
26. A. A. M. Al-shammaa and H. R. Mohamed, "Extraction of connected components Skin pemphigus diseases image edge detection by Morphological operations," *Int. J. Comput. Appl.* **46**(18), 7–13 (2012).
27. A. H. Al-fayadh, H. R. Mohamed, and R. S. Al-Shimsah, "CT angiography image segmentation by mean shift algorithm and contour with connected components image," *Int. J. Sci. Eng. Res.* **3**(8), 1–5 (2012).
28. V. Ntziachristos, A. H. Hielscher, A. G. Yodh, and B. Chance, "Diffuse optical tomography of highly heterogeneous media," *IEEE Trans. Med. Imaging* **20**(6), 470–478 (2001).
29. Y. Lou, W. Zhou, T. P. Matthews, C. M. Appleton, and M. A. Anastasio, "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *J. Biomed. Opt.* **22**(4), 041015 (2017).
30. B. Deng, D. H. Brooks, D. A. Boas, M. Lundqvist, and Q. Fang, "Characterization of structural-prior guided optical tomography using realistic breast models derived from dual-energy x-ray mammography," *Biomed. Opt. Express* **6**(7), 2366 (2015).
31. B. Deng, Q. Fang, D. A. Boas, and S. A. Carp, "Impact of errors in experimental parameters on reconstructed breast images using diffuse optical tomography," *Biomed. Opt. Express* **9**(3), 1130 (2018).
32. A. Badano, C. G. Graff, A. Badal, D. Sharma, R. Zeng, F. W. Samuelson, S. J. Glick, and K. J. Myers, "Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in Silico imaging trial," *JAMA Netw open* **1**(7), e185474 (2018).

33. G. Mettievier, A. Sarno, J. M. Boone, K. Bliznakova, F. di Franco, and P. Russo, "Virtual clinical trials in 3D and 2D breast imaging with digital phantoms derived from clinical breast CT scans," *Proc. SPIE, Med. Imaging 2020 Phys. Med. Imaging* 11312(March), 1131259 (2020).
34. S. Park, U. Villa, R. Su, A. Oraevsky, F. J. Brooks, and M. A. Anastasio, "Realistic three-dimensional optoacoustic tomography imaging trials using the VICTRE breast phantom of FDA (Conference Presentation)," in *Proc. SPIE, Photons Plus Ultrasound: Imaging and Sensing 2020* (2020), p. 112401H.
35. S. Park, A. A. Oraevsky, R. Su, and M. A. Anastasio, "Compensation for non-uniform illumination and optical fluence attenuation in three-dimensional optoacoustic tomography of the breast," in *Proc. SPIE 10878, Photons Plus Ultrasound: Imaging and Sensing 2019* (2019), (March 2019), p. 108784X.
36. Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Express* **17**(22), 20178 (2009).
37. S. Prahl, "Optical absorption of hemoglobin," omlc.ogi.edu/spectra/hemoglobin.
38. A. N. Bashkatov, E. A. Genina, V. I. Kochubey, and V. V. Tuchin, "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000nm," *J. Phys. D: Appl. Phys.* **38**(15), 2543–2555 (2005).
39. M. Cope, "The development of a near infrared spectroscopy system and its application for non invasive monitoring of cerebral blood and tissue oxygenation in the newborn infants," *Diss. Univ. London* 342 (1991).
40. K. K. Maninis, S. Caelles, J. Pont-Tuset, and L. Van Gool, "Deep extreme cut: from extreme points to object segmentation," in *Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 616–625.
41. C. Xu, H. Vavadi, A. Merkulov, H. Li, M. Erfanzadeh, A. Mostafa, Y. Gong, H. Salehi, S. Tannenbaum, and Q. Zhu, "Ultrasound-guided diffuse optical tomography for predicting and monitoring neoadjuvant chemotherapy of breast cancers: Recent progress," *Ultrason. Imaging* **38**(1), 5–18 (2016).
42. N. G. Chen, P. Guo, S. Yan, D. Piao, and Q. Zhu, "Simultaneous near-infrared diffusive light and ultrasound imaging," *Appl. Opt.* **40**(34), 6367 (2001).
43. J. Shan, H. D. Cheng, and Y. Wang, "A novel segmentation method for breast ultrasound images based on neutrosophic I-means clustering," *Med. Phys.* **39**(9), 5669–5682 (2012).
44. W. Gómez, L. Leija, A. V. Alvarenga, A. F. C. Infantosi, and W. C. A. Pereira, "Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation," *Med. Phys.* **37**(1), 82–95 (2009).