Hello everyone, thank you for coming to my PhD thesis defense. My thesis is entitled “Task-based optimization of 3D breast imaging using mathematical observer”. It is under the CIFRE contract. And it is a collaboration between CMLA, Ecole Normale Superieure de Paris-Saclay and GE Healthcare France.

Slide 2

Today, the key x-ray imaging modalities for breast cancer detection and diagnosis are full field digital mammography, or FFDM and digital breast tomosynthesis, or DBT.

FFDM is a 2D x-ray imaging technique, the breast is positioned on a support table, and then compressed. A contrasted image is obtained on the detector since different breast tissue attenuates differently the x rays. The main limitation of FFDM is breast tissue superposition, which can obscure lesions, and may create artifacts looking like lesions.

Due to these limitations, 3D breast imaging has been investigated to overcome the limitations. DBT is a 3D imaging technique. The x-ray tube moves along an arc in limited angular range to acquire multiple projections. With a reconstruction algorithm, a stack of slices parallel to the detector can be obtained. There is also an algorithm that allow to create a synthetic 2D mammogram, or S2D image from DBT reconstructed data.

The focus of this thesis is on 3D breast imaging. More specifically, we aim to assess the task-based clinical performance of 3D DBT versus the 2D FFDM.

Slide 3

For performance assessment of breast imaging technologies, clinical trials are essential. But they can be burdensome. Today many groups are investigating the potential of virtual clinical trials, or VCT, to offer a more efficient and less costly alternative.

In VCT, Patients are replaced by physical or simulated phantoms and lesions. The images are acquired using physical machine or an x-ray simulator modelling the physics. Images are reviewed by radiologist or by model observers, which are math models that aim to replace radiologists in specific clinical tasks, such as lesion detection. In VCT, the ground truth is stored in the computer

Slide 4

The main objective of this thesis is to assess the microcalcification detection performance in 3D DBT, and to compare it with 2D FFDM & S2D images using the VCT approach.

This main objective is driven by two motivations: First, variables results from existing clinical studies have been obtained regarding the comparison of 2D FFDM and 3D DBT in ucal detection performance. Secondly, limited research has been done regarding the ucal detection performance in S2D images versus FFDM.

To achieve the main objective, we also aim to develop and validate two VCT tools. First, we aim to develop & validate a new mathematical 3D breast texture model that allows to simulate realistic 2D & 3D breast x ray images, representative for the population of women seen in clinical practices. Then, we aim to develop & validate a new 3D model observer based on the *a contrario* theory for ucal detection in 2D & 3D breast images.

Slide 5

There are mainly 2 types of models used in literature to simulate x-ray breast images.

The first type of models is referred to as the 3D random field texture models. The commonly used models are 3D power-law Gaussian random field and 3D clustered lumpy background, also known as the shot-noise random field model. They are mathematical defined model and are characterized by their analytical formulations.

The other type of model is referred to as the anthropomorphic breast phantoms. These are 3D models that aim to simulated different breast anatomical structures. The model-based phantoms use geometric primitives such as ellipsoid, Voronoi diagrams for different anatomical components. The empirical data based phantoms use anatomical structures segmented from clinical images or mastectomy data. The hybrid phantoms may use different combinations of 3D random fields model-based and empirical data based phantoms.

Slide 6

Here I show you slices from 3D volumes simulated using a 3D power-law Gaussian random field on the left, and the 3D clustered lumpy background on the right. And here are the projection images created using these volumes. As a comparison, we also have a slice from a clinical breast CT reconstructed volume, and a clinical mammogram.

There are studies showing that some of the statistical characteristics of the mathematical random fields may match those of the clinical breast images; however, if we compared the images simulated using these models to the clinical images, we may find a limited realism and morphological variation.

Slide 7

Here is a mammography image created using a model based phantom that employs Voronoi digrams to simulated breast fibroglandular and adipose tissue. Here is an empirical based phantom that uses breast tissues segmented from clinical breast CT images. And here is a hybrid phantom that combines the first shown model based phantom with a 3D stochastic Perlin noise.

As we can see from these images, anthropomorphic breast phantoms offer a large variability of visual realism and morphological variations compared to clinical images. In general, these models are mathematically less traceable.

The goal of our first study is to develop a 3D breast texture model that can offer the advantages of both random fields and anthropomorphic phantoms: namely the mathematical traceability, the realism and the morphological variability compared to clinical breast images.

Slide 8

To develop such a model, we use a data base of clinical breast CT reconstructed images as input. As a first analysis, we extracted cubes inside the breast CT volumes and then segmented the fibroglandular breast tissue from them. This resulted in binary volumes, shown here, depicting the fibroglandular and adipose breast tissue.

Next, we reprojected the binary volumes using an x-ray simulator. This allowed us to obtain simulated DBT projection images, as shown here. Finally, a 3D reconstruction algorithm was applied on the DBT projection images. This allowed us to obtain DBT reconstructed slices. One of them is illustrated here.

For a visual comparison, below we have images extracted from clinical DBT projection and reconstructed images. We observe that the segmented breast CT data sets allows to simulate realistic DBT projections & reconstructed slices.

This analysis gives us the idea of modeling the tissue morphology & distribution as seen in breast CT data. We focus on the medium and the small-scale breast tissues. For the medium-scale intra-gland adipose compartments, we observed that they can be approximated by system of random ellipsoids. We also observed small-scale adipose compartment boundary irregularities. To model this small-scale aspect, we will use Voronoi diagrams.

Slide 9

Now let me dive into the algorithm to sample from our new texture model.

We start with the small scale, First, we sample points from a uniform Poisson point process, from which we create a Voronoi diagram with very small cells. Then we deal with medium scale. Independently from the first two steps, we sample points from a seed point process with distribution Ps. At each seed point, we place a random ellipsoid oriented towards the nipple, with distribution P_theta. The final step is the voxelization at resolution nu. This is done by checking each small Voronoi cell in step 2. If its center falls inside of one of the ellipsoids, we assign 0 to the entire cell; otherwise, we assign 1 to the entire cell.

Slide 10

Now let me show you some simulation results using the proposed 3D breast texture model.

In a first instance, the model parameters were empirically determined by visually comparing with the segmented breast CT data. All the images were simulated using a calibrated virtual X-ray imaging simulator, without the modelling of the mechanical breast compression.

Slide 11

These are slices from an adipose breast CT volume, and from a volume of the same density category simulated using our model. These are the mammography projections simulated using the breast CT data, and using our model. And these are DBT reconstructed slices simulated using the breast CT data, and using our model.

These are images from volumes representing two different breasts with different density categories. If we compared the images from our model to images from breast CT data, we can see that simulated images visually appear fairly realistic, in terms of medium and small scale breast textures.

Slide 12

To demonstrate the effect of small Voronoi cells, we performed a simulation using smooth ellipsoid boundaries. From left to right, these are a volume slice, a mammogram and a DBT reconstructed slice from a volume simulated with smooth ellipsoid boundaries. We can see that there are geometric artifacts in simulated images. Especially in the mammogram in the middle.

Now let’s compare the micro textures in these images simulated with images simulated with irregular boundaries.

Slide 13

Here is a volume slice, a mammogram and a DBT reconstructed slice from a volume simulated with irregular ellipsoid boundaries. The medium scale of the simulated volume is the same as the volume from the previous slide. The only difference is that, in this simulated volume, small Voronoi cells were introduced around the surface of the ellipsoids to make the boundaries irregular.

If we compare the micro-textures in these images with previous ones, we can see that the geometric artifacts disappear in these images. Small Voronoi cells indeed improves the micro-texture realism.

Slide 14

To formally assess the model realism, we performed a psycho-visual experiment referred to as the two-alternative forced choice experiment.

The experiment was carried out in a darkened room, where the observers maintained a 40cm approximate distance to the image display. All images were displayed with 100%-pixel resolution.

Here is what the observers saw on the screen for each trail. On the left one we have one DBT slice created using breast CT data. On the right, we have one DBT slice in the same density category, created using either the breast CT data, or our model, with 50% chance of each.

The task for the observers was to answer whether the image on the right was created using breast CT data or not.

Slide 15

The observer's percentage of correct answer was measured. Here is a graph showing the percentage scores for all four observers, and for all three types of simulated DBT reconstructed slices. Notice that, a 50% score means that the observer could not tell the difference btw images from our model and from breast CT data.

We can see that for the most adipose category, the simulated images were fairly realistic, since all observers had the lowest score for this category.

However, the observers reported block artifacts in the images. As shown in the images on the right. These artifacts actually helped them to identify the ground truth. It turned out that, these artifacts were due to error in the small-scale model parameters. We fixed this error right after the experiment.

Slide 16

Our next study consists of formally obtain the parameters of the texture model using a set of segmented clinical breast CT data. The idea is that, from a set of segmented clinical breast CT volumes of interest, we perform statistical inference to estimate the model parameters.

Particularly, we focus on the medium scale model parameters. That is, the distribution for the seed point process, and the distribution for the ellipsoid parameters.

Slide 17

Let's start by formulating the inference problem.

We formulate the medium scale model as a marked point process, where the ellipsoid center point process has distribution P_s, and the ellipsoid parameters have distribution P_theta. The input breast

CT volume is formulated as a binary volume, where a voxel is assign 1 if its fibroglandular tissue; and is assigned 0 otherwise. Noted that here we use medium scale volume size in order to focus on the medium scale model.

The particularity of the inference problem is that it is ill-posed, due to the fact that the ellipsoid centers are not observable. For example, here we have two different system of ellipses that both approximate well the binary image. Finding a criterion to decide which one is better can be very challenging.

Slide 18

There are classical inference methods that may not necessitate the knowledge of the ellipsoid centers.

One of the method is referred to as the minimum contrast estimator. Where the parameters are estimated by minimizing a contrast function. Often, the contrast function is defined as the integrated sum of two parts. S, which is usually some analytical 2^{nd} order statistics that does not necessitate the ellipsoid center. S-hat, which is the empirically measured counterpart of S.

The minimum contrast estimator however has challenges. First, in order to define S, we need to assume the model type, which can be non-trivial. Secondly, the analytical derivation of S can be challenging when the underlying model is complex.

Slide 19

To alleviate these challenges, we look at another approach referred as the inference from reconstruction.

It is a two-step approach. The 1^{st} step is the reconstruction step. Where we sample a good ellipsoid approximation of the ground truth using stochastic sampling technique. The second step is the inference step where we estimate the parameters from ellipsoids reconstructed in the first step.

The first advantage of this approach is that, the reconstruction step reveals the position of the ellipsoid centers, this can make the inference much easier. Also, from reconstructed ellipsoids, we can perform analysis to gain intuitions on the model type.

Slide 20

Let's now at the method of inference from reconstruction in detail.

The goal of the first reconstruction step is to find one optimal ellipsoid approximation of the input data. We made the hypothesis that the underlying marked point process has a Gibbs density, with an energy term U. The energy is decomposed into two terms. The first terms L, measures how a configuration of ellipsoids deviates from the input data. The second term P is the prior term. Where we put a constraint on the overlap ratio between ellipsoids.

With this formulation, the optimal ellipsoid configuration is the configuration that minimizes the energy term. Notice that this optimization is difficult to solve analytically. To tackle this problem, one popular method is to use Monte Carlo dynamic algorithms.

Slide 21

To develop such an algorithm, we adopted an algorithm referred to as the multiple births, deaths and shifts.

We start with an initial configuration. At a given iteration, we first perform multiple births, where random ellipsoids were generated. Then we examine all the ellipsoids to determine either to remove it or to shift it. We do it by computing a probability term r, that depends on the energy contribution of the examined ellipsoid. With probability r, we decide to remove the ellipsoid. Otherwise, we will shift the ellipsoid. The shift is done by first computing a new ellipsoid, referred to as the Legendre ellipsoid, and then replacing the current ellipsoid by the new Legendre ellipsoid. The Legendre ellipsoid is defined as the optimal ellipsoid of the black region covered by the current ellipsoid.

In the end of each iteration we update the current configuration and parameters. The final convergence was determined by looking at 10 consecutive iterations. If the energy variation during the 10 runs is smaller than a threshold, we stop the algorithm.

Slide 22

Let's visually analyze the results from the reconstruction step. Since in this study we focus on the medium scale model, for the following visualization, we will also only focus on mediums scale breast tissue.

The projection images in the following were created by computing the sum of the volume in one direction, instead of using a virtual x-ray simulator. This means that the x-ray noise is absent in the images.

Slide 23

Here is a slice from an adipose breast CT volume. And here is a slice from the volume representing the result of its reconstruction step. The reconstruction volume was created by voxelizing the ellipsoids reconstructed from the original breast CT volume. Since we focus on the medium scale, the ellipsoid boundaries were kept smooth. That is, the small-scale irregularities were not introduced.

Here are the projection images created respectively from the original breast CT volume and the reconstruction volume. And here, we also have three other sets of results from three breast CT volumes with different glandular densities.

From the visual comparison, we can see that the reconstruction step is not perfect; However, we can observe that the reconstruction preserves the medium scale characteristics of the original breast CT volumes, which is the purpose of the reconstruction step.

Slide 24

Next, the inference step. The goal of this step is to infer the center point process and ellipsoid parameters from reconstructed ellipsoids.

For the ellipsoid centers, we performed an analysis of the pair correlation function, or the PCF. The PCF is a second order statistics of point processes. It measures the probability density of having a pair of points located at two locations. For a stationary and isotropic point process, the PCF depends on the distance r between pair of points.

The PCF can reveal important distributional information of point processes. For the Poisson point process, the points are independently distributed, the PCF is constant equal to 1 in average. For a point process with cluster effects, the PCF is greater than one for some range. Here is an example of a Matérn cluster process. It is constructed by first sampling randomly positioned spheres with radius R, then generating Poisson point process inside each sphere.

By analyzing the empirical PCF of the reconstructed ellipsoid centers, we found that almost all the reconstructed ellipsoid configurations exhibit a clustering effect. This can be seen if we look at the PCF graph on the left, for a range distance values, the PCF is greater than 1.

Therefore, performed a fit of the reconstructed ellipsoid centers to the Matérn cluster process using the minimum contrast estimator. The graph on the right shows the fitted Matérn process PCF in comparison with the empirical PCF. We can see that the two PCFs matches well, meaning that the fit was good.

Slide 25

Regarding the ellipsoid axis lengths and orientation angles, we performed analyses of their empirical histograms.

Statistical tests were performed, showing that the distribution of these parameters can be modeled using either Gaussian or uniform distributions. This allowed us to fit these model parameters to either Gaussian or uniform distributions using the maximum likelihood method.

Slide 26

To visually validate the inferred medium scale model parameters, we performed new simulations using the newly inferred mediums scale model parameters.

Now let me show you some simulated images using the texture model with inferred medium scale parameters. The small-scale model parameters were the same as with the empirical model. All the image acquisitions were simulated using a calibrated virtual X-ray imaging simulator, without the modelling of breast compression.

Slide 27

Here is a slice from a volume simulated from the model representing an adipose breast. Here is the mammogram created from the same volume. And here is the DBT reconstructed slice created from the same volume. Here we also have other three sets of simulations representing three breasts with different densities.

By observing these simulated images, we can see that the simulated images exhibit high visual realism. Also, we notice the improved morphological variation in simulated images. This is can be mainly attributed to the newly inferred medium scale parameters.

Slide 28

The second part of this thesis is to develop a new model observer for 2D & 3D breast imaging based on the *a contrario* theory.

For a lesion detection task, a model observer is defined by a decision function f, which returns a decision variable lambda, indicating the observer's confidence level of the lesion presence in image g. For example, with linear observers, lambda is computed as the inner product of a template w and the input mage g.

Then, different thresholding techniques can be applied to observer's lambda values. This allows to further deduce a metric as the figure-of-merit for the detection task. For example, one commonly used figure-of-merit in binary classification tasks is the area under the receiver operator characteristics curve.

Slide 29

The channelized Hotelling observer, or CHO, is the state of the art for linear observers. It is based on linear discriminant analysis and it can be extended to 3D cases.

The CHO consists in two phases.

The first training phase aims to compute the template w_cho. We start by computing the inner product of an input image with a set of channels. This is referred to as the channelization process. It transforms the input image into a feature vector, and can reduce the dimension of the input image. The channelization process is performed for both lesion present and absent images, referred to as the training sets.

After the channelization, the lesion is estimated as the difference between the lesion-present and absent vectors. The covariance is estimated as the average of the covariance of lesion-present and absent vectors. Finally, the template w_cho is expressed as the product of the channels, the estimated lesion and the inverse of the covariance matrix. Notice that the division by the covariance matrix acts as a decorrelation or whitening step of the input images.

The second phase is the test phase, it has the objective to compute the final CHO decision variable for a given image. Given an input image g. We compute the inner product of the image with the template obtained from the training step. This yields the final decision variable for that image.

Slide 30

The *a contrario* observer is a non-linear observer

Derived from the a contrario detection principles, it is based on statistical tests and it works by quantifying the significance of a measured event in a statistically random image.

Here is an example where we want to detect a spot in a white noise image. We start by defineing a naïve model, which is a null-hypothesis H0, often referred to as the noise or background model. Then we perform a set of measurements. These measurements can be local image characteristics, such as the local contrast. Next, we compute a quantity named the number of false alarms, or the NFA. The NFA quantifies how a measurement deviates from the naïve model. Finally, we threshold on the NFA values for the detection.

The appealing property of thresholding on NFA is that we can globally control the false positive rate. That is, when thresholding the NFA at level epsilon, the expected number of false detections is bounded by epsilon.

Slide 31

The goal of our study is to develop a new *a contrario* observer. For this study, we aim to focus on microcalcification detection in 3D breast imaging. As stated earlier, this goal is driven by two motivations.

First, variable results from existing clinical studies have been obtained regarding the comparison of 2D FFDM and 3D DBT in ucal detection performance. Secondly, limited research has been done regarding the ucal detection performance in S2D images versus FFDM.

Slide 32

Let’s look at the detail of the proposed *a contrario* observer.

For a pixel x, we first define a local naïve model. This is done by restricting our analysis to a neighborhood centered at x. The neighborhood is chosen to be less than 5 by 5 centimeter squared. This is to encounter human observer's local detection characteristic, as reported in literature.

The neighborhood is defined by an inner rectangle and an outer rectangle, both centered at x. The pixels inside the inner rectangle were removed to exclude possible presence of ucal at x. Therefore, the neighborhood is finally defined as the contour area.

From a neighborhood, we estimate the local naïve model, assumed to be a white Gaussian noise. The mean and standard deviations of the local naïve model were empirically computed from the neighborhood.

Slide 33

Next, to encounter different ucal sizes, we extend the neighborhood to a multi-scale setting. This is done by simultaneously varying the size of the inner and outer rectangles.

At each scale, the measurement is chosen to be the pixel value at x. And we compute the NFA at this scale as the product of the number of total tests N_T and the Gaussian tail probability, defined by the local naïve model and the pixel value at x.

This step allows us to obtain multi-scale NFA values.

Slide 34

The next step is to repeat the multi-scale NFA computation for all pixels in the image. At each pixel, the decision variable is defined by retaining only the minimum of all NFA values at all scales. This yields an image lambda_x of decision variables for all pixels.

The proposed *a contrario* observer can be extended to location known exactly tasks in 3D DBT slices. For the extension, we consider the fact that the ucal is typically smaller than the distance between two DBT slices. Therefore, we proposed a 3D extension by separately processing the slices.

We start with a subset of slices where the microcalcification is contained inside. Then, for each slice, we restrict our computation to a small neighborhood around the known ucal location. Finally, a scalar decision variable lambda is obtained by taking the minimum of the computations in all selected slices.

Slide 35

To validate the proposed a contrario observer, we first performed a theoretical proof.

We demonstrate that, when thresholding the decision variable image λ_x at threshold epsilon, the expected number of false positives according to the naïve model, is bounded by a function C. Further, we proved that, if all neighborhoods are large enough, the function C is actually of the same order of magnitude as epsilon.

To verify whether the false positive control is also true for breast-like images, we performed a simulation experiment. 200 pairs of FFDM like images were simulated. Half of them were randomly inserted with 10 ucalc-like inserts. Then we run the a contrario observer and we estimated the empirical average false positive rate.

Slide 36

Here is the result of the simulation experiment. The red curve is the analytical upper bound C for the false positive rate. The black curves are empirically measured false positive rates.

We observe that the black curve is always bounded by the red curve. This means that the proposed *a contrario* observer indeed allows for a global control of false positive detections.

Slide 37

The final part of the thesis is to apply the developed texture model and *a contrario* observer in a complete VCT experiment to assess the ucalc detection performance in FFDM, DBT and Synthetic 2D images.

We first aim to compare the performance of the proposed a contrario model observer with the channelized Hotelling observer. The CHO is the most commonly used model observer in various VCT studies today. We also aim to compare the performance of the model observers with the performance of a human observer.

Slide 38

Here are examples of simulated FFDM, DBT reconstructed slices and S2D images for three types of backgrounds use in our experiment.

The scattered fibroglandular dense and heterogeneously dense breast images were simulated using our 3D breast texture model with empirical parameters. The uniform background images were simulated using a uniform digital phantom with glandular density equal to the average of the breast texture volumes. All the image acquisitions were simulated using a virtual x-ray projector calibrated according to a GE commercial imaging system.

Microcalcifications with different sizes and attenuation characteristics were simulated. Here we show you images where the ucal can be clearly observed at the center of the images.

Slide 39

Here is the set-up for the VCT experiment.

The task was designed to be a rating scale task, where the observer knew exactly the ucalc location and knew statistically the microcalcification appearance. 1 human observer, the a contrario observer and the 2D & 3D CHO performed the tasks.

For each image review, the observers were asked to rate his confidence level for ucalc presence in the center of the image. The observer's responses were analyzed using the receiver operator characteristics or ROC methodology. Finally, the area under ROC curve, or the AUC was used as the figure of merit.

Slide 40

Let's first compare the AUC performance of the *a contrario* observer and the CHO for different ucalc sizes and attenuation characteristics.

The results are shown for FFDM images with heterogeneously dense breast background.

The left graph shows the AUC of the two observers for different ucalcs sizes and a given attenuation. The error bars indicate twice the standard deviation. The AUC of both model observers increase with ucalc size.

The right graph shows the AUC of the two model observers for different ucalc attenuation characteristics. The size of the ucalc is 200 um. The AUC of both model observers increase with the ucalc attenuation.

These observations intuitively make sense. We found that these observations are consistent for all experimental conditions

Slide 40

Let's first compare the performance of the *a contrario* observer with the performance of the CHO.

Here the graph shows the comparison for all experimental conditions. The abscissa shows the performance of the *a contrario* observer in terms of the AUC score. The ordinate shows the performance of the CHO in terms of the AUC score. Here the error bars represent twice the standard deviation of the AUC.

We can see that the CHO always outperforms the a contrario observer. Also, there seems to be no clear linear correlation between the two observers.

Slide 41

Next, we compare the performances of the 2 model observers with the human observer, for the three types of background types simulated in our experiment

Here is the result for the *a contrario* observer. The abscissa shows the AUC for the a contrario observer and the ordinate shows the AUC for the human observer.

We can see that the *a contrario* observer under performs the human observer in all cases shown. Also, there seems to be a positive correlation btw the *a contrario* observer and the human observer performance in all background types.

Here is the result for the CHO. Similarly, the abscissa shows the AUC for the CHO and the ordinate shows the AUC for the human observer.

The CHO outperforms the human observer almost in all cases. For the uniform and the scattered fibroglandular dense backgrounds, the CHO is positively correlated with the human observer. However, it is hard to make solid conclusions on the correlation since some AUC values for CHO is close to 1. As for the heterogeneously dense background, a positive correlation between the CHO the human observer cannot be confirmed.

Slide 42

Finally, we compare the ucalc detection performance in FFDM, DBT and S2D images.

Here we illustrate the case for 200-micron meter ucalcs, with the highest attenuation. The three graphs on the top show the result for the *a contrario* observer in all three background types. The three graphs on the bottom show the result for the CHO in all three background types.

From the graphs we can see that, for both the CHO and the *a contrario* observer, there is a trend that the ucalc detection performance in S2D images is lower than in FFDM and DBT images.

However, regarding the ranking between FFDM and DBT, there seems to be no clear conclusion for neither *a contrario* observer nor the CHO.

Slide 43

After the analysis of all experimental conditions, we found that it was hard to make clear conclusion regarding the ucalc detection performance ranking between FFDM, DBT and S2D. The result varies according to different experimental conditions. This indicates that further fine tuning of the simulation set-up might be needed.

Still, we have observed that the CHO outperforms the ACO in all conditions; and there is a trend that both the performances of the CHO and the ACO are positively correlated to the performance of the human observer.

This VCT experiment demonstrate the potential of using our developed 3D breast texture model and a contrario observer for future design of more elaborated VCT studies.

Slide 44

The main results and contributions of this thesis are as follows.

In terms of tools, we first developed a mathematical 3D breast texture model that shares the advantages of the 3D random field texture models and anthropomorphic breast phantoms. It is mathematically traceable, in that it can be characterized using stochastic geometric theory. The breast images simulated using this model present a high visual realism compared with clinical breast images.

We also developed a new a contrario observer for ucalc detection in 2D breast images and in 3D DBT slices. The *a contrario* model observer provides an alternative model to the state of the art model observers such as the CHO.

In terms of methodology. We explored the use of the stochastic geometry theory in x-ray breast imaging. With our 3D texture model, we successfully demonstrated the potential of using this branch of mathematics for the modeling of different anatomical breast structures.

We also investigated a new statistical inference approach for marked point processes, referred to as the inference from reconstruction. We successfully inferred the medium scale texture model

parameters from clinical breast CT images using this approach. This has improved the morphological variability of our texture model.

Finally, in terms of application, we applied the developed texture model and *a contrario* observer to a complete VCT experiment to assess the ucalc detection performance in FFDM, DBT and S2D images. This is a clinically relevant application and it demonstrates the potential and utility of our developed VCT tools.

Slide 45

There are several directions for future research to optimize and extend the study in this thesis.

First, regarding the proposed 3D breast texture model, a statistical validation by comparing the statistical characteristics of model simulated images with clinical images should be performed to further validate our model. Also, the current texture model could be extended to simulate other breast anatomical structures such as the ductal trees.

Regarding the inference from reconstruction method, some further theoretical study and analysis of the inference accuracy should be performed. Also, the effect of the algorithm initialization should be studied. In our study, we assumed that some of the ellipsoid parameters were independent of their spatial positions. This might be unrealistic if we observe real breast anatomy. This aspect can be further addressed by studying the mark correlation function of the underlying marked point process model. In a first instance, we separately performed inference on different input breast CT volumes. This allowed us to obtain several sets of model parameters. In the future, it is also interesting to unify different sets of model parameters into a unified model. Finally, some statistical and psycho-visual validation of the inferred model parameters should also be performed.

Regarding the a contrario observer, the current 3D extension consists in processing independently each DBT slice. A future extension to model the correlation between DBT slices should be of interest. The current a contrario observer uses the white Gaussian noise as the naïve model. This naïve model does not take into account the image noise correlation. An extension to more complex naïve models, such as the power-law Gaussian random field might be worth investigating to incorporate the noise correlation. Finally, the current a contrario observer assumes that the ucalcs in images are isolated. However, ucalcs that appear in clusters are clinically important. A possible future step is to extend the current a contrario observer for the detection of ucalc clusters.

~~Finally, regarding our VCT application, some refinement of the experimental set-up should be performed, in order to better differentiate different modalities. In our study, we modeled a nominal system topology with several simplifications. As a future step, the imaging system simulation can be extended to approach real imaging systems. For example, by using non-point source and by adding electronic noise etc.~~