Biomedical Optics EXPRESS

Calibration and optimization of 3D digital breast tomosynthesis guided near infrared spectral tomography

Kelly E. Michaelsen,[1,*] Venkataramanan Krishnaswamy,[1] Linxi Shi,[2,3] Srinivasan Vedantham,[2] Steven P. Poplack,[1,4] Andrew Karellas,[2] Brian W. Pogue,[1] and Keith D. Paulsen[1]

[1]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA
[2]Department of Radiology, University of Massachusetts Medical School, Worcester, MA 01655, USA
[3]Currently at School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA
[4]Currently at Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO 63110, USA
[*]Kelly.e.michaelsen@dartmouth.edu

Abstract: Calibration of a three-dimensional multimodal digital breast tomosynthesis (DBT) x-ray and non-fiber based near infrared spectral tomography (NIRST) system is challenging but essential for clinical studies. Phantom imaging results yielded linear contrast recovery of total hemoglobin (HbT) concentration for cylindrical inclusions of 15 mm, 10 mm and 7 mm with a 3.5% decrease in the HbT estimate for each 1 cm increase in inclusion depth. A clinical exam of a patient's breast containing both benign and malignant lesions was successfully imaged, with greater HbT was found in the malignancy relative to the benign abnormality and fibroglandular regions (11 μM vs. 9.5 μM). Tools developed improved imaging system characterization and optimization of signal quality, which will ultimately improve patient selection and subsequent clinical trial results.

OCIS codes: (120.3890) Medical optics instrumentation; (170.0110) Imaging systems; (170.3830) Mammography

References and links

1. Q. Fang, J. Selb, S. A. Carp, G. Boverman, E. L. Miller, D. H. Brooks, R. H. Moore, D. B. Kopans, and D. A. Boas, "Combined optical and X-ray tomosynthesis breast imaging," Radiology **258**(1), 89–97 (2011).
2. B. J. Tromberg, N. Shah, R. Lanning, A. Cerussi, J. Espinoza, T. Pham, L. Svaasand, and J. Butler, "Non-Invasive In Vivo Characterization of Breast Tumors Using Photon Migration Spectroscopy," Neoplasia **2**(1-2), 26–40 (2000).
3. S. P. Poplack, T. D. Tosteson, W. A. Wells, B. W. Pogue, P. M. Meaney, A. Hartov, C. A. Kogel, S. K. Soho, J. J. Gibson, and K. D. Paulsen, "Electromagnetic breast imaging: results of a pilot study in women with abnormal mammograms," Radiology **243**(2), 350–359 (2007).
4. M. G. Pakalniskis, W. A. Wells, M. C. Schwab, H. M. Froehlich, S. Jiang, Z. Li, T. D. Tosteson, S. P. Poplack, P. A. Kaufman, B. W. Pogue, and K. D. Paulsen, "Tumor angiogenesis change estimated by using diffuse optical spectroscopic tomography: demonstrated correlation in women undergoing neoadjuvant chemotherapy for invasive breast cancer?" Radiology **259**(2), 365–374 (2011).
5. X. Intes, "Time-domain optical mammography SoftScan: initial results," Acad. Radiol. **12**(8), 934–947 (2005).
6. M. A. Helvie, "Digital mammography imaging: breast tomosynthesis and advanced applications," Radiol. Clin. North Am. **48**(5), 917–929 (2010).
7. Q. Zhu, N. Chen, and S. H. Kurtzman, "Imaging tumor angiogenesis by use of combined near-infrared diffusive light and ultrasound," Opt. Lett. **28**(5), 337–339 (2003).
8. N. G. Chen, P. Guo, S. Yan, D. Piao, and Q. Zhu, "Simultaneous near-infrared diffusive light and ultrasound imaging," Appl. Opt. **40**(34), 6367–6380 (2001).
9. V. Ntziachristos, X. Ma, and B. Chance, "Time-correlated single photon counting imager for simultaneous magnetic resonance and near-infrared mammography," Rev. Sci. Instrum. **69**(12), 4221–4233 (1998).
10. V. Ntziachristos, A. G. Yodh, M. D. Schnall, and B. Chance, "MRI-guided diffuse optical spectroscopy of malignant and benign breast lesions," Neoplasia **4**(4), 347–354 (2002).

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

11. B. A. Brooksby, H. Dehghani, B. W. Pogue, and K. D. Paulsen, "Near-infrared (NIR) tomography breast image reconstruction with a priori structural information from MRI: algorithm development for reconstructing heterogeneities," IEEE J. Sel. Top. Quantum Electron. **9**(2), 199–209 (2003).
12. M. A. Mastanduno, S. Jiang, R. DiFlorio-Alexander, B. W. Pogue, and K. D. Paulsen, "Remote positioning optical breast magnetic resonance coil for slice-selection during image-guided near-infrared spectroscopy of breast cancer," J. Biomed. Opt. **16**(6), 066001 (2011).
13. V. Krishnaswamy, K. E. Michaelsen, B. W. Pogue, S. P. Poplack, I. Shaw, K. Defrietas, K. Brooks, and K. D. Paulsen, "A digital x-ray tomosynthesis coupled near infrared spectral tomography system for dual-modality breast imaging," Opt. Express **20**(17), 19125–19136 (2012).
14. K. Michaelsen, V. Krishnaswamy, B. W. Pogue, K. Brooks, K. Defreitas, I. Shaw, S. P. Poplack, and K. D. Paulsen, "Characterization of materials for optimal near-infrared and x-ray imaging of the breast," Biomed. Opt. Express **3**(9), 2078–2086 (2012).
15. A. Cerussi, N. Shah, D. Hsiang, A. Durkin, J. Butler, and B. J. Tromberg, "In vivo absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy," J. Biomed. Opt. **11**(4), 044005 (2006).
16. J.-P. Bouchard, I. Noiseux, and O. Mermut, "Challenges in manufacturing optical tissue phantoms: an industrial perspective," 2012, vol. 8229, pp. 822916–822916–3.
17. R. E. Hendrick, E. D. Pisano, A. Averbukh, C. Moran, E. A. Berns, M. J. Yaffe, B. Herman, S. Acharyya, and C. Gatsonis, "Comparison of acquisition parameters and breast dose in digital mammography and screen-film mammography in the American College of Radiology Imaging Network digital mammographic imaging screening trial," AJR Am. J. Roentgenol. **194**(2), 362–369 (2010).
18. M. Jermyn, H. Ghadyani, M. A. Mastanduno, W. Turner, S. C. Davis, H. Dehghani, and B. W. Pogue, "Fast segmentation and high-quality three-dimensional volume mesh creation from medical images for diffuse optical tomography," J. Biomed. Opt. **18**(8), 086007 (2013).
19. H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, and K. D. Paulsen, "Near infrared optical tomography using NIRFAST: Algorithm for numerical model and image reconstruction," Commun. Numer. Methods Eng. **25**(6), 711–732 (2009).
20. S. Vedantham, L. Shi, A. Karellas, K. E. Michaelsen, V. Krishnaswamy, B. W. Pogue, and K. D. Paulsen, "Semi-automated Segmentation and Classification of Digital Breast Tomosynthesis Reconstructed Images," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Conf.*, **2011**, pp. 6188–6191, 2011.
21. S. Vedantham, L. Shi, K. E. Michaelsen, V. Krishnaswamy, B. W. Pogue, S. P. Poplack, A. Karellas, and K. D. Paulsen, "Digital breast tomosynthesis guided near infrared spectroscopy: volumetric estimates of fibroglandular fraction and breast density from tomosynthesis reconstructions," Biomed. Phys. Eng. Express **1**(4), 045202 (2015).
22. A. E. Cerussi, R. Warren, B. Hill, D. Roblyer, A. Leproux, A. F. Durkin, T. D. O'Sullivan, S. Keene, H. Haghany, T. Quang, W. M. Mantulin, and B. J. Tromberg, "Tissue phantoms in multicenter clinical trials for diffuse optical technologies," Biomed. Opt. Express **3**(5), 966–971 (2012).
23. S. Srinivasan, B. W. Pogue, B. Brooksby, S. Jiang, H. Dehghani, C. Kogel, W. A. Wells, S. P. Poplack, and K. D. Paulsen, "Near-infrared characterization of breast tumors in vivo using spectrally-constrained reconstruction," Technol. Cancer Res. Treat. **4**(5), 513–526 (2005).
24. B. P. Schneider and K. D. Miller, "Angiogenesis of Breast Cancer," J. Clin. Oncol. **23**(8), 1782–1790 (2005).
25. H. S. Mun, H. H. Kim, H. J. Shin, J. H. Cha, P. L. Ruppel, H. Y. Oh, and E. Y. Chae, "Assessment of extent of breast cancer: comparison between digital breast tomosynthesis and full-field digital mammography," Clin. Radiol. **68**(12), 1254–1259 (2013).
26. A. Luparia, G. Mariscotti, M. Durando, S. Ciatto, D. Bosco, P. P. Campanino, I. Castellano, A. Sapino, and G. Gandini, "Accuracy of tumour size assessment in the preoperative staging of breast cancer: comparison of digital mammography, tomosynthesis, ultrasound and MRI," Radiol. Med. (Torino) **118**(7), 1119–1136 (2013).
27. Y.-H. Huang, J.-H. Chen, Y.-C. Chang, C.-S. Huang, W. K. Moon, W.-J. Kuo, K.-J. Lai, and R.-F. Chang, "Diagnosis of solid breast tumors using vessel analysis in three-dimensional power Doppler ultrasound images," J. Digit. Imaging **26**(4), 731–739 (2013).
28. S. Ciatto, N. Houssami, D. Bernardi, F. Caumo, M. Pellegrini, S. Brunelli, P. Tuttobene, P. Bricolo, C. Fantò, M. Valentini, S. Montemezzi, and P. Macaskill, "Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study," Lancet Oncol. **14**(7), 583–589 (2013).

1. Introduction

Development of a three-dimensional non-fiber based optical imaging system co-registered with a clinical modality requires calibration prior to clinical data collection in order to validate the imaging data. Calibration is a multi-step and continuous process that includes hardware component testing, data simulation, phantom experimentation and patient image analysis. It is essential for any imaging system intended for clinical use. Multi-modality imaging techniques present some additional challenges associated with data co-registration.

Near-infrared (NIR) spectral tomography (NIRST) uses light (600-1000nm) that is preferentially absorbed by endogenous tissue metabolic markers hemoglobin, water, and lipids that are often altered in the presence of malignancy [1–5]. NIRST alone suffers from

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

low spatial resolution so it can be synergistically combined with detailed anatomic information from digital breast tomosynthesis (DBT). DBT is an x-ray based breast imaging technique approved by the FDA for breast cancer screening. It is highly spatially resolved in-plane, but the limited angle of source-detector rotation leads to partial volume averaging that reduces resolution parallel to the x-ray beam substantially. Thus, the resulting images have superb resolution in the directions perpendicular to the x-ray beam, but resolution parallel or axial to the plane is about ten times worse in the case of the NIRST/DBT system investigated here, as illustrated in Fig. 1 [6]. Hence, in the case of NIRST/DBT where DBT-derived regions-of-interest (ROIs) are incorporated into NIRST image reconstruction, the effects of (i) mischaracterization of the depth of an inclusion and (ii) under- or over-estimation of its volume need to be investigated.

Additionally, challenges are posed by combining the optical acquisition in the same geometry as the DBT which are different from previously reported multimodal integrations of optical imaging. One combination of DBT and NIR optical imaging has already demonstrated clinical success in distinguishing benign from malignant lesions [1]; however, the images were not co-registered for utilization of spatial priors in the optical image reconstruction. Ultrasound has been successfully coupled with NIR light [7,8], but in a reflectance geometry. Other combinations of NIR with MRI have appeared [9–12], but these systems incorporate fiducial markers to localize optical sources and detectors.

In our NIRST/DBT platform, the light source is not directly in contact with the breast – rather light is directed towards the tissue with mirrors as described in detail elsewhere [13,14]. This approach is less costly and allows an unlimited number of source positions to be used with arbitrary selection of locations, but correctly identifying each source position on the breast surface is more challenging.

In this paper, calibration and phantom studies are used to assess NIRST/DBT for co-registration of spatial locations between the two imaging systems, the effects of depth and volume mischaracterization of ROIs, and determination of the size and contrast resolution limits of the integrated imaging platform. These studies were facilitated by layered phantoms rapidly constructed from a series of slabs, one of which incorporated inclusions of different sizes and contrasts, that dramatically reduced the time and effort involved in the experimental system performance characterization. An example result from a clinical case is shown to demonstrate the utility of the phantom studies for system calibration and performance evaluation.

2. Materials and methods

2.1 Imaging system

The NIRST/DBT imaging system was a modified Hologic© Genesis prototype research DBT unit with integrated optical components developed at Dartmouth. Eight wavelengths between 661 and 940 nm were raster scanned at 1, 1.5, or 2 cm intervals across the top of the phantom with a detector panel comprised of seventy-five 1 cm x 1 cm photodiodes placed beneath the phantom for signal detection. Additional details on the imaging system can be found in prior publications [13,14] and it is shown in Fig. 1.

2.2 Source and detector localization

NIRST sources were placed on top of a phantom after directly visualization of the light source and metal markers placed at their most central location as shown in Fig. 2 for the most central source as well as the sources marking the corner of the rectangle. Afterwards, x-ray images were acquired of the metal markers and the detector panel. The DBT projections were reconstructed and slices corresponding to the depth of the source and detectors were analyzed. The source and detector central positions were correlated to the nearest DBT image pixel to ensure that the optical and x-ray data were completely co-registered. The source positions had been previously measured at a different phantom height with a 2 cm spacing between each source and those measurements were extrapolated for all heights using simple geometric

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

formulas to maintain constant position regardless of height. The detector positions had been modeled based on actual measurements of the detector panel. Average difference between the source position measured on DBT and the modeled source position was 3.0 mm while the average difference between the detector position on DBT and the modeled detector position was 2.2 mm with lower differences near the center of the field (2.5 mm and 1.5 mm for sources and detectors respectively).

2.3 Phantom design

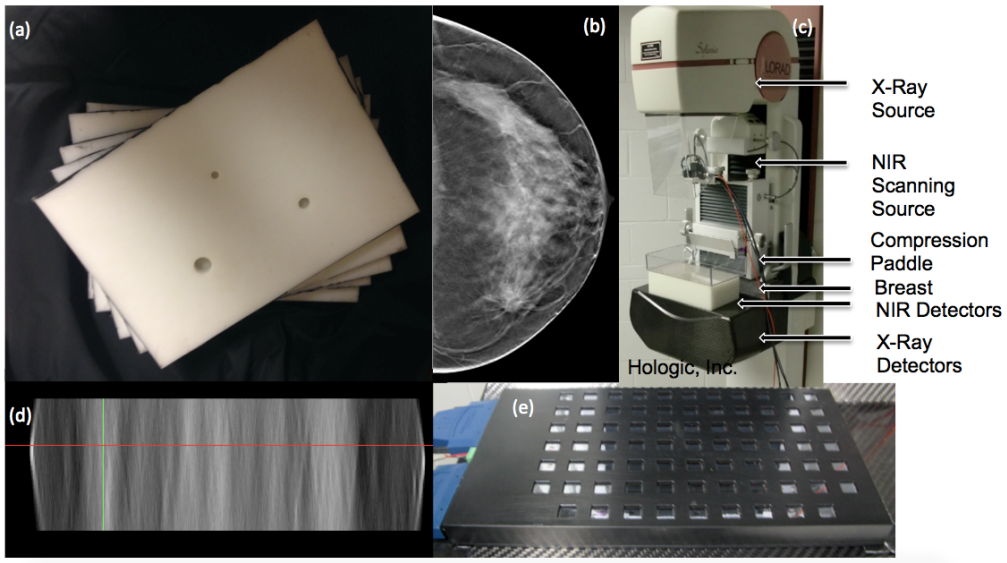

Fig. 1. (a) Photograph of the spectrally varying resin phantom slab used in the experiments. (b) DBT image slice of an invasive ductal carcinoma in the plane perpendicular to the x-ray beam. Tumor location is indicated with a white arrow. (c) Photograph of the DBT/NIRST imaging system patient interface. (d) DBT slice from same breast as (b) is shown in the parallel plane with the locations of the slice and tumor in (b) indicated by the red and green lines, respectively. (e) Photograph of the NIRST detector panel placed under the breast inside a specially designed polycarbonate case that can be removed during the procedure.

A layered phantom construction, shown in Fig. 1, was used for optical calibration studies. It was comprised of 18 x 26 cm x 1-2 cm slabs of a proprietary hard resin created by INO (Institut National D'Optique, Quebec City, CAN) with spectral variations in μ_a (0.03-0.06 mm^{-1}) and $\mu_{s'}$ (0.9-1.0 mm^{-1}), similar to the measured properties of breast tissue [15]. Optical properties were measured by INO at two wavelengths (740 nm and 785 nm) and calculated for the remaining wavelengths available in the NIRST/DBT system [16]. Slabs were either 1 cm or 2 cm in thickness and stacked in different orders and orientations to create a suite of geometries and optical properties. The 2 cm slab had cylindrical holes of different sizes to represent inclusions – a large hole of 1.5 cm diameter and 1.5 cm height, a medium hole of 1 cm diameter and 1 cm height, and a small hole of 0.7 cm in diameter and 0.7 cm in height. Slabs were arranged to mimic a lesion at four different depths and two different locations by changing the order of (and rotating) the layer with asymmetrically-placed holes. The composite phantom thickness was 6 cm, which is larger than the mean breast thickness of 5.3 cm for cranio-caudal breast mammography [17]. When only the homogeneous slabs were included, the phantom was used as a reference for calibration of clinical exams.

2.4 Inclusion measurements

The phantom was imaged in the NIRST/DBT system after a fixed liquid volume was pipetted into each small hole, starting with the lowest blood concentration (15 μM) which was

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

sequentially increased (by 15 μM) to a maximum of 60 μM. The liquid injected into each inclusion was comprised of 1.15% intralipid to approximate the scattering of the resin phantom, and included carefully measured concentrations of porcine blood and water. For a given set of inclusion positions, the slab containing the inclusion holes was not moved during sequential image acquisitions for each contrast level injection (subsequent inclusion position/depth changes did require slab reconfiguration followed by another series of contrast injections in the newly constructed, composite phantom which then remained stationary).

2.5 Data calibration

Data calibration consists of four steps: assigning the optical properties of the homogeneous reference phantom, altering the measured data based on attenuator settings, wavelength by wavelength reconstructions comparing the measured phantom data with simulated results using the known optical properties of the phantom and applying correction factors to the breast data. Lastly, generation of a homogenous initial guess for region based reconstruction of the breast data is completed. The calibration process involves a finite element mesh for the phantom and a different mesh for the breast so that exact matching of thicknesses is unnecessary. This initial guess is used for a wavelength by wavelength reconstruction hard priors guided reconstruction, leading to a single value obtained for each of the labeled regions. Only absorption is recovered, scattering for the phantom is obtained from initial measurements performed by INO.

2.6 Data reconstruction

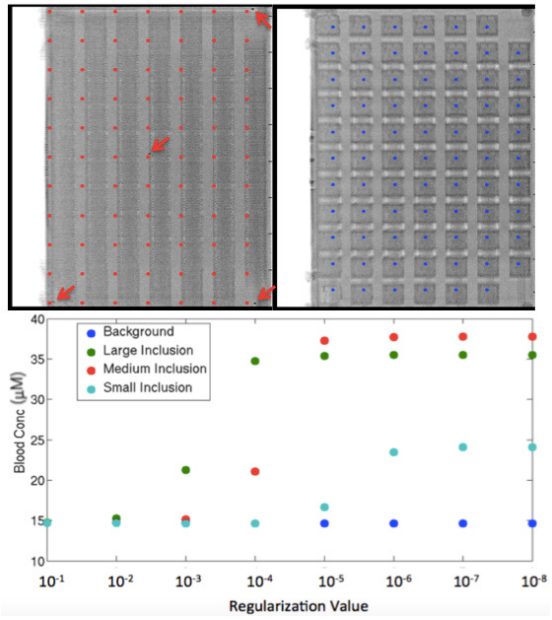

Fig. 2. DBT reconstruction slice corresponding to the height of the top of the optical phantom (left) with arrows pointing to metal markers of a few directly visualized source positions with the source positions marked in red. A DBT slice corresponding to the level of the detectors (right) with blue markings at the modeled detector locations. Graph shows dependence of inclusion HbT estimates with decreasing regularization.

The NIRFAST software package was used to reconstruct total hemoglobin (HbT) concentration images using a finite element method (FEM) technique [18,19]. FEM meshes representing the homogeneous and contrast phantoms were created using a three-dimensional (3D) algorithm that allowed higher node density in user-defined areas of interest. For each imaging geometry, meshes were created from the measured dimensions of the phantom, and

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

were refined in the inclusion regions to have a nodal spacing of 0.8 mm relative to 2.5 mm for the remainder of the background volume. The data set recorded for the lowest HbT (15 μM) was used for calibration since the hemoglobin absorption was similar to the resin at this concentration. The four most stable wavelengths - 660, 808, 830 and 852 nm – were used (the other wavelengths were less stable) and spectral image recovery was performed after the images for the individual wavelengths were reconstructed. Water and lipid contents were not assessed because of their non-physiological levels in the resin background and inclusions, and the elimination of higher wavelengths of data due to the lower stability of the light source under these illumination conditions.

Small regularization values were used to recover HbT for all the inclusions (30 μM) as shown in Fig. 2 because noise levels were low and the reference phantom was similar to the anomaly phantom. Smaller lesions required lower regularization values; in this case, 10^{-6} was necessary to fully recover the smallest inclusion, while 10^{-4} recovered HbT in the largest inclusion. All reconstructions applied prior knowledge on the size and location of the inclusions which stabilized the inversion and improved the reconstruction accuracy similarly to NIRST/MRI [11].

2.7 Clinical case

The 51 year old postmenopausal patient was enrolled in an institutional review board (IRB)-approved study after suspicious breast lesions were noted during routine screening mammography. The patient underwent NIRST/DBT scans of both the left and right breasts in cranio-caudal positioning. A calibration data set was obtained from the 2 cm homogenous section of the phantom without inclusions combined with the four homogeneous 1 cm slabs. DBT images were read by a breast radiologist (SPP), who outlined two regions of interest (ROIs) in the left breast. DBT images were segmented into adipose, fibroglandular, and two ROIs using software developed for DBT breast segmentation [20,21]. Parallel plane extension of the ROIs was assigned to be equal to the largest diameter of the lesions measured in the plane perpendicular to the x-ray beam. These segmented images were used as the basis for FEM mesh creation, and a higher node density was deployed in the two ROIs compared to the adipose and fibroglandular regions. Data measured within 15 mm of the breast periphery, outside the breast, or more than two standard deviations from a best-fit line of signal decrease with detector distance were removed prior to image reconstruction. Scattering was obtained from another near infrared imaging system with frequency domain capabilities [22]. After a wavelength-by-wavelength calibration of the data, a spectrally constrained algorithm within NIRFAST was used to calculate hemoglobin, oxygen saturation, water, and lipid fractions [23]. A single scattering average was obtained from the bulk breast tissue measurement performed using a different frequency domain imaging system. After a wavelength-by-wavelength calibration of the data, a spectrally constrained algorithm within NIRFAST was used to calculate hemoglobin, oxygen saturation, water, and lipid fractions [23]. This method has shown to have increased accuracy and couldn't be used in the case of the phantom imaging due to the presence of a nonphysiologic resin material [23]. The suspicious lesions were subsequently biopsied and the results were correlated with the NIRST/DBT image data.

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

3. Results

3.1 Depth dependence

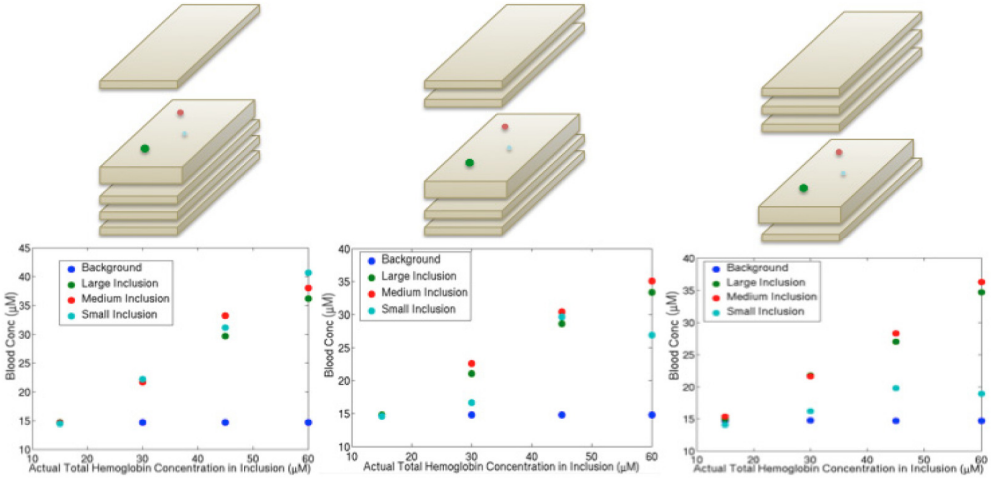

Fig. 3. Graphs depicting hemoglobin recovery vs. actual hemoglobin concentration for the background and inclusions of different sizes at three different depths. Pictorial demonstrations of the phantom set up are shown above each graph. The slab with inclusions was 2 cm thick and slabs without inclusions were 1 cm thick

Hemoglobin concentrations in inclusions placed at different depths showed linear recovery of about 70% of the actual values on average as shown in Fig. 3. Estimates in the largest (15 mm) and medium (10 mm) sized inclusions demonstrated a linear response regardless of depth, whereas linearity was maintained in the smallest (7 mm) inclusion only at the shallowest depth. The 7 mm inclusion was more susceptible to small position shifts, which may have caused the consistently lower concentration values at the larger depths. Additionally, fluctuations in source strength may have introduced variability in the estimates for the smallest inclusion at the greatest depth because the signal disturbances caused by the optical heterogeneity were also very small.

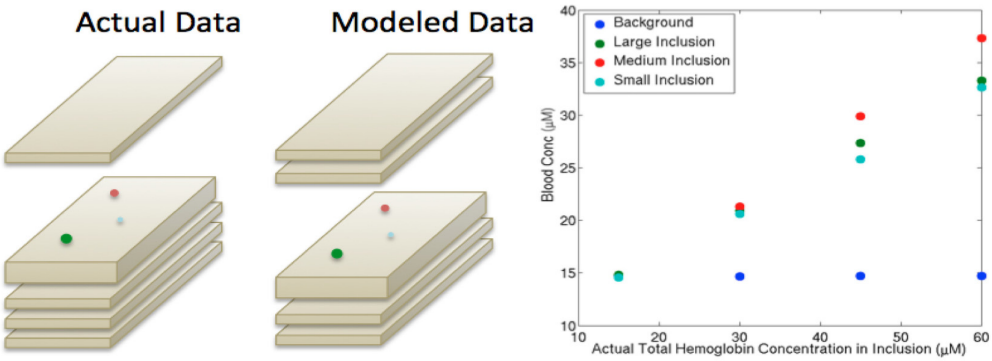

Fig. 4. Data from a phantom with actual inclusions positioned below a 1 cm slab on the top of the phantom (left picture) but modeled as below 2 cm of homogeneous background (right picture). The hemoglobin estimates are plotted (right graph) and can be compared with the left graph in Fig. 3, which uses the same data but correctly models the inclusion depths.

Mischaracterization of inclusion depth from segmentation errors caused by poor axial resolution in the DBT scans was evaluated by misplacing the inclusion locations (i.e. inclusions were segmented as being 1 cm deeper than their actual positions). These results,

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

shown in Fig. 4, indicate a reduction in the recovered hemoglobin of 10-20% relative to the values in Fig. 3. Hence, mislocalization of a breast lesion may cause differences in HbT recovery of up to 20% (lower in the case where the segmented region is deeper than its true position by ~1 cm).

3.2 Volume mischaracterization

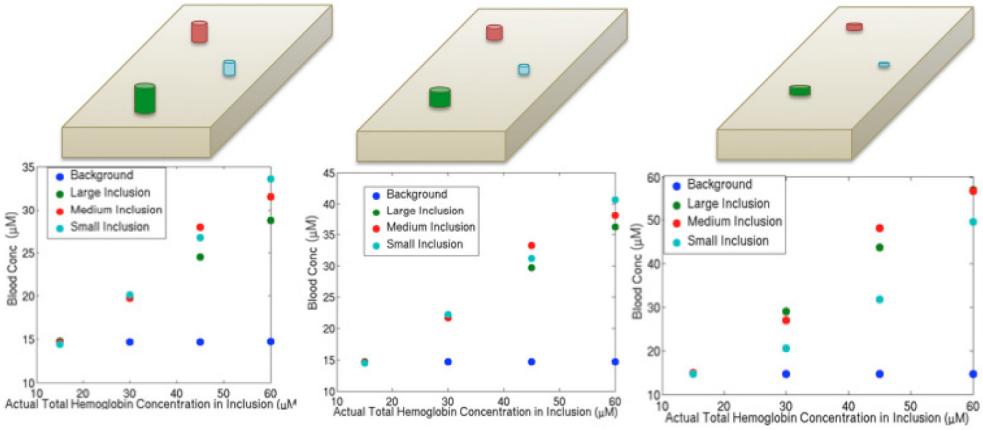

Fig. 5. Effects of volume mischaracterization when inclusions were positioned 1 cm and 2.5 cm below the top of the phantom. Pictures above each plot indicate the modeling performed: overestimation of the volume of inclusions by 50% (left), true volume of inclusions (middle), and underestimation of the volume of inclusions by 50% (right).

Table 1. Effects of inclusion size on the percent of actual hemoglobin recovered

Inclusion Size	Underestimated	Normal	Overestimated
Large	96.3	59.5	56.0
Medium	97.2	64.1	60.1
Small	74.0	67.7	60.8

Table 2. Effects of concentration on the percent of actual hemoglobin recovered

Concentration	Underestimated	Normal	Overestimated
30 μM	85.0	70.5	66.2
45 μM	91.6	63.0	58.6
60 μM	90.8	57.7	52.1

To investigate the impact of errors in volume introduced by poor resolution along the x-ray source to detector direction, inclusions were extended by 50% of their volume in height in one instance, and shrunk by 50% in another. In all cases, the average depth of inclusions remained the same. Figure 5 shows that the responses in the HbT estimates are similar (except for the 7 mm inclusion in the underestimated case) for the three volume assumptions, although the actual concentration values vary to some extent.

For the overestimated volumes, the inclusions included some actual background; thus, the average contents in the inclusions were lower in HbT because the background possessed lower HbT. Conversely, for the underestimated volumes, portions of the inclusions were mischaracterized as background. Here, the recovered HbT was not altered in the much larger background; instead, greater signal attenuation was attributed to a smaller inclusion region leading to higher HbT recovery. The small inclusion behaved less systematically because its volume became very small, and reduced optical measurement sensitivity in the

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

underestimated case. Overall, there was no correlation between inclusion size and recovery, nor was there one between the HbT concentration and the percent of actual HbT recovered as demonstrated.

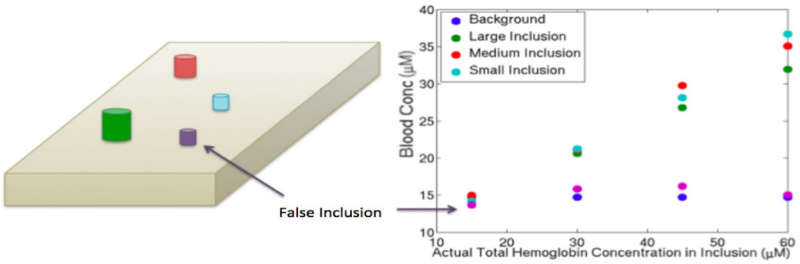

Fig. 6. A portion of the background was designated as a possible region of interest shown in purple), modeled and reconstructed along with the actual inclusions and background. Results for this false inclusion were similar to the background as shown in the graph (right).

To investigate whether contrast was caused by the specialized meshing of the inclusions, an additional region that was not present in the experimental phantom configurations was added to the model as shown in Fig. 6 and was created with the same volume and mesh resolution as the smallest inclusion. The hemoglobin levels recovered in this region were very close to those of the background, and did not vary with different HbT concentrations in the actual inclusions. These results indicate that a false positive finding is unlikely to occur, i.e., that a region in a DBT image that was labeled as suspicious but actually had no greater HbT content than the background would not likely appear to have a higher concentration.

3.3 Clinical case

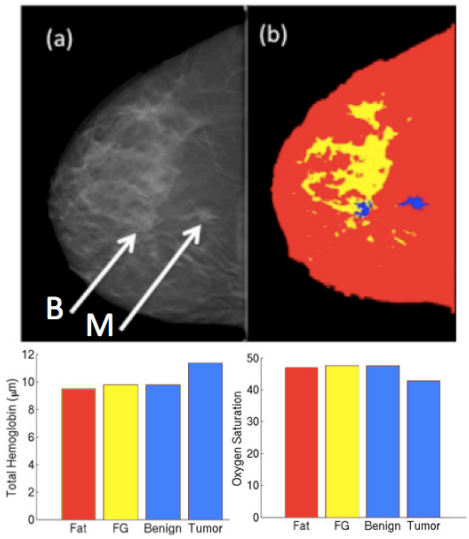

Fig. 7. (a) DBT slice for a patient with benign (left arrow) and malignant (right arrow) lesions. (b) DBT image stack was segmented into adipose, fibroglandular (FG) and two ROIs. (c) Total hemoglobin concentration and oxygen saturation obtained from the NIRST/DBT recovery for each region.

Initial clinical results from a patient with malignant pathology imaged on the NIRST/DBT system are shown in Fig. 7. After imaging, the patient underwent biopsies for suspicious lesions; one demonstrated invasive ductal carcinoma and the second, benign sclerosing adenosis. The NIRST/DBT results for this subject demonstrate increased HbT concentration

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

in the tumor (11.3 μM), compared with the fibroglandular and benign regions (9.8 μM) as well as the contralateral breast (9.9 μM). Additionally, the tumor had lower oxygen saturation, as might be expected from the higher oxygen demands and disorganized vasculature typically found in tumors [10], [24]. This clinical example nicely illustrates a case where the NIRST/DBT system successfully distinguished benign from malignant lesions in the same breast.

4. Discussion

Custom-designed rectangular resin-based phantom slab with spectrally varying absorption profiles are durable and easy to use, and offer an ability to create multiple configurations for assessing inclusion size, contrast, and depth dependence in optical images acquired with a DBT-guided NIRST system. In repeatability experiments, the same inclusion phantom was measured seven times without movement (data not shown). Mean variation in the estimated HbT in the background, large, medium, and small inclusions were calculated to be 0.5%, 5.2%, 13.3% and 13.4%, respectively. Sources of error were attributed to short-term shifts in source strength and detector sensitivity. In the large background region, almost no changes occurred in HbT as small fluctuations were averaged over a large number of data points. However, greater variation was found in the inclusion estimates because their measurement sensitivity was limited to a small number of sources and detectors, depending on their locations. Hence, small shifts in source strength corresponded to more variation in recovered HbT concentrations, especially for smaller ROIs (inclusions).

Hemoglobin contrast recovery was linear at all depths examined and for all inclusion sizes except the smallest (7 mm in height and depth). The smallest inclusion was more easily affected by changes in signal variation from the laser source, positioning errors, and the limited number of nodes created by the meshing process. Calibration showed source strength variations on the order of 0.5-4% depending on the laser diode. Unfortunately, when measuring small inclusions in a large phantom, changes in signal attenuation for the smallest inclusion at source and detector pairs nearest its location are only a few percent as well. For example, in the case of the 7 mm inclusion, a 3% difference in signal attenuation was detected for the highest contrast and the closest source detector pair relative to the homogeneous case, and is close to the size limit for contrast imaging with the current NIRST/DBT system.

Inclusions closer to the surface were recovered with higher concentrations than those in the middle of the phantom. Mischaracterization of depth led to decreases in chromophore recovery of 10-20% in a region of interest. However, the effects were much less than those caused by over- or under-estimation of inclusion volume which caused variations in HbT estimates of 15-50%.

Breast DBT segmentation for NIRST will benefit from these results, as accurate segmentation of tissue structures is important for recovering the correct HbT concentrations. However, a phantom has distinct boundaries, whereas ROIs in breast tissue can have much more uncertainty. In manual DBT segmentation, the radiologist defines the boundaries of the ROI, and a prior study with 173 lesions indicated that DBT overestimates the size by 1 cm in 14.5% of lesions and underestimates the size by 1 cm in 5% of lesions compared to pathology [25]. In another study, where 133/149 (89%) lesions were visualized by DBT, size measurements for 34/133 (26%) lesions were either under- or over-estimated (17 lesions each) by 5 mm compared to pathology [26]. The radiologist's measurements are based on the longest axis of the tumor, so total volume estimates are likely to have larger margins of error. Determining the extent of the tumor in the axial direction is particularly challenging when partial volume averaging can cause the lesion appear throughout nearly the entire axial span of the DBT images. However, Doppler ultrasound has shown that regions of increased vascularity often occur outside of the defined tumor area [27], so overestimation of the tumor boundaries may not lead to an overall decrease in hemoglobin recovery. Although segmentation errors will occur when determining the boundaries of an ROI, the effects of such errors are likely to be less in breast tissue relative to phantoms because of the lack of rigid boundaries. Any volume overestimation will result in an averaging of tissue, which can

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015

lower contrast recovery if the surrounding region is not highly vascularized. Tumor volume underestimations can result in higher-than-expected tumor contrast if the region surrounding the tumor is large and non-absorbing.

The potential of NIRST/DBT to distinguish small benign regions from malignant ROIs was demonstrated through the false inclusion phantom experiment as well as a clinical example. In the phantom case, the false ROI was defined in the same way as the other inclusions, but it was merely a section of the resin background. The HbT estimate in this region was similar to the background (14-16μM vs. 15μM for the background, within 7.5% of the actual value for a small lesion, less than the acceptable 13.4% natural variation from the repeatability studies), standard deviation < 1μM across all measurements, suggesting that the meshing and image reconstruction process are unlikely to cause false positive results. The argument was further supported by the clinical case, which showed elevated HbT (11.3 μM) in a malignant tumor compared to a benign lesion (9.8 μM). The difference between the benign and malignant lesion is 15%, greater than the variability shown in phantom experiments for similarly sized lesions (5.2%). Hence it is unlikely that this 15% difference could be attributed to measurement variability and demonstrates a true difference between the two tissue types

4. Conclusions

NIRST/DBT imaging of the breast faces several challenges related to extraction of DBT anatomical priors and the expected size of suspicious ROIs detected on a screening exam. Creation of systematic calibration and co-registration processes, and construction of a durable, heterogeneous phantom that can be configured in many different ways facilitated determination of the imaging limitations of the new platform. The phantom was easy to use and clean, and is long-lasting and durable. Significantly less time was required to perform experiments with this phantom relative to other methods for creating heterogeneous phantoms.

Linear recovery of HbT occurred for cylindrical inclusions measuring 7 mm, 10 mm and 15 mm in diameter and height in most cases. HbT concentration estimates were similar for the two larger inclusions, but values were lower for the smallest ROI in some situations. Quantitative recovery was affected by the depth of the inclusion within the phantom, mischaracterization of its depth, and most significantly, under or overestimation of its volume. The NIRST/DBT system was capable of distinguishing real changes in HbT levels from an arbitrarily defined ROI in the background without contrast, and also distinguished benign from malignant lesions in vivo of sizes similar to, or smaller than, what would normally be detected through mammographic images.

Phantoms similar to the ones described here could be used in the development of any optical imaging system to improve signal-to-noise ratios, and/or assess imaging performance or calibrate source/detector signals. The outcomes of such phantom experiments guide patient selection and optimize efficiency, and increase the likelihood of obtaining significant clinical results in the future.

The results described here will be essential in future clinical studies. Based on these results, it is possible to detect tumors at least as small as 7mm, about half the size of tumors normally detected in a screening population [28]. During patient exams, efforts will be made to position tumors centrally within the breast to minimize errors from positioning. Expanding tumor areas to include an additional 3mm boundary to account for any positioning uncertainty may improve chromophore recovery. Additionally, thresholds for distinguishing benign from malignant lesions can be estimated based on the variability measurements from this phantom work and differences from the background of less than 13.5% may be the threshold for assessing benign from malignant status.

Acknowledgments

This work was funded by National Institutes of Health (NIH) Grants R01CA139449 and F30CA168079.

#250601 Received 24 Sep 2015; revised 30 Oct 2015; accepted 1 Nov 2015; published 19 Nov 2015