Lesion Ground Truth Estimation for a Physical Breast Phantom

Suneeza Hanif[1], Frank Schebesch[1], Anna Jerebko[2], Ludwig Ritschl[2], Thomas Mertelmeier[2], Andreas Maier[1]

[1]Pattern Recognition Lab, FAU Erlangen-Nürnberg, Erlangen, Germany
[2]Siemens Healthineers, Erlangen, Germany
suneezahanif02@gmail.com

Abstract. The extraction of microstructures (like microcalcifications or masses) from a DBT phantom can be used for image quality assessment. A complete specification includes exact positions and dimensions of the microstructures, which is not always available. We propose a technique to estimate the required ground truth data from a set of multiple acquisitions. A 3D registration algorithm for DBT data is used to identify different breast phantom components and to perfectly align microstructures within the phantom. The registered data, showing variations of the same ground truth structures, is then combined to an estimate of the microstructures. This approach could be shown to improve the registration result itself, and to enable the determination of the actual parameters of the microstructures.

1 Introduction

Breast cancer is the most common cancer in women, and it is the second cause of cancer death in more developed regions [1]. Early diagnosis has higher chances of cure or can help to slow down the progress of the cancer.

Digital breast tomosynthesis (DBT) employs information from multiple images of a breast, scanned at different view angles, to allow readers to extract the 3D features in the breast tissue.

Due to the similarity of the elemental composition of normal and abnormal tissue, the optimization of image quality in DBT imaging is a critical factor, and optimal detection of masses and microcalcifications [2] is crucial. The extraction of such features from the images can be used for image quality assessment and can be carried out using physical phantoms and the knowledge about certain test objects in the phantom designed to meet specific requirements [3].

Although phantom images do not completely resemble clinical images, they have some advantages over clinical data as their complete structure is known by construction which usually delivers the ground truth in medical imaging. However, the generation of a proper testing breast phantom is still an ongoing research [4]. Several phantoms have been produced and while some of them mock up general tissue properties quite well, the specification is not necessarily accurate enough to determine the exact parameters of its contents.

To estimate the ground truth within the CIRS 20 breast phantom [5], we propose an image registration method to align its components and successively computate an approximation of the ground truth.

2 Methods

The computerized imaging reference system (CIRS) model 020 BR3D breast imaging phantom is distributed as a batch of six heterogeneous slabs. These include a single slab (we denote it as target slab) which contains an assortment of microcalcifications, fibers and masses of different sizes. The slabs can be rearranged to generate different DBT data ($6! \times 2^6 = 46080$ combinations) implying a variation of the local background structure as well as a repositioning of the target structures. Thereby, numerical errors in the reconstruction process affect result image structures differently. In our method, we want to take advantage of the partially rigid properties of the different compositions by aligning the target slab from different acquisitions, varying both the dose level and the reconstruction method.

2.1 Slab Registration

In a first step, we try to align the corresponding inserts in the target slabs of different acquisitions and phantom compositions. In order to deal with the large image data sets (approximate volume size is $1300 \times 2400 \times 60$) and to improve convergence properties of the algorithm, a rigid registration algorithm with multistage scheme and multiresolution approach is used. As an initialization we use a reference image which is supposed to have an equal or better image quality than the images to be registered (for example from a high dose acquisition). The registration is first performed at the coarsest scale level. The transform parameters determined by the registration are then used to initialize the registration at the next finer level. This process is repeated till the finest level. In this way, large misalignments can be recovered early at a low scale and more detailed ones are accounted at increasingly fine resolutions.

With a 3-stage registration algorithm we were able to avoid significant misalignments. It is based on the idea of using simpler transforms and more aggressive optimization parameters at early stages. The 3D Euler transform (contains six degrees of freedom) is used for initial coarse registration levels and upgraded to an affine transform by incorporating more degrees of freedom at the finer levels.

Due to the indistinct texture of the simulated anatomical background, feature tracking, as it is often used for rigid motion problems [6], was replaced with the idea of using abstract texture information. In the first stages, normalized cross-correlation is used to avoid large misregistrations in the first iterations of the registration algorithm, while in later stages mutual information is used to remove small misalignments. The similarity measures are evaluated on image

subsets containing the inserts regions. The intermediate and output images are evaluated by linear interpolation.

The registration problem can be formulated as a minimization problem,

$$\hat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \ \mathcal{F}\left(\boldsymbol{\mu}, I_f, I_m\right), \tag{1}$$

where the staged cost function $\mathcal{F}$ represents the respective similarity measures that are minimized. I_f and I_m are the fixed and moving images respectively. $\boldsymbol{\mu}$ contains the external transform parameters and $\hat{\boldsymbol{\mu}}$ represents the best transform parameters that align the images.

To optimize the cost function in stage 1 and 2, regular step gradient descent optimizer is used [7]

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - a_k \kappa \nabla_{\boldsymbol{\mu}} \mathcal{F}, \quad \text{for } k = 1, 2..., \tag{2}$$

where a_k is the step size, κ is a relaxation factor that can be varied between 0 and 1, and $\nabla_{\boldsymbol{\mu}} \mathcal{F}$ is the gradient of the cost function w.r.t $\boldsymbol{\mu}$. In the final registration stage, conjugate gradient descent line search optimizer is used for fine tuning. Here the update is a combination of the local gradient $\nabla_{\boldsymbol{\mu}} \mathcal{F}$ and the search direction $\boldsymbol{d}_{k-1}$ of the previous iteration

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + a_k \boldsymbol{d}_k, \quad \boldsymbol{d}_k = -\nabla_{\boldsymbol{\mu}} \mathcal{F} + \beta_k \boldsymbol{d}_{k-1}, \quad \text{for } k = 1, 2..., \tag{3}$$

where β_k is determined by Polak-Ribiere conjugate gradient and a Golden section line search is used to find the best value for the step size a_k.

2.2 Ground Truth Estimation

The described registration allows an accurate positioning of the phantom details with respect to the reference or fixed image. We assume that in a set of registered images $\{I_i\}_{i=1,\ldots,N}$, each image I_i contains a different part of the ground truth in a sense that it is overlayed by noise but always showing the same objects.

Image data obtained from different acquisition dose levels, slab combinations and reconstruction schemes contains different noise levels and variations in the spatial alignment of the microstructures. However, if depicting all these images as a (small) variation of the actual ground truth object, by minimizing the sum of squared distances the ground truth image can be estimated

$$\mathrm{SSD}\left(\hat{I}\right) = \sum_{i=1}^{N} \left(I_i - \hat{I}\right)^2 \to \min, \tag{4}$$

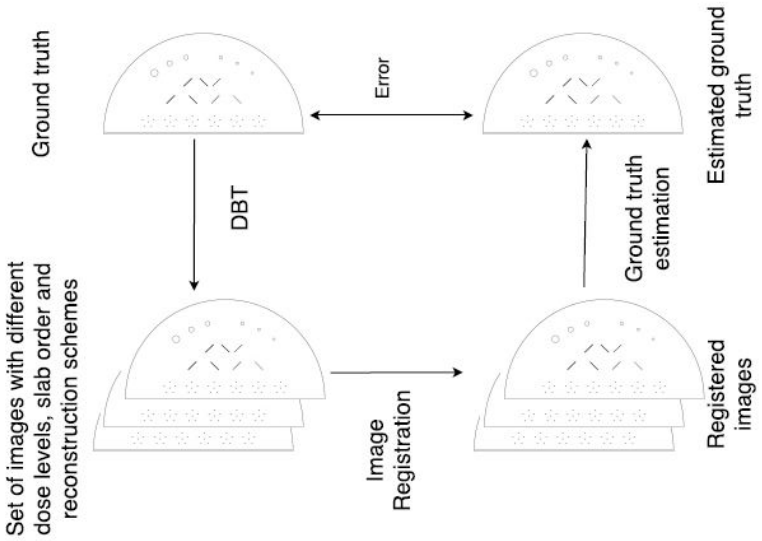
which is solved by

$$\hat{I} = \frac{1}{N} \sum_{i=1}^{N} I_i. \tag{5}$$

The reliability of this estimate depends on the amount and variety reconstruction images and the degree of misalignment per image which is why the efforts for the registration were made.

The schematic in fig. 1 illustrates the proposed idea of the CIRS 20 ground truth estimation using multiple reconstructed DBT data.

Fig. 1. Flow diagram for ground truth estimation from multiple reconstruction data



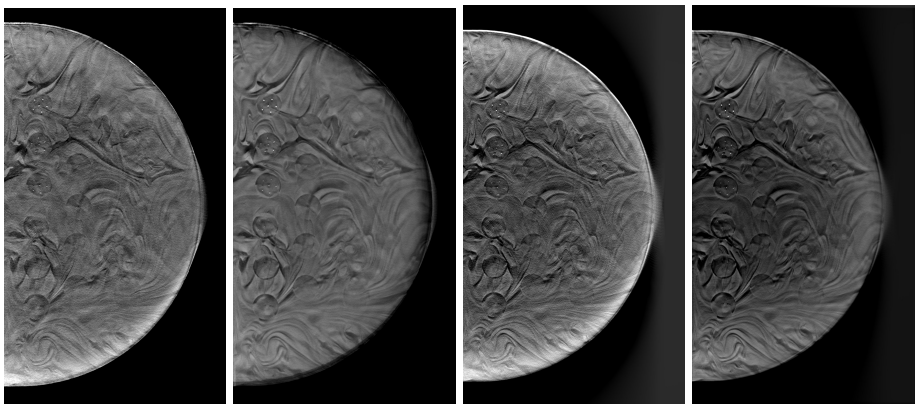
3 Experiments and Results

The image dataset consisted of images from 10 random slab phantom compositions which were acquired at three dose levels: half of, equal to, and two times reference radiation dose (determined by automatic exposure control). For reconstruction we used two DBT reconstruction schemes: filtered back projection (FBP) and super resolution statistical artifact reduction (SRSAR) [8].

We tested if the proposed method allows improved ground truth estimation from images of the same dose level and a single reconstruction scheme. The images are registered to a reference from the set and the estimated image is computed by (5), see fig. 2 for the result images. For initial registration stage, the optimizer parameters are empirically calculated and set to $\kappa = 0.5$ and $a_k = 2.72$. Similarly the minimum step size and the number of iterations are also empirically estimated. Then, at each subsequent stages and resolution levels, the minimum step size is reduced by a factor in order to allow the optimizer to focus on progressively smaller regions.

The microstructure properties of both estimated and non-estimated reference images were compared to the values specified in the datasheet of CIRS phantom (cf. table 1). The average differences in diameter with respect to the specifications of the selected microstructures (microcalcification, fiber, largest mass and smallest mass) are 0.04 mm, 0.02 mm, 0.04 mm, and 0.17 mm in the estimated image, whereas for the single reference they are 0.1 mm, 0.12 mm, 0.25 mm and 0.21 mm, respectively. This means that the dimensions of the microstructures in the estimated image are closer to the actual values as compared to the single image.

Fig. 2. Single vs. estimated reference images for SRSAR and FBP reconstruction schemes



(a) Single SRSAR image (b) Estimated SRSAR image (c) Single FBP image (d) Estimated FBP image

Table 1. Ground truth estimation comparison with FBP reconstruction

Microstructure	⊘ Datasheet	⊘ Single	⊘ Estimated
Microcalcification (1)	0.40 mm	0.30 ±0.11 mm	0.36 ±0.05 mm
Fiber (7)	0.60 mm	0.48 ±0.04 mm	0.58 ±0.02 mm
Mass (14)	6.30 mm	6.05 ±0.10 mm	6.26 ±0.07 mm
Mass (19)	1.80 mm	1.59 ±0.03 mm	1.63 ±0.02 mm

In a second experiment the estimated image itself is then used as a reference image to study the improvements in overall image registration process as compared to the image taken from data set. table 2 illustrates the average outcome in terms of mutual information for the target slab evaluated at the end of the registration process.

4 Conclusion

The implemented registration algorithm is tested on a large number of physically varied sample set and the results showed that the microstructures are perfectly registered with only differences occurring because of the different microstructures sizes, resulted from different dose levels and reconstruction schemes.

Although the estimation from registered data is not able to acquire all possible noisy image variations, the results indicate that we can still find a reasonable estimate in which the overall noise structure seems to cancel out sufficiently using a subset.

Table 2. Mutual information after registration

MI	FBP	Estimated FBP	SRSAR	Estimated SRSAR
Target slab	0.689 ± 0.221	0.844 ± 0.240	0.635 ± 0.342	0.801 ± 0.181

In terms of actual improvement, the computed diameters of the CIRS 20 phantom microstructures in the estimated image show values in a closer range to the specification values. This behaviour seems to be independent of the reconstruction scheme, however, the general idea allows a mixed set composed of images from different reconstruction schemes, as long as the alignment of details by the underlying registration aligns is working.

The estimation accuracy using the proposed approach is limited by the data nonetheless, so in consideration of an automated image quality assessment a sufficient variety of reconstruction data must be available. In future work it would also be interesting to see a validation of the proposed method using high resolution acquisition techniques like micro-CT.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al.. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. Lyon, France; 2013. Accessed on 24/8/2016. Internet. Available from: `http://globocan.iarc.fr`.
2. White D, Tucker A. A test object for assessing image quality in mammography. The British journal of Radiology. 1980;53(628):331–335.
3. Suryanarayanan S, Karellas A, Vedantham S, Glick SJ, D'Orsi CJ, Baker SP, et al. Comparison of tomosynthesis methods used with digital mammography. Academic Radiology. 2000;7(12):1085–1097.
4. Cockmartin L, Bosmans H, Marshall NW. Comparative Power Law Analysis of Structured Breast Phantom and Patient Images in Digital Mammography and Breast Tomosynthesis. Medical Physics. 2013;40(8):081920–1–17.
5. BR3D Breast Imaging Phantom - Model 020 - Data Sheet;. Accessed: 2016-01-20. `http://www.cirsinc.com/file/Products/020/020%20DS%20102915.pdf`.
6. Klüppel M, Wang J, Bernecker D, Fischer P, Hornegger J. On Feature Tracking in X-Ray Images. In: Deserno TM, Handels H, Meinzer HP, Tolxdorff T, editors. Bildverarbeitung für die Medizin 2014. Informatik aktuell. Berlin Heidelberg; 2014. p. 132–137.
7. Van der Bom I, Klein S, Staring M, Homan R, Bartels L, Pluim J. Evaluation of optimization methods for intensity-based 2D 3D registration in x-ray guided interventions. In: SPIE Medical Imaging. International Society for Optics and Photonics; 2011. p. 796223–796223.
8. Abdurahman S, Jerebko A, Mertelmeier T, Lasser T, Navab N. Out-of-plane artifact reduction in tomosynthesis based on regression modeling and outlier detection. In: International Workshop on Digital Mammography. Springer; 2012. p. 729–736.