Two Methods for Simulation of Dense Tissue Distribution in Software Breast Phantoms

Joseph H. Chui, Rongping Zeng*, David D. Pokrajac†, Subok Park*,
Kyle J. Myers*, Andrew D. A. Maidment, and Predrag R. Bakic

University of Pennsylvania, Department of Radiology, Philadelphia, PA
* Center for Devices and Radiological Health, Food and Drug Administration, Silver Spring, MD
† Computer and Information Sciences Department, Delaware State University, Dover, DE

{Joseph.Chui | Andrew.Maidment | Predrag.Bakic}@uphs.upenn.edu
DPokrajac@desu.edu {Rongping.Zeng | Subok.Park | Kyle.Myers}@fda.hhs.gov

ABSTRACT

Software breast phantoms have been developed for use in evaluation of novel breast imaging systems. Software phantoms are flexible allowing the simulation of wide variations in breast anatomy, and provide ground truth for the simulated tissue structures. Different levels of phantom realism are required depending on the intended application. Realistic simulation of dense (fibroglandular) tissue is of particular importance; the properties of dense tissue – breast percent density and the spatial distribution – have been related to the risk of breast cancer. In this work, we have compared two methods for simulation of dense tissue distribution in a software breast phantom previously developed at the University of Pennsylvania. The methods compared are: (1) the previously used Gaussian distribution centered at the phantom nipple point, and (2) the proposed combination of two Beta functions, one modeling the dense tissue distribution along the chest wall-to-nipple direction, and the other modeling the radial distribution in each coronal section of the phantom. Dense tissue distributions obtained using these methods have been compared with distributions reported in the literature estimated from the analysis of breast CT images. Qualitatively, the two methods produced rather similar dense tissue distributions. The simulation based upon the use of Beta functions provides more control over the simulated distributions through the selection of the various Beta function parameters. Both methods showed good agreement to the clinical data, suggesting both provide a high level of realism.

Keywords: Breast cancer imaging, anthropomorphic breast phantoms, software breast phantoms, validation, fibroglandular tissue distribution, Beta functions.

1. INTRODUCTION

Virtual Clinical Trials (VCTs) are emerging as a preclinical complement to clinical trials of breast imaging systems which are often longer and more expensive. In VCT, the simulations of breast anatomy, image acquisition, and model observers are combined to form a simulation pipeline. Realistic simulation of dense tissue is of particular importance since several properties of dense tissue may be used as imaging biomarkers of breast cancer risk. In this paper, we consider improvements to the glandular tissue distribution of the breast anatomy simulation component of the pipeline.

The 2D and volumetric fractional amount of dense tissue (called breast percent density) and the spatial distribution of dense tissue (called parenchymal texture) are known to correlate with cancer risk [1-7]. In this work we have compared two methods for simulation of dense tissue distribution in our software breast phantom design. The simulated distributions have been compared with distributions reported in the literature estimated from the analysis of breast CT images [8]. Covariance profiles estimated from phantom images, created with different methods for dense tissue simulation, are also compared to clinical data reported in the literature [11].

2. METHODS

Our proposed simulation of dense tissue is performed on breast phantoms created using a method proposed by Bakic et al [9]. In this proposed method, each simulated breast is divided into a predefined set of compartments. Each compartment consists of a seed point and a shape function. The seed point defines the compartment location, while the shape function defines its orientation. The values of seed points and shape functions are generated randomly based on a number of parameters defined by users. Each compartment is then labeled with different material types. Figure 1a) shows the cross section of a breast phantom where the compartments have not been labeled and figure 1b) shows a labeling of compartments to dense or fat tissues. A compartment is labeled as dense tissue if the criteria defined by the methods are met; or as fat tissue otherwise.
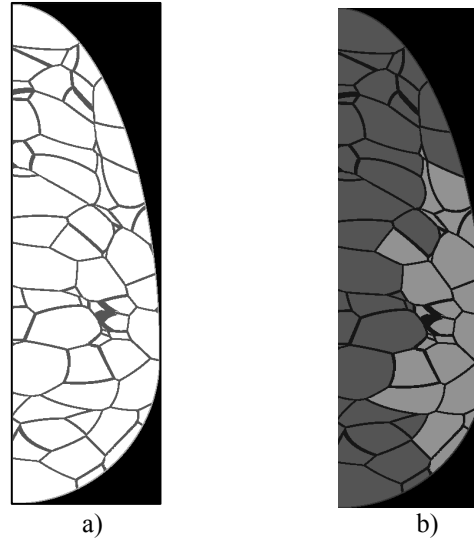

a)   b)

Figure 1. a) A cross section of simulated phantom. b) Each compartment inside the phantom is labeled with tissue types (light gray – dense, dark gray – fat).

A target number of compartments being labeled as dense is first determined based on a desired target volumetric breast density (VBD) defined by users. Each compartment is then assigned a probability of it being labeled as dense tissue, based on factors such as its location in the breast. Finally, each compartment is then labeled randomly based on its probability value.

2.1. Gaussian distribution method

The Gaussian distribution method for labeling dense components was originally proposed by Bakic et al [9]. In this method, the probability of a compartment labeled as dense tissue is determined by a Gaussian function:

$$p_G(\mathbf{s}_i) = \frac{\exp\left(-\sigma \cdot f_M^{\,2}(s_{xi} - a, s_{yi}, s_{zi})\right)}{Z} \tag{1}$$

where $f_M(.)$ represents the compartment shape function consistent with the quadratic decision boundaries described by a maximum *a posteriori* (MAP) classifier; a is the x coordinate of a simulated nipple point ($y = z = 0$), $\mathbf{s_i}$ (s_{xi}, s_{yi}, s_{zi}) are coordinates of seed vectors for the i-th compartment, and σ is a scaling coefficient. Z is a normalization constant chosen based upon a user-specified VBD of the phantom. In this method, the compartments near the nipple have a higher probability of being labeled as dense tissue compared to the ones further from the nipple.

2.2. Beta distribution method

In the Beta distribution method, the probability of a compartment labeled as dense tissue is determined by a function given by the product of two separate Beta functions:

$$p_B(\mathbf{s}_i) = \frac{Beta(\frac{s_{xi}}{a}; p_1, q_1) \times Beta(\frac{r_i}{R_i}; p_2, q_2)}{Z} \tag{2}$$

where:

$$Beta(x; p, q) - \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}, 0 \le x \le 1, B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1} dt,$$

s_{xi} is the distance of i-th seed point from the chest wall in posterior to anterior direction; r_i is the radial distance of seed point from center of its coronal slice; and R_i is the maximum radius of the coronal plane containing the i-th seed point.

Since these two beta functions are functions of distance in different directions, the distributions of dense tissue can be controlled separately in the chest wall to nipple direction and radial direction. Moreover, the shape of the beta function changes with different (p, q) values. Figure 2 shows two examples of beta function where the shapes are one-sided or two-sided depending on the values of (p, q).
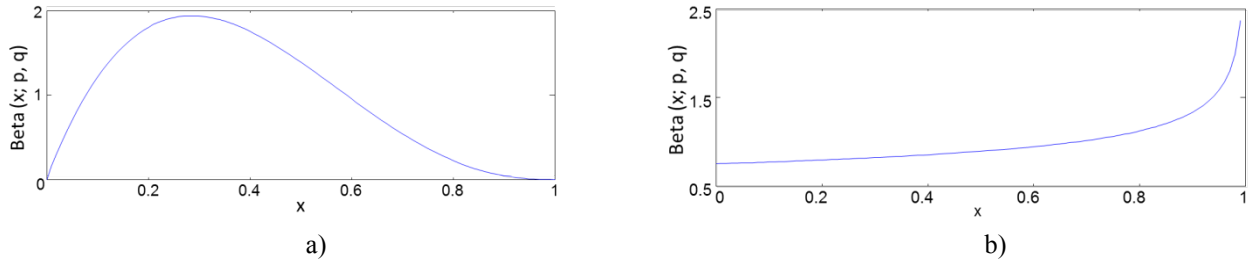

a)   b)

Figure 2: Beta distributions of two different (p, q) values: a) p = 2, q = 3.5; b) p = 1.0, q = 0.75.

2.3. Simulated acquisition of phantom images

Simulated mammograms are generated using software phantoms in a 2-step procedure. First, the breast deformation due to clinical mammographic compression is simulated by a finite element model [10]. The finite element model is implemented, using Abaqus (version 6.10-EF; Dessault Systèmes Americas, Waltham, MA), assuming a hyperelastic, almost incompressible material model for breast tissue, and 50% reduction in phantom thickness. Second, projection images of compressed phantoms are simulated using a ray-tracing method. The x-ray image acquisition model assumes a mono-energetic x-ray beam with the energy of 20 keV and an ideal detector with 100μm pixel size. The quantum noise is simulated by Poisson random variations and added to all simulated images.

2.4. Statistical analysis of phantom data

2.4.1. Analysis of dense tissue distributions

The dense tissue simulated using different methods are quantified using metrics defined in Huang et al [8], called Radial Glandular Fraction (RGF) and Coronal Glandular Fraction (CGF).

RGF and CGF are defined as

$$RGF_n(r) = \left\langle \frac{N_D(r, x)}{N_D(r, x) + N_A(r, x)} \right\rangle_{x \in n} \tag{3}$$

$$CGF(x) = \frac{N_D(x)}{N_D(x) + N_A(x)} \tag{4}$$

where N_D is the number of pixels labeled as dense tissue, N_A is the number of pixels labeled as adipose tissue, and n indicates portion of breast, $n \in$ {Posterior, Middle Breast, Anterior}.

The RGF is used to quantify the distribution of simulated dense tissue based on the distance from center of the coronal slice in each region breast region (Posterior, Middle, and Anterior), while the CGF is used to quantify the distribution of simulated dense tissue in the posterior to anterior direction.

2.4.2. Analysis of covariance in simulated Images

Projection images were simulated for software phantoms with dense tissues labeled using either Gaussian or Beta distribution method with different sets of parameters. Covariance matrix elements [11] are defined as

$$K_{ij} = \langle (g_i - \overline{g_i})(g_j - \overline{g_j}) \rangle \ , \qquad (5)$$

where g_i and g_j are pixel values at the positions whose covariance is being estimated. The covariance matrix of each image set is assumed to be stationary; i.e. the covariance is independent of the location of their ROIs. We estimated covariance matrix elements along two orthogonal directions (chest-to-nipple and top-to-bottom) using ROIs of 4.35cm × 4.35cm in the regions of simulated images with constant thickness of the compressed phantom. A total number of 25 windows (1.45cm × 1.45cm) of 50% overlap in each ROI were used to calculate covariance matrix elements. To compare the covariance matrices of images simulated using different methods and parameters as well as with clinical data, the full width at half maximum (FWHM) of the average normalized covariance was used as the metric.

2.5. Materials

We compared the statistical properties of the phantom data to the clinical data reported in the literature [8] [11]. All phantoms were simulated with a breast volume of 450ml with resolution of 200µm per voxel. Each phantom contains 333 compartments randomly located and oriented inside the breast. Three phantom instances were simulated for each pair of distribution method/parameter and VBD value. The distribution parameters: σ in Gaussian method, and (p_1, q_1, p_2, q_2) in Beta method, were chosen manually based on user experience. Two sets of beta distribution parameters with different sidedness were chosen in order to interrogate the effects of the sidedness in beta functions. Table 1 shows the three sets of parameters used in the study.

Table 1. Distribution methods and parameter values used in the study.

Distribution methods	Parameter values
Gaussian	σ = 5.0
Beta1	p_1=2.0, q_1=0.5, p_2=2.0, q_2=3.5
Beta2	p_1=4.0, q_1=0.5, p_2=1.0, q_2=4.5

In our study, we created a total of 27 phantoms which consisted of 3 phantoms for each pair of VBD (20, 30, and 40%) and parameter in Table 1.

3. RESULTS

3.1. Probability maps of simulated dense tissue distribution

Probability values of dense tissue plotted on phantom surface provide a useful insight on the spatial characteristics of the method. Figure 3 shows the probability maps of phantoms created using the distribution parameters in Table 1. The probability map of Beta1 indicates a more uniformly distribution of probability, while Beta2 indicates a more concentrated probability near the nipple.
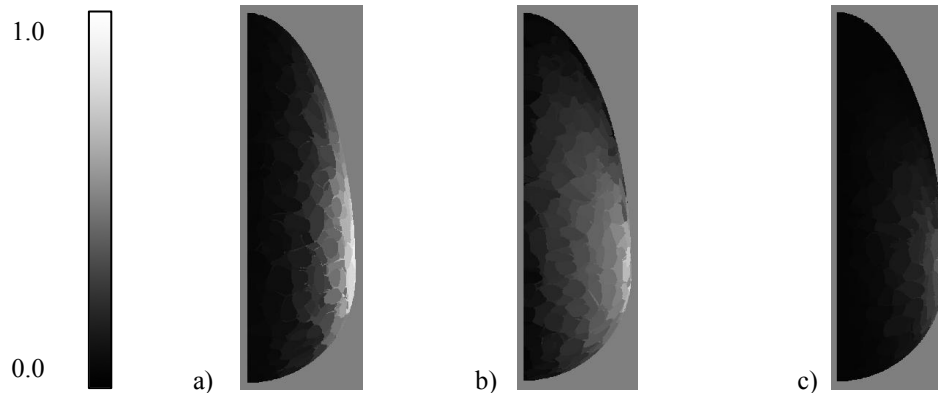
Figure 3: Probability maps of compartments labeled as dense tissues: a) Gaussian b) Beta1 and c) Beta 2.

3.2. Dense tissue simulation

Based on the probability values, calculated for each phantom compartment using the selected simulated method, the compartments are randomly labeled as containing dense or adipose tissue. Figure 4 shows the examples of phantoms of the same definition of compartment locations and orientations with dense tissue simulated using three sets of distribution parameters in Table 1. Compared to Gaussian and Beta2, dense tissue simulated using Beta1 are more distributed inside the breast.
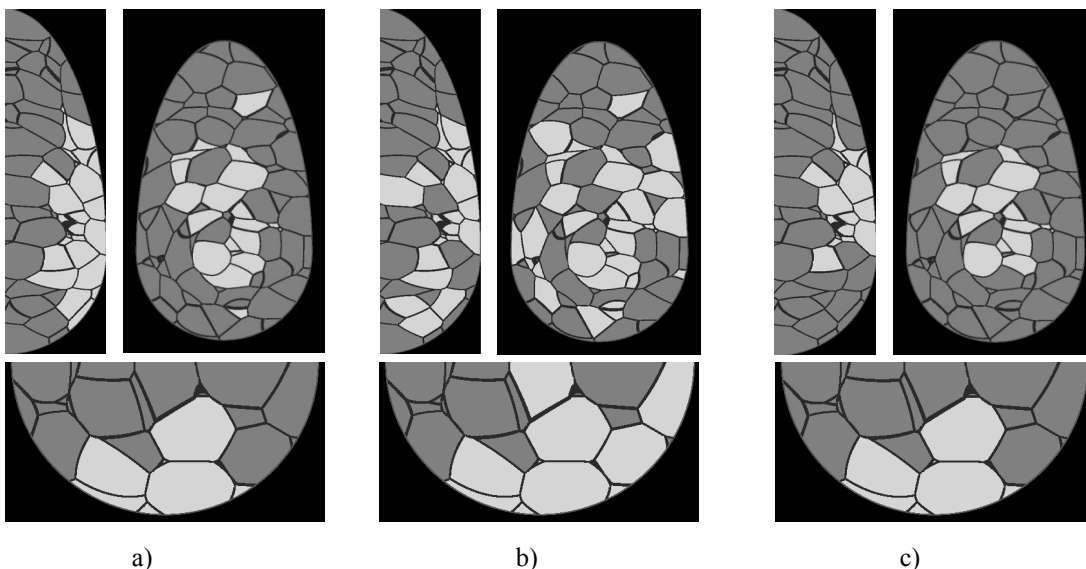

Figure 4: Phantoms with dense tissue simulated using a) Gaussian; b) Beta1; and c) Beta2.

3.3. RGF

The analysis of RGF is intended to compare dense tissue distributions simulated using different methods and parameters to clinical data, based on the radial distance from the center of the coronal plane. In order to have a close comparison with the clinical data, each breast phantom was divided into three equal thickness regions (Posterior, Middle, and Anterior). The RGFs were then measured separately in each of these three regions. Figure 5 a) to c) show the average RGFs of simulated phantoms using the three pre-defined parameters. Figure 5 d) to f) (the dash lines) show the average RGFs measured from the clinical data by Huang et al [8]. It was observed that the RGFs of the three distribution parameters in Table 1 result in similar trend as clinical data. Among the three distribution parameters, Beta2 method results in the best fit to the clinical data qualitatively.
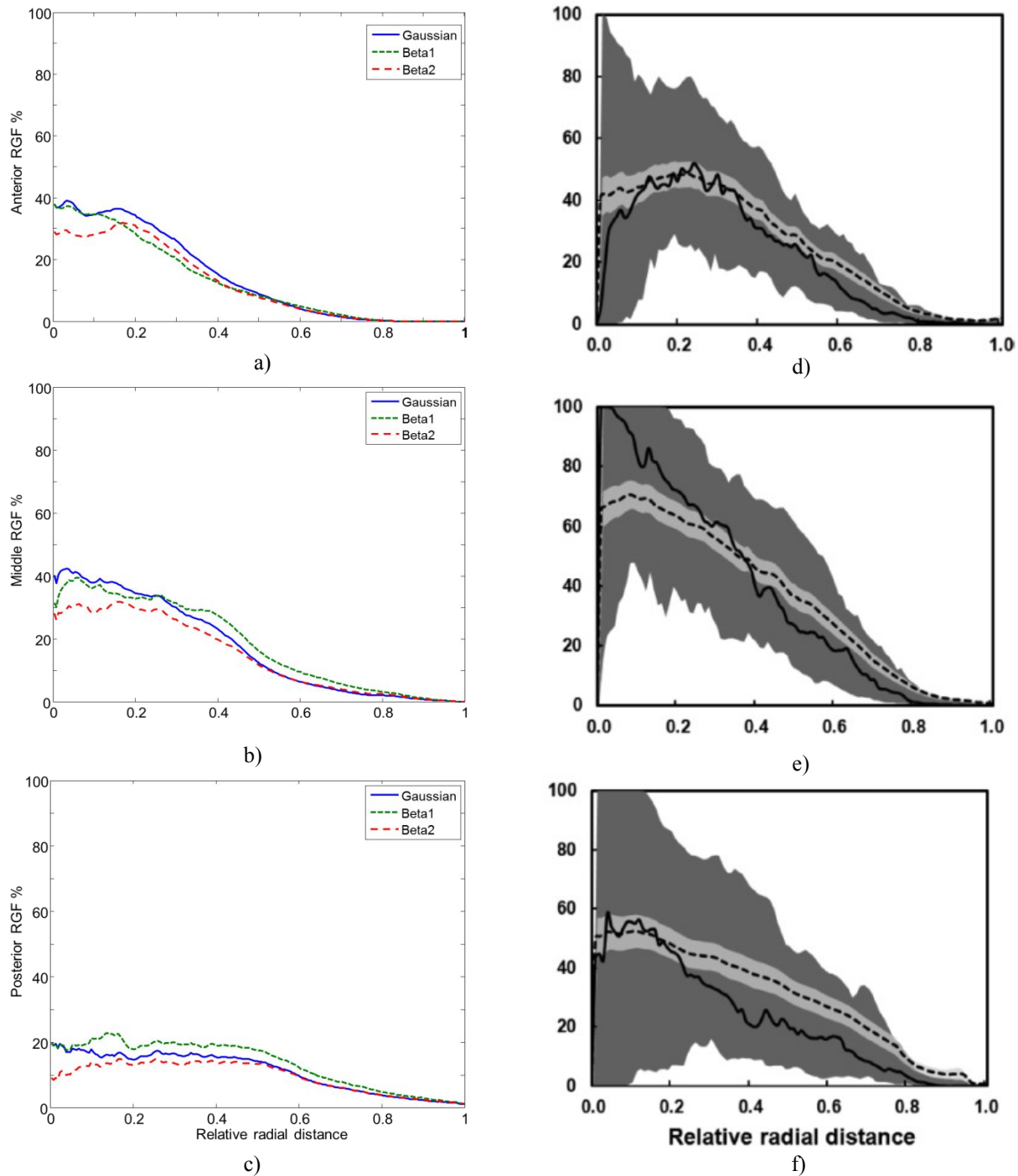
Figure 5: Average RGFs estimated from the simulated and clinical data. a)–c) The average RGFs in anterior, middle and posterior regions of simulated data, respectively. d)—f) The average RGFs in the respective regions, estimated from clinical data; (reprinted with permission from Huang et al. [8]). The total and 25-75 percentile regions are indicated by dark and bright colors, respectively. The mean and median values are indicated by dash and solid lines, respectively.

3.4. CGF

The analysis of CGF is intended to compare dense tissue distributions simulated using different methods and parameters to clinical data, based on the distance from the chest wall. Figure 6 a) shows the average CGFs measured from simulated distribution parameters in Table 1, and the dash line in Figure 6 b) shows the average CGFs measured from clinical data. Similar to the clinical data, the average CGFs measured from the simulated data increases with the coronal distance from the chest wall.
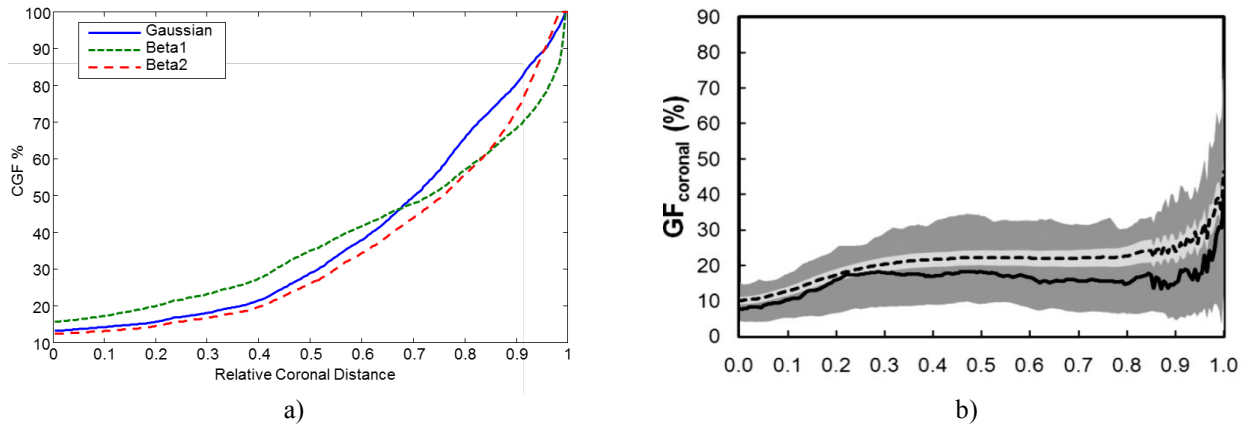

Figure 6: a) Average CGFs estimated from simulated data. b) Average CGFs (dash line) estimated from clinical data; (reprinted with permission from Huang et al [8].)

3.5. Simulated phantom images

Software phantoms created with different distribution methods and parameters are deformed to simulate breast compression during mammography acquisitions. Figure 7 shows the simulated acquisitions of phantoms with dense tissue created using the parameters in Table 1. Compared to Gaussian and Beta2, the dense tissue is more widely distributed in the image simulated using Beta1.
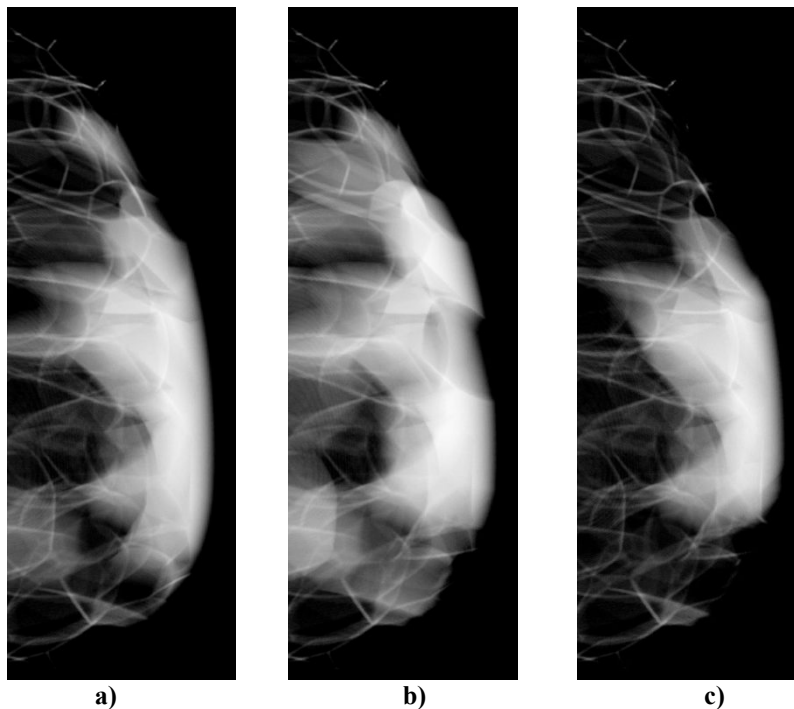

Figure 7: Simulated x-ray acquisition of phantoms using a) Gaussian method; b) Beta1 method; and c) Beta2 method.

3.6. Covariance analysis of simulated images

Average normalized covariance matrices measured from simulated acquisitions are shown in Figure 8 as function of the relative distance. The relative distance is equal to the spatial distance normalized by the window size used for estimating the covariance in the simulated images. Two windows are completely overlapped when the relative distance is 0, while only one row or column of pixels are overlapped when the relative distance is 1 or -1. The FWHMs of the average normalized covariance matrices are 0.381 (Gaussian), 0.433 (Beta1) and 0.344 (Beta2) for posterior-to-anterior direction; and 0.296 (Gaussian), 0.366 (Beta1) and 0.237 (Beta2) for top-to-bottom direction.
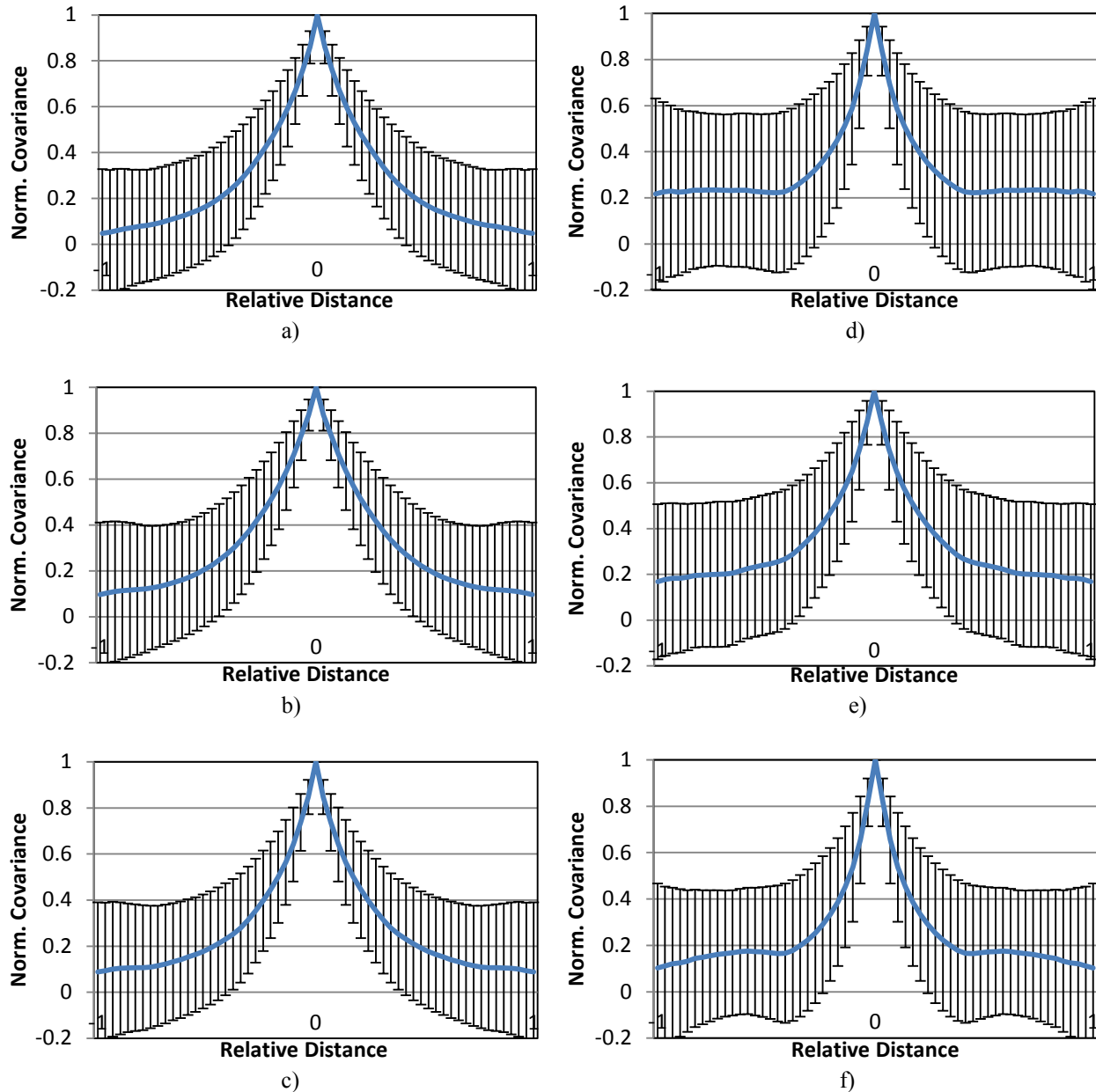
Figure 8: The profiles of average normalized covariance matrix measured from simulated acquisitions. Figure a) to c) are the profiles in posterior-to-anterior direction measured from a) Gaussian method; b) Beta1; and c) Beta2 method. Figures d) to f) are the respective profiles in top-to-bottom direction.

3.7. Comparing covariance profiles between simulated and clinical data

The covariance profiles in both posterior-to-anterior and top-to-bottom directions measured in simulated data are compared to ones measured from clinical data [11], using FWHM as the metric. Among the three distribution parameters in Table 1, Beta1 most closely matches the clinical data in both directions. Figure 9 shows the covariance profiles estimated from phantom images and clinical data. (Similar as in Figure 8, the relative distance is calculated as normalized by the window size used for the covariance calculation in the simulated images.)
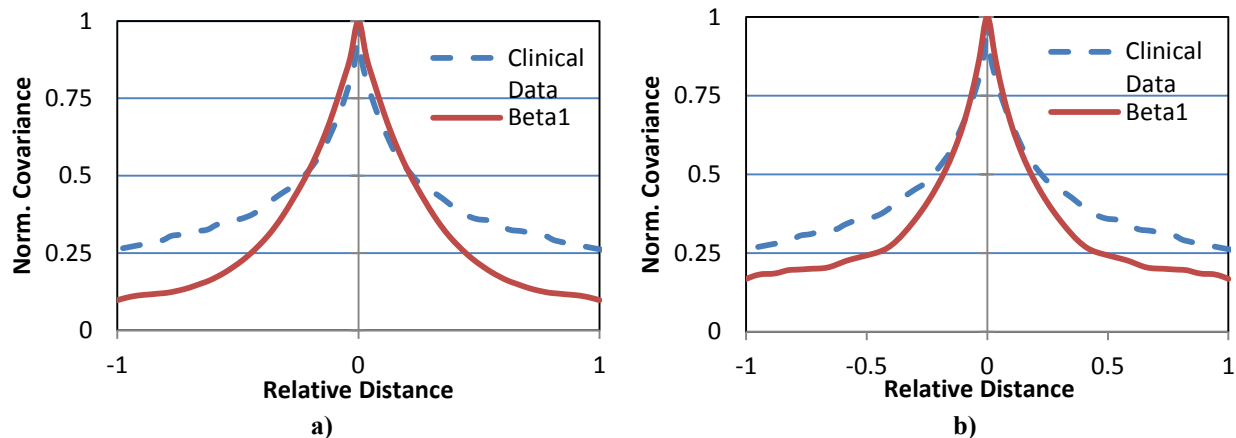

Figure 9: Profiles of average normalized covariance matrices in a) posterior-to-anterior direction; and b) top-to-bottom direction from simulated data created with Beta1 method and clinical data (modified from Freed et al. [11]). FWHMs measured from clinical data are 0.450(posterior-to-anterior) and 0.466 (top-to-bottom). FWHMs measured from simulated data using Beta1 are 0.433 (posterior-to-anterior) and 0.366 (top-to-bottom).

4. DISCUSSION

We implemented and compared the simulations of dense tissue using two different methods. Comparing to the Gaussian method using the Cartesian distance from the nipple, the Beta method separates the distance into radial and coronal distance. The use of beta functions offers higher control in the shape of distribution function such as its skewness and sidedness. The combination of extra flexibility in direction and the distribution functions provides additional freedom for the user to control the result of the dense tissue simulation.

The statistical properties such as RGF, CGF, and covariance matrices from data simulated using different methods and parameters shows Beta method has a better match to clinical data compared to Gaussian method. Careful optimization of the parameters in Gaussian method and in Beta method would be desirable to improve the matching between statistical properties measured in simulated and clinical data. Computing methods such as simulated annealing [12] and genetic programming [13] could be utilized for the tasks of parameter optimizations.

When comparing the RGFs of simulated versus clinical data near anterior breast region, a sudden drop off of glandular fraction is observed in the clinical data (Figure 5 a) and d)). The drop off is likely caused by the existence of subcutaneous fat around the nipple area. The existing simulation model used in our study does not correctly model the subcutaneous fat.

We observe a difference in normalized covariance between simulated and clinical data when the sampling windows are distant from each other. We believe that this could be the result of two factors. First it could be caused by the difference of ROI sizes used in simulated and clinical data. Smaller ROIs are used in measuring the simulated data, because of a restriction of the smaller region where uniform thickness exists in the compressed phantom. Second, there could be long distance correlations in the clinical data that are not modeled in the simulation.

5. CONCLUSION

We have proposed a novel method to assign dense compartments based upon Beta distributions. The new method offers better user control in spatial directions and shape of distribution function. We compared the simulated results with clinical images using CGF and RGF, and showed qualitative agreement. Future work includes quantitative evaluation of the agreement and selection of optimal distribution parameters.

ACKNOWLEDGEMENT

REFERENCES

[1] Wolfe, J.N.: Breast patterns as an index of risk for developing breast cancer. American Journal of Roentgenology 126, 1130-1139 (1976).

[2] Wolfe, J.N.: Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 37, 2486-2492 (1976).

[3] Boyd, N.F., Rommens, J.M., Vogt, K., Lee, V., Hopper, J.L., Yaffe, M.J. and Paterson, A.D.: Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol 6, 798-808 (2005).

[4] Li, H., Giger, M.L., Olopade, O.I., Margolis, A., Lan, L., Chinander, M.R.: Computerized texture analysis of mammographic parenchymal patterns of digitized mammograms. Academic Radiology 12, 863-873 (2005).

[5] Huo, Z., Giger, M.L., Olopade, O.I., Wolverton, D.E., Weber, B.L., Metz, C.E. and Zhong, W., Cummings, S.A.: Computerized analysis of digitized mammograms of brca1 and brca2 gene mutation carriers. Radiology 225, 519-526 (2002)

[6] Torres-Mejia, G., De, S.B., Allen, D.S., Perez-Gavilan, J.J., Ferreira, J.M., Fentiman, I.S. and Dos, S.S. I.: Mammographic features and subsequent risk of breast cancer: A comparison of qualitative and quantitative evaluations in the Guernsey prospective studies. Cancer Epidemiol Biomarkers Prev 14,1052-1059 (2005)

[7] Bakic, P.R., Carton, A.K., Kontos, D., Zhang, C., Troxel, A.B. and Maidment, A.D.: Breast percent density estimation from mammograms and central tomosynthesis projections. Radiology 2009-252(1):40-9.

[8] Huang, S.Y., Boone, J.M., Yang, K., Packard, N.J., McKenney, S.E., Prionas, N. D., Lindfors, K.K., Yaffe, M.J., The characterization of breast anatomical metrics using dedicated breast CT. Medical Physics 384, 2180-91 (2011).

[9] Pokrajac, D.D., Maidment, A.D.A., and Bakic, P.R. "Optimized generation of high resolution breast anthropomorphic software phantoms," Medical Physics, vol. 39, 2290-2302 (2012).

[10] Ruiter, N.V., Zhang, C., Bakic, P.R., Carton, A.-K., Kuo, J., Maidment, A.D.A.: "Simulation of tomosynthesis images based on an anthropomorphic breast tissue software phantom," In Visualization, Image-guided Procedures, and Modeling, Proc. SPIE 6918, edited by M.I. Miga, K.R. Cleary (2008).

[11] Freed, M., Badal, A., Jennings, R.J., De Las Heras, H., Myers, K. J., Badano, A., "X-ray properties of an anthropomorphic breast phantom for MRI and x-ray imaging," Phys Med Biol., vol. 56, 3513-33 (2011).

[12] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., "Optimization by Simulated Annealing," Science, vol. 220, no. 4598, 671–680 (1983).

[13] Banzhaf, W., Nordin, P., Keller, R.E., and Francone, F.D., Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications, Morgan Kaufmann (1998).