Statistical Parameter Inference of a 3D Solid Breast Texture Model

Zhijin Li, Ann-Katherine Carton, Thomas Almecija, Serge Muller, Pablo Milioni de Carlvalo and Agnès Desolneux

***Abstract*—The abstract goes here.**

***Index Terms*—X-ray breast imaging, virtual clinical trials, 3D breast texture model, stochastic geometry, statistical parameter inference.**

I. Introduction and motivation

SINCE clinical imaging studies are expensive and time consuming, there has been an increased research investment into less costly and more time-efficient simulation studies ([1], section 12.5). Reliable simulation studies in the field of digital breast tomosynthesis (DBT), and more generally in 3D imaging of the breast, require a realistic numerical 3D anthropomorphic test object model, a realistic numerical image acqui-sition chain and task-based mathematical observer models.

Several numerical 3D anthropomorphic breast object models have been proposed [2] [3] [4] [5] [6] [7]. Each model presents advantages and restrictions in terms of the visual realism of breast object images, the consistency between statistical features of test object and clinical images, and the potential to mimic the anatomical variability seen in real breast images. These aspects may impact the model validity for certain performance assessment studies.

Previously, we proposed a 3D solid breast texture model that is capable of simulating a large variability of realistic 2D and 3D breast images in terms of small and medium scale fibroglandular and inter-glandular adipose tissue in breast [8]. The proposed model is inspired by the morphology and distribution of medium and small scale fibroglandular and inter-glandular adipose tissue observed in clinical breast computerized tomography (bCT) reconstructed volumes. The reconstructed bCT volumes are stacks of gray-value coronal slices reconstructed with a filtered back-projection algorithm from projections images acquired using a prototype bCT machine developed at the University of California Davis Medical Center [9]. The volumes were pre-processed by an automatic 3D segmentation algorithm developed in-house to obtain volumes depicting breast fibroglandular and adipose tissue [10]. Figure 1 shows an example of bCT reconstructed volume and its 3D segmentation.

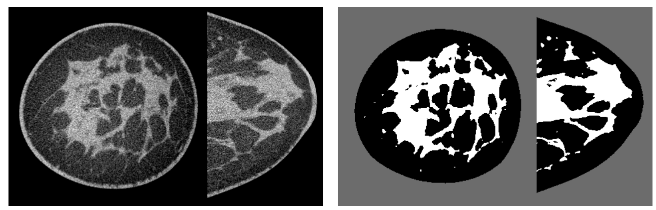

Fig. 1. Left: coronal and sagittal slices from an example of bCT reconstructed volumes. Right: coronal and sagittal slices from segmented volume of the same bCT volume, depicting fibroglandular tissue (white regions) and adipose tissue (black regions).

The previously proposed 3D solid breast texture model uses stochastic geometric elements to simulate the morphology and distribution of the medium and small scale breast fibroglandular and intra-glandular adipose tissue. For medium scale breast tissue, we used a system of stochastically distributed overlapping ellipsoids to depict the adipose compartments, and the complement of the ellipsoid system to depict the fibroglandular tissue. Small scale intra-glandular adipose compartment irregularities were introduced by replacing the smooth ellipsoid boundaries by Voronoi cells with average volume less than $1\,\mathrm{mm}^3$. The texture model outputs medium size rectangular volumes in voxelized binary format, with typical side length between $3.5\,\mathrm{cm}$ and $5\,\mathrm{cm}$. Figure 2 illustrates the construction of the proposed texture model.

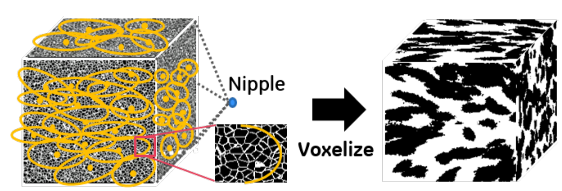

Fig. 2. Illustration of the construction of the proposed 3D solid breast texture model. The medium intra-glandular adipose compartments are modeled by a system of random overlapping ellipsoids. Small scale intra-glandular adipose compartment irregularities were introduced by replacing the smooth ellipsoid boundaries by Voronoi cells. The texture model produces voxelized medium size rectangular binary volumes depicting fibroglandular tissue (white regions in the image on the right) and intra-glandular adipose tissue (black regions in the image on the right).

Zhijin Li is with centre de mathématique et leurs applications of Ecole Normale Supérieure de Paris-Sacaly, Cachan, 94235 France and General Electric Healthcare France, Buc, 78530, France. E-mail: jonathan.li@ge.com.

Agnès Desolneux is with centre de mathématique et leurs applications of Ecole Normale Supérieure de Paris-Sacaly, Cachan, 94235 France.

Ann-Katherine Carton, Thomas Almecija, Serge Muller and Pablo Milioni de Carlvalo are with General Electric Healthcare France, Buc, 78530, France.

Manuscript received December, XXXX; revised January XX, XXXX.

In our previous study, we proposed a prototype model implementation where we subjectively determined the texture model parameters by visually comparing the sampled texture volumes with reference bCT volumes. A preliminary two-alternative forced choice experiment was performed showing the capability of the subjectively determined texture model to simulate realistic DBT reconstructed slices representing a large variety of breast density types. However, further observations show that the simulated DBT slices using the subjective model parameters have a limited morphological variability within each breast density type. This might be due to the limitation that the subjectively determined model parameters might not objectively capture the variability of breast tissue in clinical bCT volumes. To allow the 3D solid breast texture model to simulate a larger variability of breast textures, we propose to infer the parameters of the 3D breast texture model from clinical bCT volumes using objective statistical inference approach. A larger texture variability is of interest in simulation studies, where the impact of various breast tissue types on clinical task performance needs to be understood.

II. Problem formulation

We start by mathematically formulating the texture model and the ground truth data. For this study, we focus on the inference of the medium scale model parameters, which are responsible for the distribution and morphology properties of the system of random ellipsoids. The reason to consider only the medium scale model is mainly due to the relatively low spatial resolution of the clinical bCT volumes, where the voxel size is typically between 0.2 mm and 0.4 mm. When the number of the small Voronoi cells is large, the size of the cells may be in the same order of magnitude as the voxel size in the bCT data. Due to this fact, the bCT data may not allow for an accurate estimation of the small scale model parameters.

A. The medium scale texture model

The medium scale texture model is formulated as a *marked point process* (MPP)

$$\mathbf{Y} = \{\mathbf{\Phi_s}, \boldsymbol{\theta}\},$$

defined on a product space $Y \subset \mathbb{R}^3 \times \mathbb{R}^6$ [11].

- $\mathbf{\Phi_s}$ is a simple point process with distribution $\mathbf{P^s}$. Its realization represents the center points of the ellipsoids.
- $\boldsymbol{\theta}$ is a random vector with distribution $\mathbf{P}^{\boldsymbol{\theta}}$. Its realization represents the parameter vector of an ellipsoid. That is, $\boldsymbol{\theta} = (L_a, L_b, L_c, \delta_{\phi_a}, \delta_{\phi_b}, \delta_{\phi_c})$, where L_a, L_b, L_c are the half lengths of the principle axes of the ellipsoids and $\delta_{\phi_x}, \delta_{\phi_y}, \delta_{\phi_z}$ are random tilt angles. For each ellipsoid, the random tilt angles are added to three extrinsic rotation angles along x, y and z axes, determined by the center of the ellipsoid and the pre-defined nipple position. In marked point process theory, $\boldsymbol{\theta}$ is often referred to as the *mark* of the point process $\mathbf{\Phi_s}$, and $\mathbf{P}^{\boldsymbol{\theta}}$ is often referred to as the *mark distribution* [11].

B. The ground truth

To limit the inference to the medium scale breast tissue, only volumes of interest (VOI) with size 3.5 cm × 3.5 cm × 3.5 cm of the segmented bCT volumes were considered. The VOIs were extracted from the center of the segmented bCT volumes, leaving at least 2 cm distance to the breast skin and the chest-wall, depending on the actual breast size. A collection of 16 sets of segmented clinical bCT VOIs were selected for this study. The selected VOIs have BI-RADS densities (5th Edition, [12]) that cover category a, b and c, and represent a subset of breasts with considerable variability in medium scale fibroglandular and interglandular adipose tissue. None of the VOIs contained a lesion.

Mathematically, a segmented bCT VOI can be expressed as a binary volume $\mathcal{D}$, where for any $x \in \Omega \subset \mathbb{Z}^3$,

$$\mathcal{D}(x) = \begin{cases} 1 & \text{if voxel } x \text{ represents fibroglandular tissue,} \\ 0 & \text{if voxel } x \text{ represents adipose tissue.} \end{cases} \tag{1}$$

Here Ω represents the 3D discrete spatial domain of $\mathcal{D}$.

Notice that, the input segmented clinical bCT VOIs depict adipose compartments, from which it might be difficult to discern the individual ellipsoids. This means that the positions of the ellipsoid centers are *unobservable* from the ground truth data $\mathcal{D}$. This unobservability makes it difficult to propose an appropriate parametric model for the ellipsoid center point process. Also, multiple configurations of ellipsoids may be used to approximate the adipose tissue in $\mathcal{D}$, making the inference problem ill-posed.

III. Existing parameter inference methods for marked point processes with unobservable point positions

Due to the data unobservability problem described in the previous section, direct application of common MPP parameter inference approaches such as the *likelihood*-based approaches [13] is impossible, since they require the knowledge of the positions of the ellipsoid centers [14]. Some classical inference methods in stochastic geometry have be investigated to mitigate the issue of data unobservability. One of such methods is referred to as the *minimum contrast estimator* (MCE).

Let $\mathbf{\Theta}$ denote the set of all parameters for $\mathbf{Y}$. Let $\mathcal{C}(\mathcal{D}, \mathbf{\Theta})$ denote a contrast function depending on the ground truth $\mathcal{D}$ and the parameters $\mathbf{\Theta}$. The MCE is defined as [11]:

$$\hat{\mathbf{\Theta}} = \text{argmin}_{\mathbf{\Theta}} \mathcal{C}(\mathcal{D}, \mathbf{\Theta}). \tag{2}$$

Often,

$$\mathcal{C}(\mathcal{D}, \mathbf{\Theta}) = \iint_{Y \times Y} \left(S(y_1, y_2; \mathbf{\Theta}) - \hat{S}(y_1, y_2; \mathcal{D}) \right)^2 \mathrm{d}y_1 \mathrm{d}y_2, \tag{3}$$

where $S(\cdot,\cdot;\mathbf{\Theta})$ is the analytical formula of one, or a weighted sum of several second-order summary statistics of $\mathbf{Y}$; and $\hat{S}(\cdot,\cdot;\mathcal{D})$ denotes the empirical counterpart of $S(\cdot,\cdot;\mathbf{\Theta})$ measured from the ground truth $\mathcal{D}$. Popular choices of summary statistics are the two-point function [15] and the contact distribution function [16] etc.

The MCE method has limitations that make its application in our case difficult. The major limitation is the difficulty in the derivation of analytical summary statistics when the underlying model $\mathbf{Y}$ becomes very complex. In this case, the derivation of analytical summary statistics may involve some numerical integration technique and the minimization can also be non-trivial.

IV. The method of inference from reconstruction

Recently, Thiedmann *et al.* proposed a novel inference method for parametric MPP models using a two-step approach that can effectively mitigate the data unobservability problem [17]. Inspired by their work, we hereby formally present this method and refer to it as the *inference from reconstruction* method.

The inference from reconstruction method consists of two steps:

1) *The reconstruction step* aims at recovering the unobserved point positions and marks from the observed data through stochastic sampling.
2) *The inference step* is a parametric inference step where a marked point process models is proposed and fitted to the reconstructed points and marks.

This two-step mechanism provides two main advantages over classical MCE method. First, once the point positions and marks are recovered from the reconstruction step, it is easier to perform statistical analysis to gain intuitions to determine which model should be used for the inference step. Second, the inference step using reconstructed point positions and marks becomes more straightforward since all point positions are observable. Due to these advantages, we decided to apply the inference from reconstruction method to fit a parametric MPP model to the ground truth segmented bCT data.

In the following section we describe in detail the methodology used for the reconstruction step and the inference step in our study.

A. Reconstruction step: approximating segmented bCT data using ellipsoids

Consider the case where the density function $f_{\mathbf{Y}}$ of the underlying MPP model $\mathbf{Y}$ has a Gibbsian form [18]. Let $\mathbf{u} \in Y$ be a set of ellipsoids, refered to as a configuration, then we have:

$$f_{\mathbf{Y}}(\mathbf{u}) = \frac{1}{Z}\exp\left(-\frac{1}{T}U(\mathbf{u})\right). \tag{4}$$

Here the term $U(\cdot)$ is called the *energy*. The parameter $T \in \mathbb{R}^+$ is often referred to as the temperature as in physics. The term $Z = \int \exp\left(-U(\mathbf{u})\right) \mathrm{d}\mathbf{u}$ is a normalization constant. The energy term is often decomposed into two parts [19] [20]:

$$U(\mathbf{u}) = \mathcal{L}(\mathbf{u},\mathcal{D}) + \mathcal{P}(\mathbf{u}), \tag{5}$$

where $\mathcal{L}(\mathbf{u},\mathcal{D})$ is a *data term* representing how well a configuration $\mathbf{u}$ matches the ground truth dataset $\mathcal{D}$; and $\mathcal{P}(\cdot)$ is a *prior term* containing the a-priori information of the underlying MPP model $\mathbf{Y}$. Finding the best set of ellipsoids $\mathbf{u}^*$ is equivalent to solving the following optimization problem:

$$\mathbf{u}^* = \arg\max_{\mathbf{u}} f_{\mathbf{Y}}(\mathbf{u}) = \arg\min_{\mathbf{u}} \left(\mathcal{L}(\mathbf{u},\mathcal{D}) + \mathcal{P}(\mathbf{u})\right). \tag{6}$$

We formulated $\mathcal{L}(\mathbf{u},\mathcal{D})$ and $\mathcal{P}(\mathbf{u})$ as follows:

- The *data term* $\mathcal{L}(\mathbf{u},\mathcal{D})$ consists of two terms $\mathcal{L}_1$ and $\mathcal{L}_2$. The first term $\mathcal{L}_1$ measures the approximation error ratio of the current configuration $\mathbf{u}$ given input $\mathcal{D}$. For a given ellipsoid $\mathcal{E} \in \mathbf{u}$, its approximation error ratio $d_{\mathcal{E}}(\mathcal{D})$ can be expressed as:

$$d_{\mathcal{E}}(\mathcal{D}) = \frac{|\{x \in \Omega | x \in \mathcal{E} \text{ and } \mathcal{D}(x) = 1\}|}{|\{x \in \Omega | x \in \mathcal{E}\}|}. \tag{7}$$

  Then $\mathcal{L}_1$ is defined by summing the errors of all ellipsoids. That is,

$$\mathcal{L}_1(\mathbf{u},\mathcal{D}) = \sum_{\mathcal{E}\in\mathbf{u}} d_{\mathcal{E}}(\mathcal{D}). \tag{8}$$

  The second term $\mathcal{L}_2$ measures the proportion of the input $\mathcal{D}$ that is not covered by the current configuration $\mathbf{u}$. That is,

$$\mathcal{L}_2(\mathbf{u},\mathcal{D}) = 1 - \frac{|x \in \Omega | x \in \mathbf{u} \text{ and } \mathcal{D}(x) = 0|}{|x \in \Omega | \mathcal{D}(x) = 0|}. \tag{9}$$

  The sum of $\mathcal{L}_1$ and $\mathcal{L}_2$ defines the final data term $\mathcal{L}$. That is,

$$\mathcal{L}(\mathbf{u},\mathcal{D}) = \mathcal{L}_1(\mathbf{u},\mathcal{D}) + \mathcal{L}_2(\mathbf{u},\mathcal{D}). \tag{10}$$

- Regarding the *prior term* $\mathcal{P}(\mathbf{u})$, we prefer to obtain a model estimate without imposing too much a-priori information on the distribution of the ellipsoids. Hence, only a weak constraint on the overlap ratio between ellipsoids is used to formulate $\mathcal{P}(\mathbf{u})$. That is:

$$\mathcal{P}(\mathbf{u}) = \sum_{\mathcal{E}\in\mathbf{u}} q(\mathcal{E}, \mathbf{u}\setminus\mathcal{E}), \tag{11}$$

  where

$$q(\mathcal{E}, \mathbf{u}\setminus\mathcal{E}) = \begin{cases} 0 & \text{if } \frac{|\{x\in\Omega|x\in\mathcal{E} \text{ and } x\in\mathbf{u}\setminus\mathcal{E}\}|}{|\{x\in\Omega|x\in\mathcal{E}\}|} \leq 0.95, \\ +\infty & \text{otherwise.} \end{cases} \tag{12}$$

The optimization given by (6) finally becomes:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \mathcal{L}_2(\mathbf{u},\mathcal{D}) + \sum_{\mathcal{E}\in\mathbf{u}} \left(d_{\mathcal{E}}(\mathcal{D}) + q(\mathcal{E}, \mathbf{u}\setminus\mathcal{E})\right) \tag{13}$$

The analytical solution of (13) is difficult to obtain in pratice. In MPP literature, the *reversible jump Markov chain Monte Carlo* (RJMCMC) sampling is the most commonly employed technique to tackle this type of optimization problem [21] [22] [23] [20] [24] [19] [25] [26]. A typical RJMCMC procedure consists of iteratively simulating a Markov chain of configurations $\{\mathbf{u}_t\}_{t\in\mathbb{N}}$ whose density converges to the target density $f_{\mathbf{Y}}(\mathbf{u})$. At each iteration t, a modification of the current configuration $\mathbf{u}_t$ is proposed to create the next configuration $\mathbf{u}_{t+1}$. The term “jump” refers to the fact that the cardinality, *i.e.* the number of marked points of the current configuration might change during the modification. A modification in RJMCMC is performed according to a density function $Q(\mathbf{u}_t, \mathbf{u}_{t+1})$, referred to as a proposition kernel. The modifications are local, in the sense that for each iteration only one or two marked points in the current configuration are modified. Typically, $Q(\mathbf{u}, \cdot)$ is a combination of several sub-proposition kernels:

$$Q(\mathbf{u}, \cdot) = \sum_n p_n Q_n(\mathbf{u}, \cdot), \tag{14}$$

where p_n is the probability of the occurrence of the sub-proposition kernels $Q_n(\cdot, \cdot)$, such that $\sum_n p_n \leq 1$ [27] [28]. Frequently investigated sub-proposition kernels include:

- The *Birth proposal*, in which a marked point is added to current configuration $\mathbf{u}_t$, according to a birth kernel denoted as $Q_b(\mathbf{u}, \cdot)$. That is, $\mathbf{u}_{t+1} = \mathbf{u}_t \cup \{u\}$, where u is drawn according to $Q_b(\mathbf{u}, \cdot)$.
- The *Death proposal*. This is the reverse process of the birth proposal, in which a marked point is chosen and deleted from current configuration $\mathbf{u}$ according to a death kernel $Q_d(\mathbf{u}, \cdot)$. That is, $\mathbf{u}_{t+1} = \mathbf{u}_t \setminus \{u\}$, with $u \in \mathbf{u}_t$ chosen according to $Q_d(\mathbf{u}, \cdot)$.
  The combination of birth and death proposals ensures that the Markov chain is able to switch between configurations with different cardinality.
- *Perturbation proposal*. This proposal consists of changing the parametrization of a marked point u in the current configuration $\mathbf{u}$ according to a perturbation kernel $Q_p(\mathbf{u}, \cdot)$.

In practice, the RJMCMC procedure might suffer from prohibitive rate of convergence, since each iteration brings only one or two marked points into play [27]. To mitigate this issue, a *multiple births and deaths* algorithm has recently been proposed by Descombes *et al.* [20], allowing for the births of multiple marked points at each iteration, offering a more effective sampling procedure. In our study, we adapted the multiple births and deaths algorithm design proposed in [20]. For each iteration, the original multiple births and deaths algorithm consisted of a birth step of multiple marked points and a death step that examines all marked points in current configuration. Additionally, to achieve a more effective exploration of each configuration, we proposed to add an extra perturbation step to the multiple births and deaths algorithm, named as the *shift* step.

To describe the shift step used in our adapted multiple births, deaths and shifts (MBDS) algorithm, we adopt the notion of *Legendre ellipsoid* of a convex body in classical mechanics [29]. Given a volume $K \subset \mathbb{Z}^3$, the Legendre ellipsoid $\mathcal{L}(K)$ is the unique ellipsoid defined as [30]:

$$\mathcal{L}(K) = \{x \in K | x^T \Sigma^{-1} x \leq 1\}, \tag{15}$$

where

$$\Sigma = \frac{\sum_{x\in K}(x-\mu)(x-\mu)^T}{|K|}, \tag{16}$$

with $\mu = \frac{1}{|K|}\sum_{x\in K} x$. Notice that $\mathcal{L}(K) = K$ if K is itself an ellipsoid.

In our proposed MBDS algorithm, the shift of an ellipsoid $\mathcal{E}$ consists in replacing it by the Legendre ellipsoid computed from the part of the adipose tissue of the observed data $\mathcal{D}$ inside $\mathcal{E}$. That is,

$$\mathcal{E} \to \mathcal{L}(K_{\mathcal{E}}(\mathcal{D})) \text{ with } K_{\mathcal{E}}(\mathcal{D}) = \{x \in \Omega | x \in \mathcal{E} \text{ and } \mathcal{D}(x) = 0\}. \tag{17}$$

The complete description of the proposed MBDS algorithm is given by Algorithm I. We assumed that the proposal distribution $f_{\boldsymbol{\theta}}$ consists of independent densities f_{L_a}, f_{L_b}, f_{L_c}, $f_{\delta\phi_1}$, $f_{\delta\phi_2}$ and $f_{\delta\phi_3}$ to sample the half lengths L_a, L_b, L_c and the tilt angles $\delta\phi_x$, $\delta\phi_y$, $\delta\phi_z$. We set these densities to the ones used in our previous study [8] for the BI-RADS breast density b type breasts because these previously validated values might provide a good starting-point for the MBDS algorithm and might accelerate the convergence of the algorithm.

B. Inference step: fitting a marked point process model to reconstructed ellipsoids

Once a segmented bCT VOI is represented as a system of ellipsoids with known spatial positions and shape parameters, we fit a parametric MPP model to the reconstructed ellipsoids. To reduce the complexity of the inference step, we assume in the first instance that the marks of the MPP model are independent from each other and that they are independent from the ellipsoid centers. This allows us to fit the center point process of the ellipsoids and their shape parameters separately.

1) The point process for reconstructed ellipsoid centers: To gain some intuition on the type of point process model $\mathbf{\Phi_s}$ for the fit, we first analyzed the *pair correlation function* (PCF) of the reconstructed ellipsoid centers for each ground truth data set. The PCF is a commonly studied second-order statistical descriptor in stochastic geometry that can help reveal comprehensive structural information of a point process [31].

Definition IV.1. *The pair correlation function $g(\cdot, \cdot)$ of a simple point process $\mathbf{\Phi}$ with intensity function $\lambda(\cdot)$ is defined as:*

$$g(x, y) = \frac{\rho^{(2)}(x, y)}{\lambda(x)\lambda(y)}. \tag{21}$$

Here $\rho^{(2)}(\cdot, \cdot)$ is the second order moment density of $\mathbf{\Phi}$, satisfying

$$\int_{B_1\times B_2} \rho^{(2)}(x, y)\, \mathrm{d}x\mathrm{d}y = \sum_{x,y\in\mathbf{\Phi}}^{x\neq y} \mathbb{E}\left(\mathbb{1}(\{x, y\} \in B_1 \times B_2)\right), \tag{22}$$

Algorithm: Multiple births, deaths and shifts

- **Initialization**.
  Set the initial configuration $\mathbf{u}_0 = \emptyset$, the initial temperature $T = 100$. Let $\mathbf{\Phi}$ be a Poisson point process with initial intensity $\lambda = 0.005$. Let $f_{\boldsymbol{\theta}}$ be a multivariate proposal distribution used to sample $(L_a, L_b, L_c, \delta\phi_1, \delta\phi_2, \delta\phi_3)$, *i.e.* the marks.
- **Iterations**
  Iterate the following steps in order, until convergence is reached.
  - **Multiple births**
    Generate a random configuration u_b with ellipsoid centers drawn from $\mathbf{\Phi}$ and ellipsoid parameters drawn from $f_{\boldsymbol{\theta}}$. Update current configuration $\mathbf{u}$:
    $$\mathbf{u} \to \mathbf{u} \cup u_b. \tag{18}$$
  - **Computation of death probabilities**
    For each ellipsoid $\mathcal{E} \in \mathbf{u}$, the death probability p_d is given by [20]:
    $$p_d = \frac{r\lambda}{1 + r\lambda}, \tag{19}$$
    with $r = \exp\left(\frac{U(\mathbf{u}) - U(\mathbf{u}\setminus\mathcal{E})}{T}\right)$, where $U(\cdot)$ is given in (13).
  - **Deaths and shifts**
    For each ellipsoid $\mathcal{E} \in \mathbf{u}$, draw a random variable $v \sim \text{uniform}(0, 1)$.
    if $v < p_d$, perform a death operation: $\mathbf{u} \to \mathbf{u} \setminus \mathcal{E}$
    else, shift $\mathcal{E}$ to its Legendre ellipsoid: $\mathcal{E} \to \mathcal{L}(\mathcal{K}_{\mathcal{E}(\mathcal{D})})$, as described in (17).
  - **Update parameters**
    Decrease the Poisson intensity λ and the temperature T:
    $$\lambda \to \lambda \cdot 0.99 \text{ and } T \to T \cdot 0.99, \tag{20}$$
- **Test of convergence**
  For each iteration t, record its energy value and the energy values of its next nine successive iterations. This yields a set of ten energy values $E_t = \{U_t, U_{t+1}, \cdots, U_{t+9}\}$. Then compute $\max E_t$ and $\min E_t$. If for a predefined small real number ϵ, $\max E_t - \min E_t \le \epsilon$, then the convergence is considered reached.

TABLE I
Algorithm I

for arbitrary bounded Borel sets B_1 and B_2.

The PCF can be used to interpret the interaction between points in a point process [32]. For all Poisson processes where there are no interactions between points, $g(x, y) = 1$. If $g(x, y) > 1$, an *attraction* between points at locations x and y exists. If $g(x, y) < 1$, a *repulsion* between points at locations x and y exists.

In our study, we assumed that the centers of the reconstructed ellipsoids come from a stationary and isotropic point process. This indicates that the PCF depends only on the relative distance between two spatial positions. That is, $g(x, y) = g(r)$, where $r = \|x - y\|$ is referred to as the *interpoint distance*. Under this assumption, we applied the PCF estimator described in [32, p232] to estimate the empirical PCF of reconstructed ellipsoid centers from each segmented bCT VOI. Analytically, the PCF estimator is expressed as:

$$\hat{g}(r; \Phi_s) = \sum_{x,y \in \Phi_s \cap W}^{x \neq y} \frac{\mathbf{k}(\|x - y\| - r)}{4\pi r^2 \nu(W_x \cap W_y)\hat{\lambda}^2}, \tag{23}$$

where Φ_s is the collection of all ellipsoid centers reconstructed from a dataset and $\hat{\lambda}$ is an estimate of its intensity parameter, expressed as

$$\hat{\lambda} = \frac{|\Phi_s|}{\nu(W)}. \tag{24}$$

The function $\mathbf{k}(\cdot)$ is a smoothing kernel. We use $\nu(\cdot)$ to denote the volume measure and W is the observation window; *i.e.* a $3.5\,\text{cm} \times 3.5\,\text{cm} \times 3.5\,\text{cm}$ cube in our case. Finally, W_x denotes the translation of W by x. The division by $\nu(W_x \cap W_y)$ instead of by $\nu(W)$ acts as an edge correction for points falling outside of the observation window W [33].

The estimation was performed using the `pcf3est` function implemented in the `R` software package `spatstat` [34] with its default setting. In this setting the *Epanechnikov* smoothing kernel [11] was used. Mathematically, the *Epanechnikov* kernel is defined as:

$$\mathbf{k}(s) = \begin{cases} \frac{3}{4\delta}(1 - \frac{s^2}{\delta^2}) & \text{if } -\delta \le s \le \delta, \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

It has a bandwidth parameter δ to tune. In the default setting of `pcf3est`, δ is set according to the rule-of-thumb: $\delta = \frac{0.26}{\sqrt[3]{\hat{\lambda}}}$ [34].

Once the empirical estimates of PCFs were obtained, we first checked if the Poisson property can or can not be rejected based on the PCF estimates. This check was performed using the envelope test described in [35]. We will show in Section V-B that the analysis of the PCFs revealed a clustering interaction between reconstructed ellipsoid centers for the majority of the cases. To model the clustering interaction, we proposed to fit a three-dimensional *Matérn cluster process* [31] to the reconstructed ellipsoid centers. The 3D Matérn cluster process is a two-step process described as follows.

1) First, a set of "parent points" $\{y_i\}_{i\in\mathcal{I}} \subset \mathbb{R}^3$ are sampled from a homogeneous Poisson point process with intensity parameter κ. For each "parent point" y_i with $i \in \mathcal{I}$, a sphere with radius R centered at y_i is generated.
2) Then, inside each obtained sphere, a set of "children points" are sampled from another homogeneous Poisson point process with intensity parameter λ_0. A realization of the Matérn cluster process is obtained as the collection of all "children points".

A Matérn cluster process model $\Phi_{\mathcal{M}}$ is a stationary and isotropic point process completely determined by the three parameters: κ, λ_0, and R. Theoretical formula for the PCF of a 3D Matérn cluster process is analytically accessible and is given by the following proposition [32, p376].

Proposition IV.1. *Let $\Phi_{\mathcal{M}}$ be a Matérn cluster process with parameters κ, λ, R, then its intensity parameter λ is:*

$$\lambda = \frac{4}{3}\pi R^3 \kappa \lambda_0. \tag{26}$$

Its pair correlation function g is:

$$g(r;\kappa,\lambda_0,R) = \begin{cases} 1 + \frac{3(R-\frac{r}{2})^2(2R+\frac{r}{2})}{8\pi\kappa R^6} & \text{if } 0 < r \leq 2R, \\ 1 & \text{if } r > 2R. \end{cases} \tag{27}$$

To fit a Matérn cluster process with parameters κ, λ, R to a set of reconstructed ellipsoid centers Φ_s, we applied the MCE method described in Section III. For a given Φ_s, the contrast function $\mathcal{C}$ was computed based on the analytical and empirical PCF. That is,

$$\mathcal{C}(r;\kappa,\lambda_0,R,\Phi_s) = \sum_{r\in\mathcal{R}} (g(r;\kappa,\lambda_0,R) - \hat{g}(r;\Phi_s))^2, \tag{28}$$

where $\hat{g}(r,\Phi_s)$ is given in (23) and $g(r;\kappa,\lambda_0,R)$ is given in (27). Here $\mathcal{R}$ denotes the set of interpoint distances considered by the MCE estimator. Additionally, (26) was used as an equality constraint. Let $\boldsymbol{\Theta} = (\kappa, \lambda_0, R)$, for a given set of reconstructed ellipsoids Φ_s, the MCE estimator $\widehat{\boldsymbol{\Theta}}$ is formally expressed as:

$$\begin{aligned} \widehat{\boldsymbol{\Theta}} &= \arg\min_{\boldsymbol{\Theta}} \sum_{r\in\mathcal{R}} \left(g(r;\kappa,\lambda_0,R) - \hat{g}(r;\Phi_s)\right)^2, \\ &\text{subjected to: } \hat{\lambda} = \frac{4}{3}\pi R^3\kappa\lambda_0. \end{aligned} \tag{29}$$

The optimization (29) was numerically solved using the function `fmincon` implemented in the `Matlab` software (version `2016b`, The MathWorks Inc., Natick, Massachusetts, United States). The default setting of `fmincon` was used. In this setting, the *interior point* optimization method is applied, with the Hessian of the contrast function $\mathcal{C}(\cdot)$ approximated using the *Broyden–Fletcher–Goldfarb–Shanno* algorithm. We set $\mathcal{R}$ to be a set of values increasing from 0.2 mm to 30 mm with a step size of 0.2 mm. All optimizations were run with an initial condition $\kappa = 0.1$, $\lambda_0 = 0.1$ and $R = 1$. The convergence was considered reached when the contrast function was non-decreasing in all feasible directions, within the a tolerance value of 1×10^{-6}. Since we were able to obtain fairly good fit for all input datasets (Section IV-B1), the impact of different configurations of the optimization function was considered out-of-scope in our study and was not investigated.

2) The mark distribution: Empirical statistics of each individual mark were examined separately from the ellipsoid centers. Histograms of the half lengths L_a, L_b, L_c and the tilts angles $\delta\phi_x$, $\delta\phi_y$, $\delta\phi_z$ of the reconstructed ellipsoids were obtained to visualize the empirical distributions of L_a, L_b and L_c. Based on the empirical statistical analysis, independent distributions were proposed for each mark.

V. Results

In this section, we demonstrate the results of the inference from reconstruction method focusing on examples of three bCT VOIs, referred to as VOI #1, #2 and #3 hereunder. We chose these three VOIs since their glandular densities belong to BI-RADS density category a, b and c respectively, and hence they represent a fairly large variability of medium scale fibroglandular and interglandular adipose breast tissue. The same analyses were performed for all 16 input datasets and similar results were obtained.

A. Reconstruction step

First, to demonstrate the convergence of the MBDS algorithm, Figure 3 shows the decrease of the energy with the increasing number of iterations for the three bCT VOIs. From this figure we can see that the convergence was reached after about 7500 iterations for all the cases. Similar results were obtained for other investigated bCT VOIs.

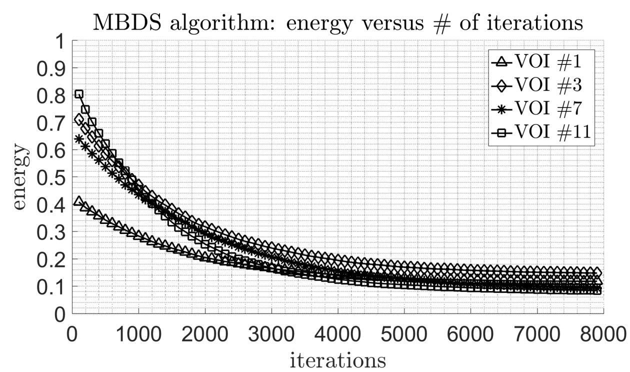

Fig. 3. Illustration of the energy defined in (5) as a function of the number of iterations in the MBDS algorithm. The illustration considers the four examples of segmented clinical bCT VOIs shown in Figure 4. The convergence of the MBDS algorithm was reached after about 7500 iterations for all the cases.

The result of the reconstruction step for VOI #1, #2 and #3 is demonstrated in Figure 4 by comparing the original VOIs with the reconstructed VOIs. Reconstructed VOIs are binary volumes having the same spatial and voxel size as their corresponding bCT VOIs. They were created by voxelizing the reconstructed ellipsoids from the MBDS algorithm and assigning value 0 to the ellipsoid interior. The coronal and sagittal slices of the original VOIs and the reconstructed VOIs at the same positions are shown in Figure 4. Moreover, projections images of the original VOIs and the reconstructed VOIs are also demonstrated. The projections were obtained by averaging the VOIs in the direction perpendicular to the transverse plane.

From Figure 4 we can see that the reconstructions by ellipsoids are not perfect. Despite this, the distribution and morphology of the medium scale fibroglandular and interglandular adipose tissue in reconstructed VOIs agree fairly well with the original segmented clinical bCT VOIs. Also, the medium scale texture variations in the projections of the original VOIs are preserved in the projections of the reconstructed VOIs. The result of the reconstruction step provides sufficiently good input for follow-up inference step.

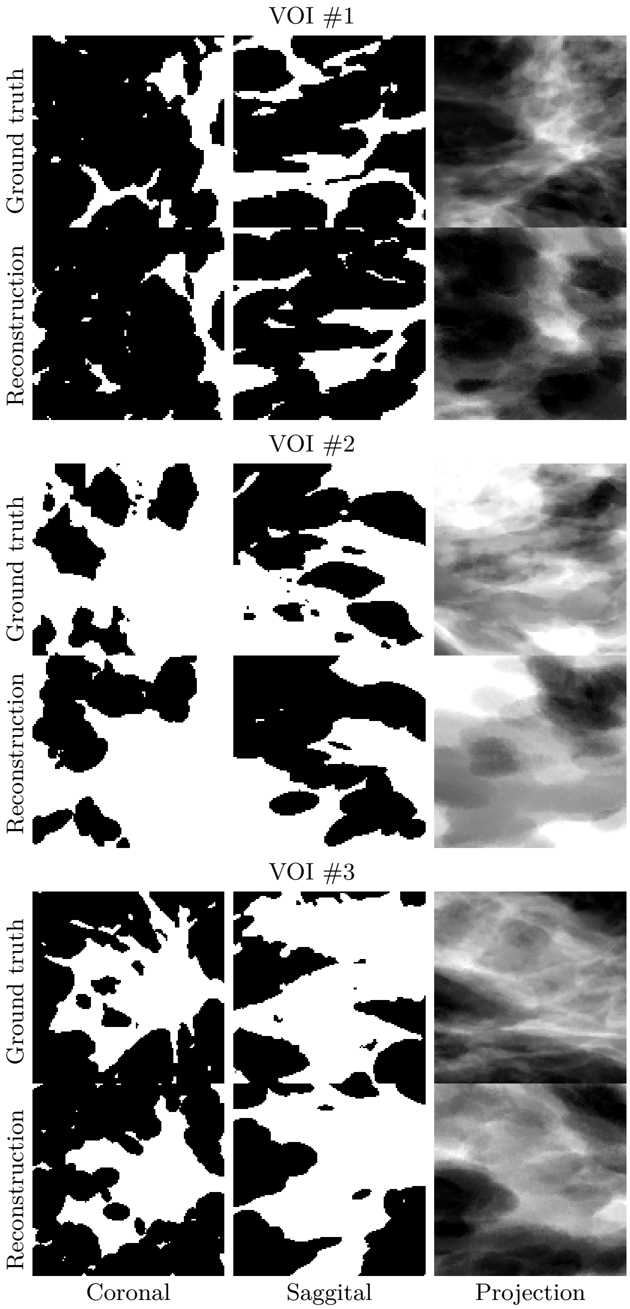

Fig. 4. Illustration of coronal slices, saggital slices and projections of VOIs #1, #2, #3, and their reconstructed VOIs. The sizes of the VOIs are 3.5 cm × 3.5 cm × 3.5 cm. Reconstructed VOIs are binary volumes with the same size and resolution as their corresponding segmented clinical bCT VOIs. They were created by voxelizing the reconstructed ellipsoids from the MBDS algorithm and assigning value 0 to the ellipsoid interior. Projections images were obtained by averaging the VOIs in the direction perpendicular to the transverse plane. The distribution and morphology of the medium scale fibroglandular and inter-glandular adipose tissue in reconstructed VOIs agree fairly well with the original segmented clinical bCT VOIs

B. Inference step

By analyzing the estimated PCFs of the reconstructed ellipsoid centers from the 16 input bCT VOIs using the envelope test method described in [35], we found that for four VOIs, the Poisson null-hypothesis can not be rejected. For the remainder of our study, these VOIs were excluded since they represented a much smaller population among all bCT VOIs.

The three figures on the left in Figure 5 compare the empirical PCFs estimated from the reconstructed ellipsoid centers for VOI #1, #2, #3 and the theoretical Poisson PCFs. The envelopes of the Poisson PCFs generated using the method described in [35] are also shown. Figure 5 indicates that the empirical PCFs estimated from the reconstructed ellipsoid centers of the three VOIs exhibit a clustering effect. This can be seen if we look at the interpoint distance $r = 3\,\mathrm{mm}$ emphasized in Figure 5. The envelope test rejects the null-hypothesis for all demonstrated VOIs at $r = 3\,\mathrm{mm}$. The same clustering effect was found for all other investigated bCT VOIs.

1) Parameters of the 3D Matérn cluster process: As discussed in Section IV-B1, a 3D Matérn cluster process was fitted to each set of reconstructed ellipsoid centers using the MCE method. The three figures on the right in Figure 5 compare the empirical PCFs estimated from the reconstructed ellipsoid centers for VOI #1, #2, #3 and the theoretical PCFs of the Matérn cluster processes fitted to the same VOIs. The PCF envelopes of the fitted Matérn cluster processes were also generated using the method described in [35], but with the fitted Matérn cluster process as the null hypothesis model. It can be seen that for the three VOIs, the the empirically estimated PCFs fall inside the envelopes at all interpoint distances. Based on this, we conclude that the fits are fairly good. Similar result was found for all other investigated bCT VOIs.

2) The mark distribution: Figure 6 shows the histograms of the half lengths L_a, L_b, L_c and the tilt angles $\delta\phi_x$, $\delta\phi_y$, $\delta\phi_z$ of the reconstructed ellipsoids for VOI #1, #2 and #3. From visual observations, we notice that the distributions of L_a, L_b, L_c, $\delta\phi_y$ and $\delta\phi_z$ seem to be Gaussian and the distribution of $\delta\phi_x$ seems to be uniform between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Similar observations were also obtained for reconstructed ellipsoids from other bCT VOIs. Following these observations, we then fitted the distributions of L_a, L_b, L_c, $\delta\phi_y$ and $\delta\phi_z$ to Gaussian distributions using the maximum likelihood estimator. A series of Kolmogorov-Smirnov tests were also performed to verify the statistical validity of the fit. The tests show that for more 80% of the cases, the fit was good statistically under 5% confidence interval.

As a summary, Table I lists the fitted medium scale parameters for VOI #1, #2 and #3. By applying the inference from reconstruction method on all the 12 input bCT VOIs, 12 new sets of medium scale parameters were finally obtained.

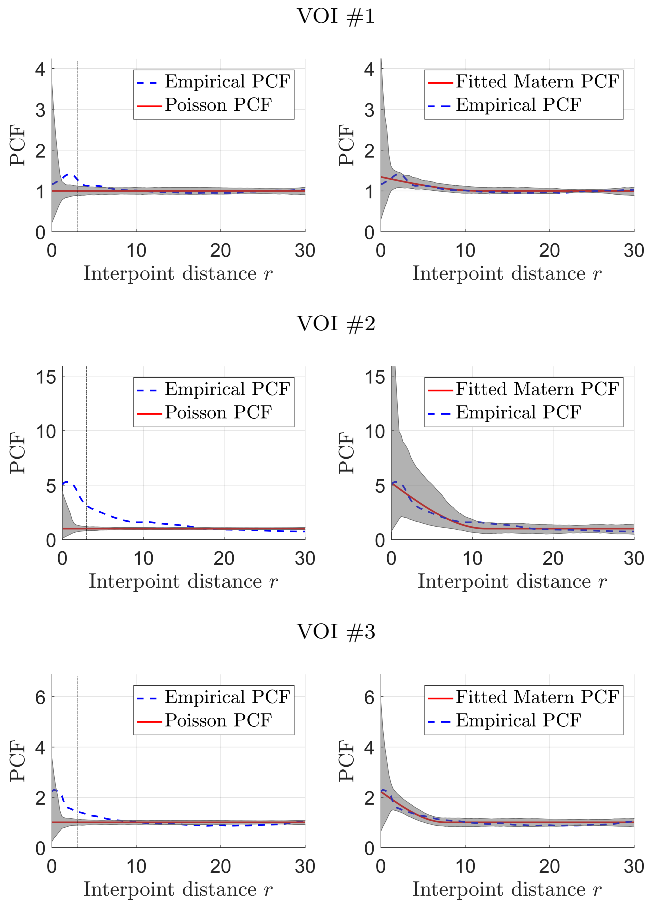

Fig. 5. Figures on the left compare the empirical PCFs (dashed lines) estimated from the reconstructed ellipsoid centers for VOI #1, #2, #3 and the theoretical Poisson PCFs (solid lines). Gray surfaces are the envelopes of the Poisson PCFs generated using the method described in [35]. The envelope test rejects the Poisson null-hypothesis for all four VOIs at interpoint distance $r = 3\,\text{mm}$. Figures on the right compares the same empirical PCFs and the theoretical PCFs (solid lines) of the Matérn cluster processes fitted to the same VOIs. Gray surfaces are the envelopes of the PCFs of the fitted Matérn cluster processes generated using the method described in [35]. The empirical PCFs fall inside the envelopes at all interpoint distances, indicating good fits.

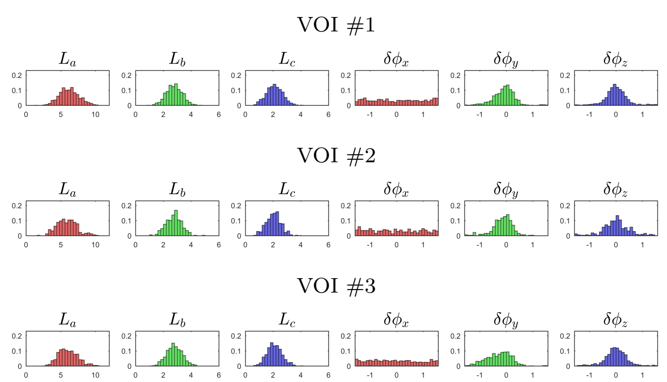

Fig. 6. Histograms of the half lengths L_a, L_b, L_c and the tilt angles $\delta\phi_x$, $\delta\phi_y$, $\delta\phi_z$ of reconstructed ellipsoids for VOI #1, #2 and #3. Visual observations show that these parameters seem to follow either Gaussian or uniform distributions.

VI. Validation of inferred medium scale texture model parameters

A. Visual inspection

To visually inspect the realism of simulated x-ray breast images using the fitted medium scale texture model parameters, a simulation experiment was conducted. Three texture volumes were simulated using respectively the three set of medium scale parameters listed in Table I. The average size of the small Voronoi cells was set to be $0.1\,\text{mm}^3$. The nipple position and the voxel size for each volume were set to be the same as in the corresponding ground truth bCT data. Mammograms and DBT projection images were simulated by virtually projecting the texture volumes using the previously described breast x-ray imaging simulator [36]. The x-ray simulator was adjusted to model the topology of the GE SenoClaire DBT imaging system. Each VOI was placed at the center-left of the detector, such that the left edge of the VOI cube was aligned with the detector's left edge. A mono-energetic x-ray source at 20 keV was used to simulate nine DBT projection images with an angular range of $(-12.5°, 12.5°)$. Only Poisson x-ray noise was modeled and the electronic noise was not modeled. X-ray scattering and blurring on the detector were not modeled. The size of the isotropic detector pixels is 100 µm. Mechanical breast texture deformation to mimic breast compression during the DBT image acquisition was not modeled. Projection images were processed by the `ASIR-DBT` 3D reconstruction algorithm (GE Healthcare, Buc, France) to obtain DBT reconstructed slices with 1 mm slice thickness.

TABLE I
Medium scale parameters for the 3D stochastic breast texture model fitted from VOIs #1, #2 and #3. Here $\mathcal{N}(\mu, \sigma)$ denotes the Gaussian distribution with mean μ, standard deviation σ; and $\mathcal{U}(l, u)$ denotes the uniform distribution with lower bound l and upper bound u.

Parameters	VOI #1	VOI #2	VOI #3
κ (mm^{-3})	$3.24e-03$	$3.41e-03$	$2.87e-04$
λ_0 (mm^{-3})	$5.98e-03$	$1.92e-02$	$3.09e-02$
R (mm)	5.98	3.85	5.82
p_{L_a} (mm)	$\mathcal{N}(6.21, 1.41)$	$\mathcal{N}(5.88, 1.44)$	$\mathcal{N}(6.06, 1.39)$
p_{L_b} (mm)	$\mathcal{N}(2.77, 0.58)$	$\mathcal{N}(2.75, 0.56)$	$\mathcal{N}(2.79, 0.56)$
p_{L_c} (mm)	$\mathcal{N}(2.10, 0.57)$	$\mathcal{N}(2.03, 0.52)$	$\mathcal{N}(2.10, 0.54)$
$p_{\delta_{\phi_x}}$ (rad)	$\mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$	$\mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$	$\mathcal{U}(-\frac{\pi}{2}, \frac{\pi}{2})$
$p_{\delta_{\phi_y}}$ (rad)	$\mathcal{N}(-0.09, 0.4)$	$\mathcal{N}(-0.15, 0.38)$	$\mathcal{N}(-0.28, 0.47)$
$p_{\delta_{\phi_z}}$ (rad)	$\mathcal{N}(0, 0.26)$	$\mathcal{N}(0.01, 0.5)$	$\mathcal{N}(-0.01, 0.43)$

Figure 7 shows examples of 3.5 cm × 3.5 cm slices through the simulated texture volumes, as well as mammographic projections and DBT reconstructed slices simulated from these volumes. Visual inspection of the simulated images indicates fairly high visual realism. Also, compared with images simulated using the prototype implementation [8], the new model parameters are capable of simulating mammographic projections and and DBT reconstructed slices with a larger morphological variability. These observations indicate an improvement of the model's realism and morphological variability compared with the prototype implementation with empirical parameters proposed in the previous chapter.

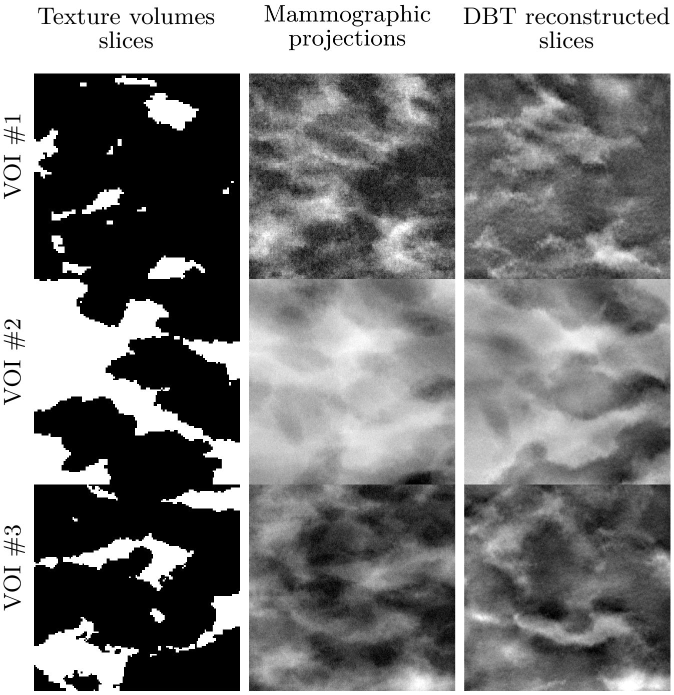

Fig. 7. The first column shows slices through volumes simulated from the 3D breast texture model with the four sets of parameters listed in Table I. The second column shows mammographic projections simulated from the simulated texture volumes. The third column shows DBT reconstructed slices simulated from the simulated texture volumes. The sizes of the images are 3.5 cm × 3.5 cm.

B. Statistical validation

As a statistical validation of the realism of simulated breast images using the new model parameters, we performed statistical estimations of the β metric on simulated breast images. The β metric measures the inverse slope of the log-scale noise power spectrum (NPS) of an image and is the most commonly used statistical metric for the realism of simulated breast images. For each of the obtained 12 sets of medium model parameters, 100 texture volumes were simulated under the same set-up as described in the previous section. From each simulated texture volume, 3.5 cm × 3.5 cm mammographic projection and DBT reconstructed slices were obtained using the same x-ray simulator and 3D reconstruction algorithm as described in previous section. We used the method described by Wu *et al.* [37] to estimate the β values from the mammographic projection and DBT slices. In this method, each input image was first divided into 2.5 cm × 2.5 cm regions of interest (ROI) with 50% overlap. Then the ROIs were multiplied by a Hann window and their average power spectrum was used as the NPS estimate. Finally, a linear regression of the NPS versus the frequency in log scale was performed within a frequency range of $\left[0.1\,\mathrm{mm}^{-1}, 0.7\,\mathrm{mm}^{-1}\right]$, and the inverse of the regression slope was used as the β estimate. We found that the β values of the simulated mammography projections have an average of XXX and a standard deviation of XXX; the β values of the simulated DBT central slices have an average of XXX and a standard deviation of XXX. These values are consistent with those reported in literature [REFERENCE].

C. Psychovisual validation validation

A two-alternative forced choice (2AFC) experiment was performed to formally assess the visual realism of simulated DBT reconstructed images. Pairs of $3.5 \times 3.5\,\mathrm{cm}^2$ ROIs, extracted from images simulated from our texture model and clinical bCT data were displayed side-by-side. Images on the left always came from the clinical bCT data set, while images on the right had 50% chance to be from the texture model and 50% chance to be from clinical bCT data. The reader had to tell whether the image on the right was from the clinical bCT data set or from the texture model. Similar level of glandular density was maintained for each image pair. In total, 144 image pairs were presented; 52 pairs from BI-RADS density category a, 56 pairs from BI-RADS density category b and 36 pairs from BI-RADS density category c. Images were displayed on 5M pixels grayscale portrait monitors (SMD 21500 G, Siemens AG; Munchen, Germany) at 100% resolution. Each image pair was displayed during 5 seconds, followed by the display of a uniform gray–level image for another 5 seconds; the readers were thus imposed to make a decision within 10 seconds. Four readers participated in the experiment, all GE Healthcare engineers. Reader 1, 2 and 3 have no prior knowledge of our texture model while reader 4 knows the algorithm of the texture model construction. A short training session with 10 image pairs of known ground truth was performed before the real experiment with no time constraint. The reading distance was set to be one meter. The percentage of correct answers, P_c, was calculated as an indication for the realism of the simulated images from the texture model. Under the null hypothesis that images simulated from the model and clinical bCT images data cannot be distinguished, $P_c = 0.5$.

TO BE MODIFIED [The left chart of Figure 5 shows the Pc value of the four readers of the 2AFC experiments. Readers 1 and 2 had overall Pc values of 0.65 and 0.66, while reader 3 had a Pc value of 0.73 indicating that he could more easily distinguish texture images from clinical bCT images. Reader 4 had the highest Pc value of 0.82. He reported that he was able to infer simulated images from his prior knowledge of the intermediate steps of the model construction. Readers 1, 3 and 4 reported that they had observed a quantification effect appearing like "small square blocks" in some simulated images from the texture model, as illustrated in the right image of Figure 5. When this effect was present, they used it as criterion to distinguish simulated images using the texture model from those using clinical bCT data. Reader 2 did not report the same primary criteria. This 'small square block artifact' can likely be attributed to the size of the Voronoi cells at the transition of glandular/adipose tissue. Table 2 reports Pc for each reader by type of breast density. The p-value indicates whether the reader was or was not able to differentiate simulated data from real data. The lower Pc values for al-most entirely adipose breasts can be explained by the lower number of adipose/dense transitions, designed by Voronoi cells, in this type of breasts.]

VII. Conclusion and discussion

In this paper, we applied a novel statistical inference methodology referred to as the inference from reconstruction to objectively infer the medium scale parameters of a previously introduced 3D breast texture model from $3.5\,\mathrm{cm} \times 3.5\,\mathrm{cm} \times 3.5\,\mathrm{cm}$ volumes of interest (VOI) of segmented clinical breast computerized tomography (bCT) reconstructed datasets.

A multiple births, deaths and shifts (MBDS) algorithm was employed to first reconstruct a set of random ellipsoids from each ground truth bCT VOI. Visual inspection of the volumes recreated by voxelizing the reconstructed ellipsoids shows a fairly good approximation of the medium scale breast tissue in the original bCT VOIs. A Matérn cluster process was then fitted to the reconstructed ellipsoid centers using the minimum contrast method based on the pair correlation functions (PCF). This introduces a clustering interaction between ellipsoids to the previous prototype model. Statistical diagnostic analysis using the PCF suggested a fairly good fit of the reconstructed ellipsoid centers to the proposed Matérn cluster process. Distributions of the ellipsoid half lengths and orientations were finally estimated from their empirical distributions. Twelve sets of new medium scale model parameters were obtained. Preliminary evaluation of the 2D and 3D breast images simulated from 3D texture volumes generated using new parameters shows fairly high visual realism. The breast tissue variability in new simulated images is larger than the images simulated using previously proposed prototype implementation with empirical parameters. The values of the β metric measure from simulated mammographic projections and DBT slices are consistent with values reported in literature. Finally, a formal two-alternative forced choice XXXX.

The proposed method has several limitations. First, the impact of the reconstruction step on the estimated parameters in the inference step was not investigated. To address this, one possible approach is to study how the inference result deviates from a ground truth model when the latter is a known Poisson marked point process. Second, to reduce the optimization complexity in the inference step, the distribution of ellipsoid half lengths were estimated independently from the ellipsoid centers. This may be a simplification compared with the distribution of inter-glandular adipose compartments in real breasts. The correlation between the the half lengths and the centers of the ellipsoids could be further investigated by studying the mark correlation function of the reconstructed ellipsoids.

Acknowledgment

This research is partially funded by the French national association for research and technology (ANRT) under the CIFRE grant n°2013/1052. The authors would like to thank Prof. John Boone for kindly providing the clinical bCT data sets.

References

[1] M. K. Markey, *Physics of mammographic imaging.* CRC press, 2012.

[2] C. M. Li, W. P. Segars, G. D. Tourassi, J. M. Boone, and J. T. Dobbins, “Methodology for generating a 3d computerized breast phantom from empirical data,” *Medical physics*, vol. 36, no. 7, pp. 3122–3131, 2009.

[3] D. M. Mahr, R. Bhargava, and M. F. Insana, “Three-dimensional in silico breast phantoms for multimodal image simulations,” *IEEE transactions on medical imaging*, vol. 31, no. 3, pp. 689–697, 2012.

[4] A.-K. Carton, A. Grisey, P. M. de Carvalho, C. Dromain, and S. Muller, “A virtual human breast phantom using surface meshes and geometric internal structures,” in *International Workshop on Digital Mammography*, pp. 356–363, Springer, 2014.

[5] P. R. Bakic, D. D. Pokrajac, R. De Caro, and A. D. Maidment, “Realistic simulation of breast tissue microstructure in software anthropomorphic phantoms,” in *International Workshop on Digital Mammography*, pp. 348–355, Springer, 2014.

[6] F. Chen, P. R. Bakic, A. D. Maidment, S. T. Jensen, X. Shi, and D. D. Pokrajac, “Description and characterization of a novel method for partial volume simulation in software breast phantoms,” *IEEE transactions on medical imaging*, vol. 34, no. 10, pp. 2146–2161, 2015.

[7] C. G. Graff, “A new open-source multi-modality digital breast phantom,” in *SPIE Medical Imaging*, pp. 978309–978309, International Society for Optics and Photonics, 2016.

[8] Z. Li, A. Desolneux, S. Muller, and A.-K. Carton, “A novel 3d stochastic solid breast texture model for x-ray breast imaging,” in *International Workshop on Digital Mammography*, pp. 660–667, Springer, 2016.

[9] K. K. Lindfors, J. M. Boone, T. R. Nelson, K. Yang, A. L. Kwan, and D. F. Miller, “Dedicated breast ct: Initial clinical experience 1,” *Radiology*, vol. 246, no. 3, pp. 725–733, 2008.

[10] T. Almecija, “Development of an anthropomorphic software model of the female breast,” master thesis, Ecole Supérieure d’Electricité, 2015.

[11] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications.* John Wiley & Sons, 2013.

[12] C. J. D’Orsi, *ACR BI-RADS Atlas: Breast Imaging Reporting and Data System.* American College of Radiology, 2013.

[13] J. Moller and R. P. Waagepetersen, *Statistical inference and simulation for spatial point processes.* CRC Press, 2003.

[14] D. Dereudre, F. Lavancier, and K. S. Helisová, “Estimation of the intensity parameter of the germ-grain quermass-interaction model when the number of germs is not observed,” *Scandinavian Journal of Statistics*, vol. 41, no. 3, pp. 809–829, 2014.

[15] P. J. Diggle, “Binary mosaics and the spatial pattern of heather,” *Biometrics*, pp. 531–539, 1981.

[16] L. Heinrich, “Asymptotic properties of minimum contrast estimators for parameters of boolean models,” *Metrika*, vol. 40, no. 1, pp. 67–94, 1993.

[17] R. Thiedmann, O. Stenzel, A. Spettl, P. R. Shearing, S. J. Harris, N. P. Brandon, and V. Schmidt, “Stochastic simulation model for the 3d morphology of composite materials in li–ion batteries,” *Computational Materials Science*, vol. 50, no. 12, pp. 3365–3376, 2011.

[18] J. Møller and R. P. Waagepetersen, “Modern statistics for spatial point processes,” *Scandinavian Journal of Statistics*, vol. 34, no. 4, pp. 643–684, 2007.

[19] F. Lafarge, G. Gimel farb, and X. Descombes, “Geometric feature extraction by a multimarked point process,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 9, pp. 1597–1609, 2010.

[20] X. Descombes, R. Minlos, and E. Zhizhina, "Object extraction using a stochastic birth-and-death dynamics in continuum," *Journal of Mathematical Imaging and Vision*, vol. 33, no. 3, pp. 347–359, 2009.

[21] C. J. Geyer and J. Møller, "Simulation procedures and likelihood inference for spatial point processes," *Scandinavian journal of statistics*, pp. 359–373, 1994.

[22] P. J. Green, "Reversible jump markov chain monte carlo computation and bayesian model determination," *Biometrika*, pp. 711–732, 1995.

[23] Z. Tu and S.-C. Zhu, "Image segmentation by data-driven markov chain monte carlo," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 657–673, 2002.

[24] W. Ge and R. T. Collins, "Marked point processes for crowd counting," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 2913–2920, IEEE, 2009.

[25] A. Utasi and C. Benedek, "A 3-d marked point process model for multi-view people detection," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 3385–3392, IEEE, 2011.

[26] K. Sirinukunwattana, D. R. Snead, and N. M. Rajpoot, "A stochastic polygons model for glandular structures in colon histology images," *IEEE transactions on medical imaging*, vol. 34, no. 11, pp. 2366–2378, 2015.

[27] X. Descombes, *Stochastic geometry for image analysis*. John Wiley & Sons, 2013.

[28] Y. Verdié and F. Lafarge, "Detecting parametric objects in large scenes by monte carlo sampling," *International Journal of Computer Vision*, vol. 106, no. 1, pp. 57–75, 2014.

[29] E. Lutwak, D. Yang, and G. Zhang, "A new ellipsoid associated with convex bodies.," *Duke Mathematical Journal*, vol. 104, no. 3, pp. 375–390, 2000.

[30] M. Ludwig, "Ellipsoids and matrix-valued valuations," *Duke Mathematical Journal*, vol. 119, no. 1, pp. 159–188, 2003.

[31] A. Baddeley, I. Bárány, and R. Schneider, "Spatial point processes and their applications," *Stochastic Geometry: Lectures given at the CIME Summer School held in Martina Franca, Italy, September 13–18, 2004*, pp. 1–75, 2007.

[32] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical analysis and modelling of spatial point patterns*, vol. 70. John Wiley & Sons, 2008.

[33] J. Ohser, "On estimators for the reduced second moment measure of point processes," *Statistics: A Journal of Theoretical and Applied Statistics*, vol. 14, no. 1, pp. 63–71, 1983.

[34] A. Baddeley and R. Turner, "Spatstat: an r package for analyzing spatial point patterns," *Journal of statistical software*, vol. 12, no. 6, pp. 1–42, 2005.

[35] A. Baddeley, P. J. Diggle, A. Hardegen, T. Lawrence, R. K. Milne, and G. Nair, "On tests of spatial pattern based on simulation envelopes," *Ecological Monographs*, vol. 84, no. 3, pp. 477–489, 2014.

[36] P. Milioni de Carvalho, *Low-dose 3D quantitative vascular X-ray imaging of the breast*. PhD thesis, University Paris 11, 2014.

[37] G. Wu, J. G. Mainprize, and M. J. Yaffe, "Spectral analysis of mammographic images using a multitaper method," *Medical physics*, vol. 39, no. 2, pp. 801–810, 2012.