Medical Physics
The International Journal of Medical Physics Research and Practice

Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture

Predrag R. Bakic, Michael Albert, Dragana Brzakovic, and Andrew D. A. Maidment

Citation: Medical Physics **29**, 2140 (2002); doi: 10.1118/1.1501144
View online: http://dx.doi.org/10.1118/1.1501144
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/29/9?ver=pdfcov
Published by the American Association of Physicists in Medicine

Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture

Predrag R. Bakic and Michael Albert
Department of Radiology, Thomas Jefferson University, Suite 3390, Gibbon Building, 111 South 11th Street, Philadelphia, Pennsylvania 19107–5563

Dragana Brzakovic
Office of Integrative Activities, National Science Foundation, Arlington, Virginia 22230

Andrew D. A. Maidment[a)]
Department of Radiology, Thomas Jefferson University, Suite 3390, Gibbon Building, 111 South 11th Street, Philadelphia, Pennsylvania 19107–5563

(Received 25 February 2002; accepted for publication 14 June 2002; published 27 August 2002)

We have evaluated a method for synthesizing mammograms by comparing the texture of clinical and synthetic mammograms. The synthesis algorithm is based upon simulations of breast tissue and the mammographic imaging process. Mammogram texture was synthesized by projections of simulated adipose tissue compartments. It was hypothesized that the synthetic and clinical texture have similar properties, assuming that the mammogram texture reflects the 3D tissue distribution. The size of the projected compartments was computed by mathematical morphology. The texture energy and fractal dimension were also computed and analyzed in terms of the distribution of texture features within four different tissue regions in clinical and synthetic mammograms. Comparison of the cumulative distributions of the mean features computed from 95 mammograms showed that the synthetic images simulate the mean features of the texture of clinical mammograms. Correlation of clinical and synthetic texture feature histograms, averaged over all images, showed that the synthetic images can simulate the range of features seen over a large group of mammograms. The best agreement with clinical texture was achieved for simulated compartments with radii of 4–13.3 mm in predominantly adipose tissue regions, and radii of 2.7–5.33 and 1.3–2.7 mm in retroareolar and dense fibroglandular tissue regions, respectively. [DOI: 10.1118/1.1501144]

Key words: mammography simulation, 3D, synthetic mammograms, texture analysis

I. INTRODUCTION

We have proposed an approach to generate synthetic mammograms based upon a 3D simulation of mammography.[1] Synthetic mammographic texture is produced by projecting simulated 3D breast anatomic structures. In clinical images, the overlapped projections of normal anatomic tissue structures generate a background texture in mammograms which can mask the existing abnormalities or introduce false ones. The simulation can be used to optimize positioning, compression and acquisition in order to improve the visibility of the breast tissue, and to test new breast imaging modalities.

The proposed mammography simulation consists of three major components. First, a 3D software breast phantom contains two ellipsoidal regions of large scale tissue elements: predominantly adipose tissue (AT) and predominantly fibroglandular tissue (FGT) regions. Internal structures of these regions, namely the adipose compartments and breast ductal network, are approximated by realistically distributed medium scale phantom elements: shells filled with simulated adipose tissue and a synthetic ductal tree. Second, a compression model of the breast deformation occurring during a mammographic exam is based upon tissue elasticity properties. Deformation is simulated separately for layers of tissue positioned normal to the compression plates. Each slice is approximated by a rectangular beam composed of AT and FGT regions. The slices are computationally deformed, assuming clinical values of the compression force. Deformed slices are stacked together to produce a model of the compressed breast. Third, mammogram image acquisition is modeled assuming monoenergetic x rays and a parallel beam geometry without scatter. Details of the simulation are given in the accompanying paper.[1]

Ideally, each of the three components of the simulation should be evaluated separately by a 3D imaging technique. There is, however, a significant difference in tissue properties captured by the clinically available 3D breast imaging modalities (ultrasound and MRI) and mammography which is the focus of our simulation. Breast ultrasound and MRI also have different resolution and compression than mammography. With these issues in mind we have evaluated the tissue model indirectly, assuming that a relationship exists between the distribution of 3D breast tissue structures and the 2D parenchymal pattern. It is our hypothesis that the texture properties computed in synthetic and clinical images have similar distributions.

There are two approaches to mammogram synthesis found in the literature: (i) direct modeling of 2D distribution of pixels and (ii) simulation of 3D tissue distribution and the

mammographic imaging. Bochud *et al.*[2] modeled mammogram texture as a ''clustered lumpy background'' by random placement of ''blob'' clusters, visually resembling tissue appearance in mammograms. Synthetic images were evaluated by comparing their power spectra and statistical moments with the values from 32 clinical mammograms. Good agreement of the first and the second moments in clinical and synthetic images were observed, with similar statistical properties overall. Heine *et al.*[3] modeled a mammogram as evolving from a process of passing a random field (colored noise) through a linear filter with a self-similar characteristic, based upon the analysis of 60 clinical mammograms. Such an approach can match some of the statistical properties of clinical images but cannot relate the 3D tissue structures and their mammographic appearance. Both papers do not model breast ducts or the large scale tissue regions. Consequently, the images of the same simulated breast, with modified positioning, compression, or x-ray parameters cannot be consistently synthesized.

Taylor *et al.*[4] generated synthetic images by mammography simulation, in an approach similar to our work. The focus of their simulation is on modeling breast ducts based upon the fractal properties of the duct length and diameter. They have evaluated the synthetic images so obtained by comparing the Fourier spectrum with that computed in images of tissue slices with contrast enhanced ducts. Good agreement using a small number of samples was observed.

Separate evaluations were performed for the simulation of the ductal network, the compression model, and the synthetic parenchymal pattern. Initial feasibility tests of the ductal model and compression simulation are presented elsewhere.[5,6] This paper describes the analysis of the synthetic mammogram texture.

Synthetic images were generated by simulating the x-ray image acquisition on a computationally compressed phantom. Images of the phantoms were generated containing different sizes of simulated medium scale elements: spherical shells and blobs. The synthetic mammograms, so obtained, were evaluated by comparing them with clinical images taken from the MIAS database of digitized mammograms.[7] Subimages taken from regions corresponding to different tissues were compared separately, including the subcutaneous AT, retromammary AT, retroareolar FGT, and dense FGT regions. Three texture features were used for description of the parenchymal pattern: (i) the average size of image structures computed using mathematical morphology, (ii) the texture energy, and (iii) the fractal dimension. Feature values were computed over each clinical and synthetic subimage and statistically compared using the Kolmogorov–Smirnov test and histogram correlation. Details of the analysis of synthetic and clinical mammographic texture are given in Sec. II and the results of the comparison are discussed in Sec. III.

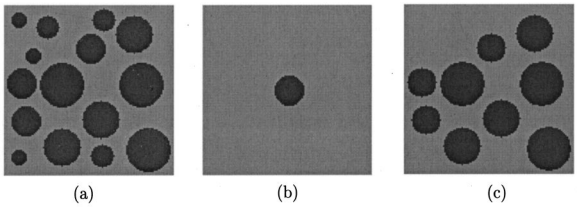

FIG. 1. Illustration of morphological closing. (a) The original image with objects of different size. (b) The image of the structuring element. (c) The resulting image obtained by the morphological closing with the structuring element from (b) applied to the image from (a).

II. TEXTURE ANALYSIS OF SYNTHETIC MAMMOGRAMS

A. Texture descriptors

The following texture descriptors were used for the evaluation of synthetic mammogram texture. First, size analysis was performed by a sequence of morphological closings with disks of increasing size as structuring elements.[8] Average image brightness increases after the closing operation. The change in brightness as a function of the disk radius is related to the size distribution of radiolucent (adipose) areas in the mammograms. Second, texture energy analysis was performed by convolving each image with a small mask.[9] Treating gray scale image intensity as the height of a 3D object, this mask is sensitive to local roughness of the image surface. Third, fractal dimension was computed by the blanket box counting method of self-similarity analysis.[10]

1. Morphological analysis of image structure size

Morphological image analysis is based upon the shape of image objects and is used to simplify image data while preserving shape characteristics. The theory of mathematical morphology is discussed in the books of Matheron[11] and Serra.[8] An application oriented tutorial of morphological image processing is given by Haralick.[12]

Morphological operations are performed on a set of image pixels using a second set of pixels called the structuring element. Definitions of the basic operations are given in the Appendix. The opening operation is used for size analysis of bright objects, and closing for the analysis of dark objects. This analysis is sensitive to the radiolucent areas of the mammogram, corresponding to the adipose tissue which appears darker than the surrounding tissue. X rays are less attenuated by adipose tissue, producing greater film density than connective tissue.

The gray scale closing first replaces each pixel with the maximum from its neighborhood defined by the structuring element (a disk). The original values are then recovered for all of the pixels, except for those from regions which are both darker than their surroundings and smaller than the structuring element. As an illustration, Fig. 1(a) shows an image with several objects of different size. After the closing operation with the structuring element from Fig. 1(b), the resulting image is given in Fig. 1(c). It can be seen that dark objects smaller than the structuring element have been eliminated; the resulting image is thus brighter than the original.

This is the basis for morphological size analysis, whereby the change in average image brightness (i.e., the total pixel sum after the closing) is used to describe the size distribution of the image objects. The derivative of the brightness as a function of size shows the contribution of the objects equal in size to the structuring element.

Morphological size analysis of mammograms has been reported previously in the literature.[13,14] Behrens and Dengler[13] reported examples of applying morphological size analysis at global, regional, and local image levels; analysis of calcifications was presented as a local processing. Miller and Astley[14] used morphological size analysis to segment the FGT region from mammograms. They used the opening operation which is dual to closing; it replaces the bright regions smaller than the structuring element by their dark surrounding pixels. The overall image brightness is, thus, reduced. However, the authors did not analyze the relationship between morphological feature values and the physical properties of the anatomic structures. Our research represents a novel application of morphological image analysis as a result of using the simulated 3D tissue structures to synthesize parenchymal patterns.

2. Texture energy analysis

Texture energy features are the statistical estimates of the outputs from a filter bank implemented in the form of local linear transformations. They were introduced with the goal of achieving texture segmentation and description at each image pixel, corresponding to a hypothetical low level function of the human visual system.[9] The filter bank consists of small 2D convolution masks whose coefficients are computed as the product of 1D masks with different numbers of zero crossings. Contrast invariance of the filter outputs is achieved by the normalization with the output of the filter sensitive to the average local image intensity. The absolute values or variances of the convolved images are used for analysis. A generalization of this approach can include a larger set of local linear transformations, and the estimation of higher order moments of the output channel histograms.[15]

In mammogram processing applications[14,16] texture energy was usually computed using a single or a few convolution masks. The mask sensitive to image ''ripple'' was found to be the most efficient in segmenting potentially abnormal regions in mammograms, a task which is related to the local roughness of the image surface. Texture energy features have also been used in mammogram registration.[17,18]

The mask coefficients are given in the Appendix. A 5×5 ''ripple'' convolution mask, R5R5 [Eq. (A2)], was used, capturing local roughness of the image surface. The absolute values of the convolved data were averaged on a 15×15 window and normalized by the ''level'' mask, L5L5 [Eq. (A2)], providing contrast invariance.

3. Fractal analysis

Fractal dimension describes self-similarity of image properties at different spatial scales. It is common to perform fractal analysis on the area of the image surface, obtained by considering the pixel values as local surface heights. This area is related to the roughness of the image texture. A complete definition of the fractal dimension of image surface area is given in the Appendix.

There are numerous reports in the literature on fractal analysis of mammograms.[19–22] Caldwell *et al.*[19] analyzed the fractal dimension of various parenchymal patterns and the difference between the fractal dimensions computed over the whole image and within a region near the nipple. A feature space defined by these two fractal features was segmented and a relatively good agreement with the original Wolfe classification[23] was observed. Mammographic calcifications have been segmented using a variety of methods for computing fractal dimension, including box counting,[20] iterated function systems,[21] and fractal Brownian motion.[22]

We computed fractal dimension by the blanket algorithm.[10] This method has been used previously in the detection of calcifications in mammograms.[20] The fractal dimension is computed for each pixel by analyzing the local image surface around the pixel. A 15×15 window was selected, centered on each pixel. This corresponds to the nonlinear averaging window size used in the texture energy method. A log–log plot of $A_{\text{local}}(\epsilon)$ is generated for the local surface around each pixel. The local fractal dimension value D_{local} is computed as the slope through three points on the log–log plot, corresponding to the scale parameter values of $\epsilon=2$, 3, and 4 pixels.

B. Image selection

The following criteria were used for selection of the clinical and synthetic mammograms to be used for comparison. First, the clinical images had to represent normal breast tissue. Second, the glandularity seen in the mammograms should approximately represent the average breast glandularity (not too dense and not predominantly adipose). Third, spatial resolution of the clinical and synthetic mammograms should be matched. The clinical images were selected from the MIAS database[7] of digitized mammograms and the synthetic mammograms were generated for varying properties of the medium scale elements, i.e., different sizes of simulated adipose compartments in the AT and FGT regions. In addition, the comparison was repeated for the same set of clinical and synthetic mammograms at a reduced resolution. The images with reduced resolution were generated by averaging 2×2 blocks of pixels from the original mammograms.

1. Clinical mammograms

Sixty-five mammograms from the Mini-MIAS database of clinical mammograms were used, having a spatial resolution of 200 μm/pixel. The Mini-MIAS database was obtained by averaging 4×4 pixel blocks in the original MIAS mammogram database.[7] This resolution is sufficient for the evaluation of our synthetic mammograms since presently they do not include fine, small scale tissue detail. The selected images represent normal cases in the MIAS database with the background tissue classified as ''fatty-glandular.'' As the sizes of adipose compartments differ for various tissue regions, up to four 25 mm$\times$25 mm subimages per mammo-

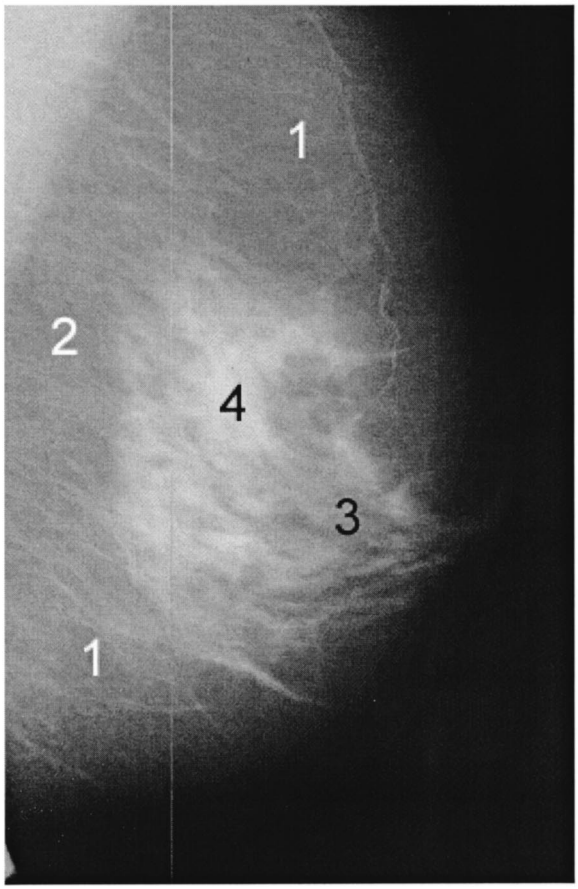

FIG. 2. Tissue regions used in texture analysis, illustrated on a clinical mammogram from the MIAS database: (1) subcutaneous adipose tissue, (2) retromammary adipose tissue, (3) retroareolar fibroglandular tissue (immediately posterior to the nipple), and (4) dense fibroglandular tissue.

gram were selected manually, giving a total of 219 subimages, from the following regions (see Fig. 2): (1) subcutaneous fat; (2) retromammary fat; (3) retroareolar glandular tissue, immediately posterior to the nipple; and (4) dense glandular tissue. If the extent of a tissue region could not be unambiguously determined, or if it was too small for a subimage window, the corresponding tissue sample was excluded from analysis.

2. Synthetic mammograms

Synthetic images were generated at a spatial resolution of 200 μm/pixel, matching that of the database. Four subimages per synthetic mammogram were selected from different regions in the same manner as for the clinical images. The positions of the subimages were determined from the known extent of the large scale model elements, the AT and FGT regions. Model parameters controlling the distribution of medium scale tissue structures, modeled by shells in the AT and spheres in the FGT regions, were varied to match the statistical properties of real images. Three groups of synthetic mammograms were tested. The groups consisted of ten synthetic mammograms each, generated randomly using the same range of size of simulated adipose tissue compartments. The ranges of compartment sizes differed between the groups by 30% (see Table I and Fig. 8 in the accompanying paper[1]).

C. Statistical comparison

Two methods were used for statistical comparison of the texture features. First, feature histograms were computed for each subimage. Synthetic histograms were then averaged over all subimages of the same tissue type and were compared with similarly computed clinical histograms. The correlation between the corresponding clinical and synthetic averaged histograms was used to measure how well the synthetic images approximated the clinical images. Next, mean feature values (i.e., the histogram first moments) were computed for each subimage. Distributions of these means for all subimages of the same tissue type were then analyzed and compared with the distributions of means of the clinical images, using the Kolmogorov–Smirnov (KS) test.[24] The maximum difference between the cumulative distribution functions (CDFs) of the clinical and synthetic mean feature values was used as another measure of quality of mammogram synthesis. In both methods the average texture features were compared, thereby testing the ability of the simulation to match the average properties of a large set of clinical mammograms, rather than simulating an image of a particular breast.

1. Analysis of feature histograms

As a measure of similarity between the real and synthetic feature distributions, the correlation between the feature histograms was calculated for each of the clinical and synthetic subimages, and averaged over all subimages of the same tissue type. In the case of size analysis, the correlation was computed between the brightness gradient (as a function of the structuring element radius) of clinical and synthetic images. In the following text, these derivative values are referred to as the ''average histogram of the size analysis feature.''

The coefficient of correlation, R, between the real, h_R, and synthetic, h_S, histograms averaged over all subimages (in a given category) is computed as:

$$R(h_R,h_S)=\frac{\Sigma_i h_R(i)h_S(i)}{\sqrt{\Sigma_i[h_R(i)]^2\Sigma_i[h_S(i)]^2}}, \quad (1)$$

where the summation runs over histogram bins i.

2. Kolmogorov–Smirnov (KS) test

The KS test compares two random distributions based upon the maximum difference between their CDFs.[24] It belongs to a group of nonparametric methods which make no assumptions about the types of distributions used. The maximum difference between two CDFs, D, is a measure of the discrepancy between the two sets of samples. Kolmogorov showed that for two sets of samples with the same parent distribution, the CDF of D is given asymptotically by:[24]

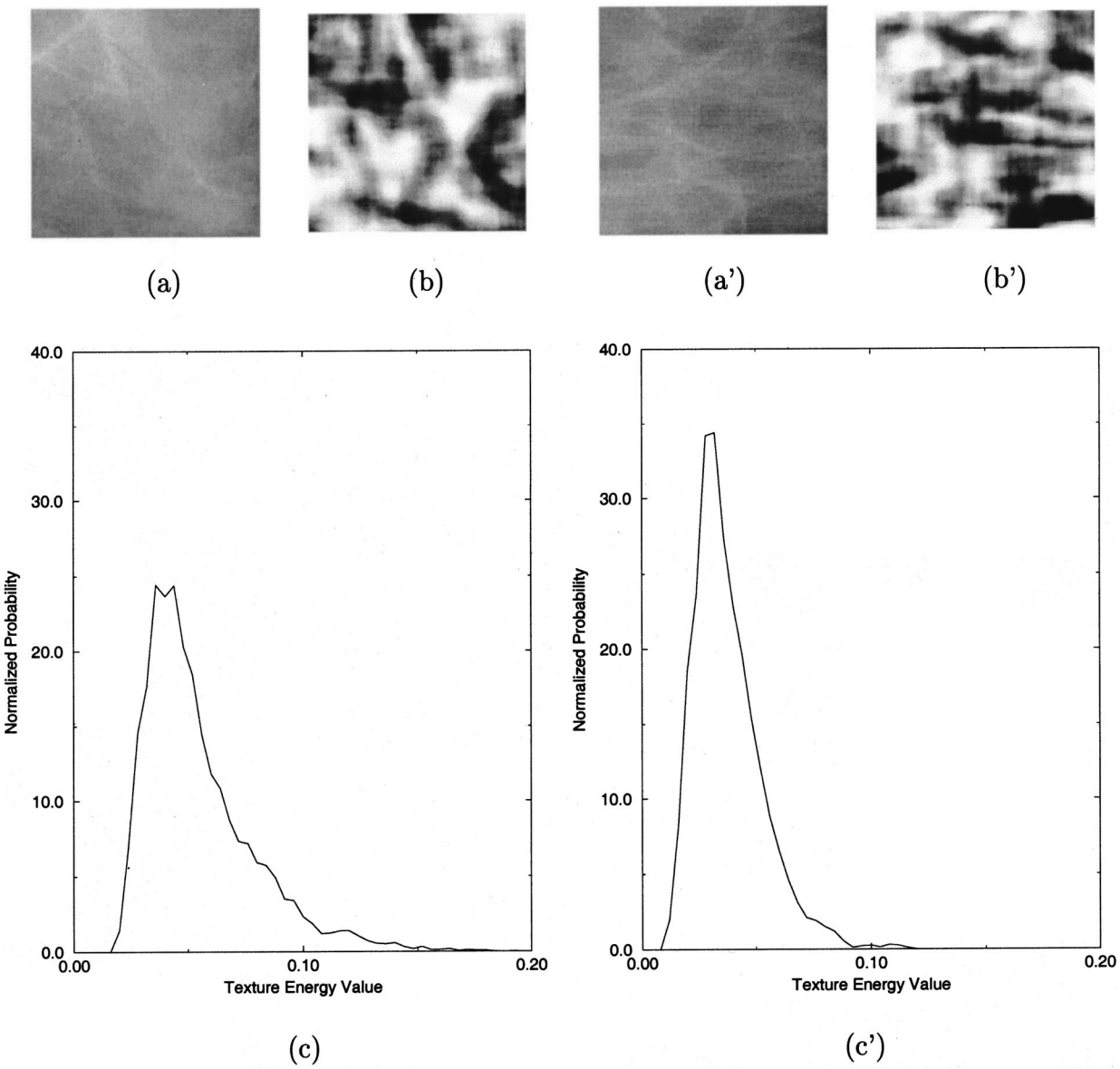

FIG. 3. Texture energy histograms of the FGT from clinical (left) and synthetic (right, primed) mammograms. (a) Sample subimage. (b) Image of texture energy values. (c) Texture energy histogram (normalized for the range of feature values).

$$\lim_{m,n\to\infty} P(D_{m,n}\leqslant z) = 1 - 2\sum_{r=1}^{\infty} (-1)^{r-1} \times \exp\left[-2r^2z^2\left(\frac{1}{m}+\frac{1}{n}\right)\right], \quad (2)$$

where m and n are the numbers of samples in the two sets, $D_{m,n}$ is the maximum CDF difference for the given number of samples, and P is the probability that $D_{m,n}$ is less than a given value z. The level of significance, α, is defined as

$$P(D_{m,n} > d_\alpha) = \alpha, \quad (3)$$

where d_α is the critical value of $D_{m,n}$ corresponding to the significance α. Thus, the observed discrepancy between a CDF drawn from clinical mammograms and a CDF drawn from simulated mammograms can be quantified in terms of the significance, α, which is the probability that a greater discrepancy than observed would occur due to chance alone. The relationship between α and d_α for various sample sizes is tabulated in several textbooks.[24,25]

The CDFs of statistics for each subimage from the clinical and synthetic mammograms were compared. In the texture energy analysis and the fractal analysis, for each subimage the appropriate feature value was averaged over all of the pixels in the subimage, and this average was used as a sample for the KS test. In the morphological analysis, for each subimage the first moment of the brightness gradient was used as a sample value.

Both the KS test and the histogram correlation show how well, *on average*, the synthetic images can approximate the properties of the clinical mammographic texture. The differ-

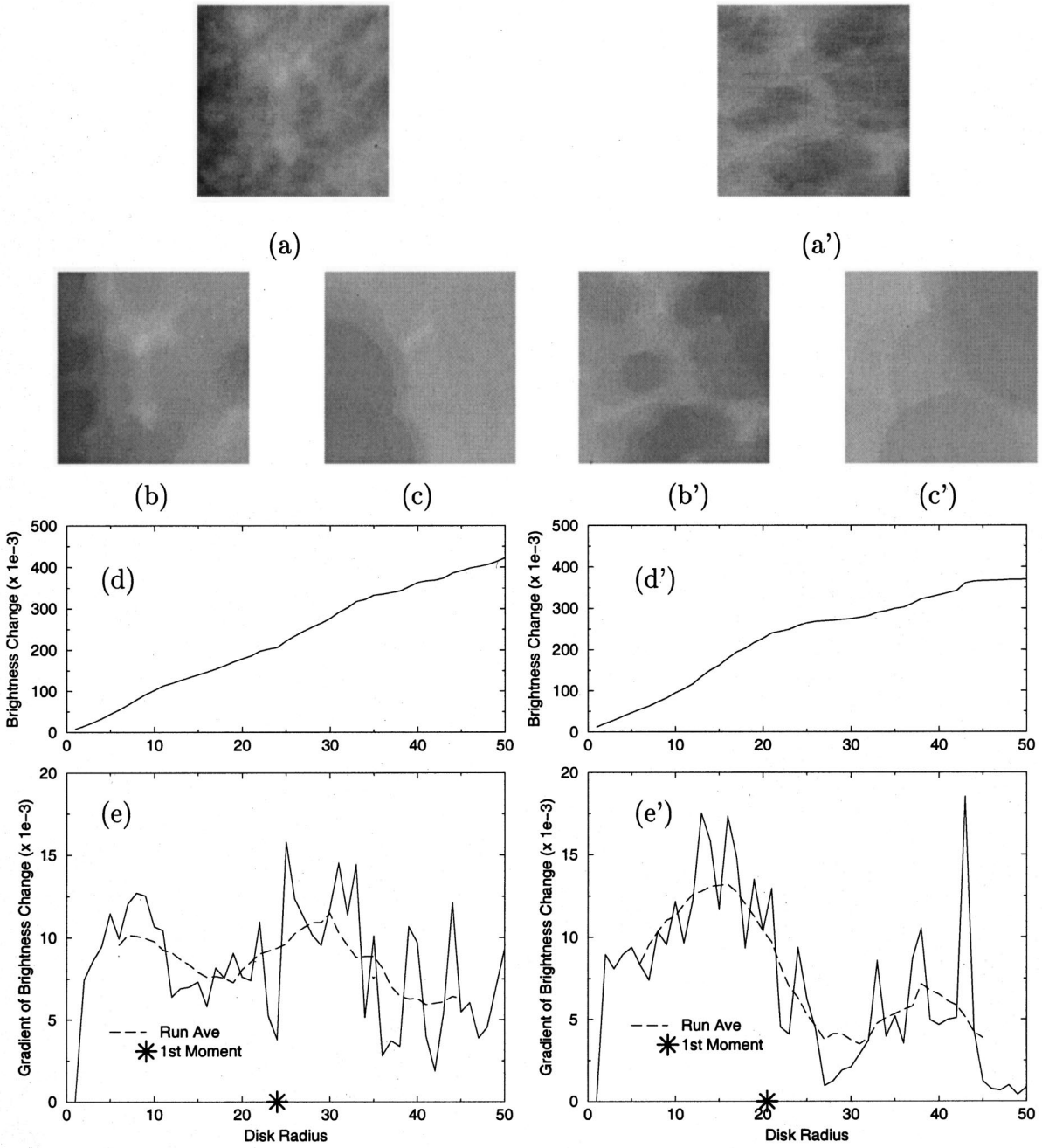

FIG. 4. Size analysis of the FGT from clinical (left) and synthetic (right, primed) mammograms. (a) Sample subimage. (b) Result of closing with a 10 pixel disk structuring element. (c) Result of closing with a 40 pixel disk. (d) Change in brightness (sum of all pixels) before and after closing. (e) Gradient of the brightness.

ence between the two methods is that the KS test compares the mean feature values averaged over each subimage, while the histogram correlation takes into account the range of feature values computed locally at each pixel.

3. Illustration of the analysis

An illustration of the histogram analysis is given in Fig. 3 by the texture energy features computed on subimages of retroareolar glandular tissue. The histogram of a clinical FGT subimage is shown on the left and of a synthetic subimage on the right. Histograms averaged over all clinical and over all synthetic subimages are shown in Fig. 5(c). Figure 4 illustrates the analysis of object size distribution for the retroareolar glandular tissue. The left-hand side shows the results for the clinical FGT and the right-hand side for the synthetic FGT. The upper graphs show the average brightness (offset for the brightness of the original image) after each morphological closing as a function of structuring element (disk) size in pixels. Note that the output images get brighter with increasing disk size, as seen in the examples of the images obtained for the disk radii of 10 and 40 pixels. The graph in Fig. 4(e) shows the gradient of the features graphed in Fig.

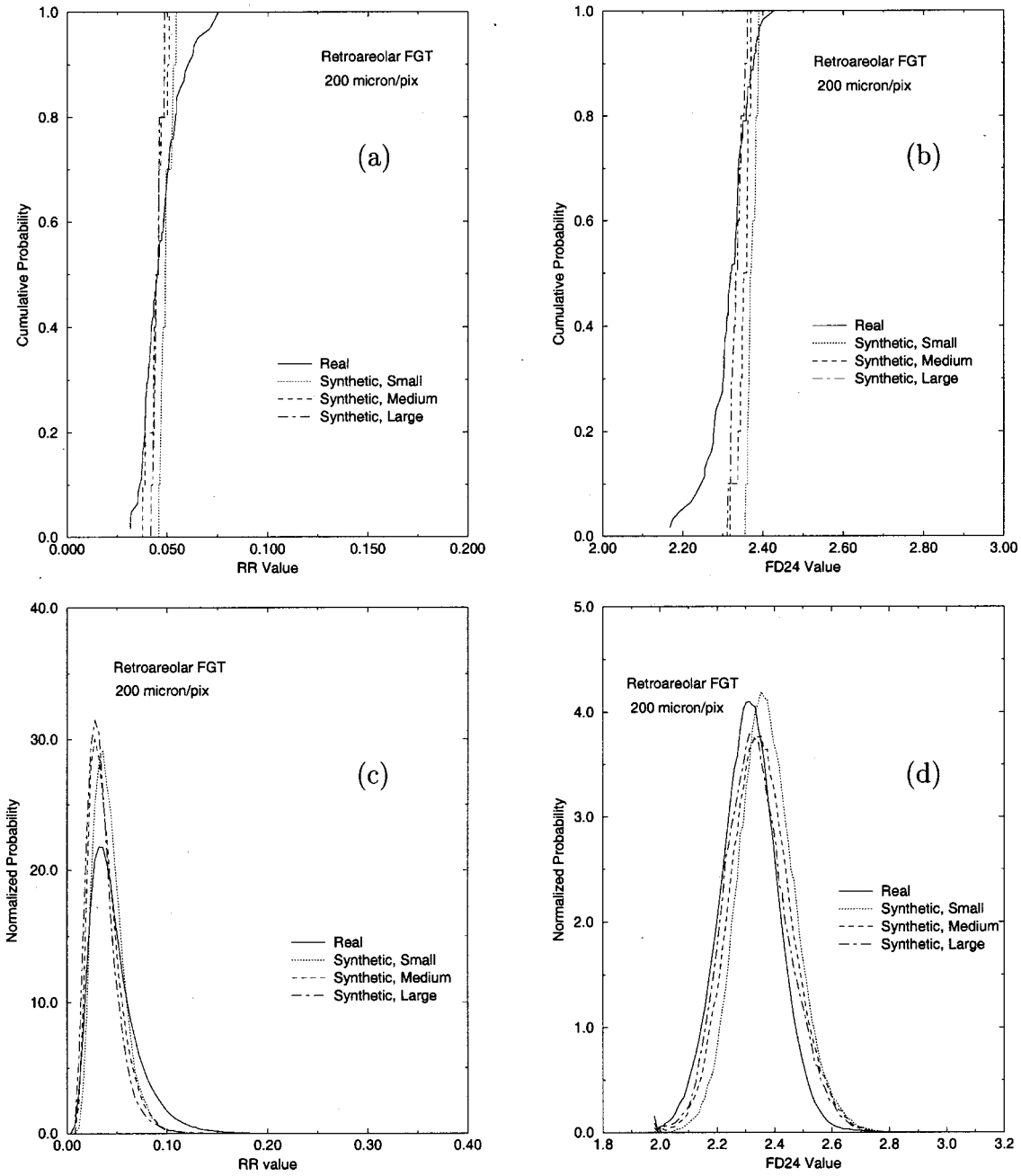

FIG. 5. Comparison between clinical and synthetic images of retroareolar fibroglandular tissue: CDFs of the (a) mean texture energy feature and (b) mean fractal analysis feature; average histograms of the (c) mean texture energy feature and (d) mean fractal analysis feature.

4(d). The gradient is used in place of the feature histogram for the size analysis. (Running averages of the gradients are shown for clarity.) Mean feature values, whose distributions were analyzed by the KS test, are also indicated on the graphs in Fig. 4(e). For texture energy and fractal dimension, the mean feature values were computed as the first moments of the feature histograms of each subimage. CDFs for all clinical and three groups of synthetic subimage means are shown in Fig. 6(b).

Figure 5 shows the CDF of the mean feature values and the average histograms for texture energy and fractal dimension computed for the samples of retroareolar glandular tissue regions in clinical and synthetic mammograms. Figure 6 gives the cumulative distributions of the first moments of the brightness gradients and the average brightness gradients for the size analysis of the clinical and synthetic samples of the subcutaneous adipose tissue (left) and the retroareolar glandular tissue (right).

III. RESULTS AND DISCUSSION

Results of the synthetic texture evaluation are presented in the form of graphs of the histogram correlation coefficients and maximum CDF differences computed for several texture features. Simulations were performed for three ranges of compartment size and for two spatial resolutions. Figures

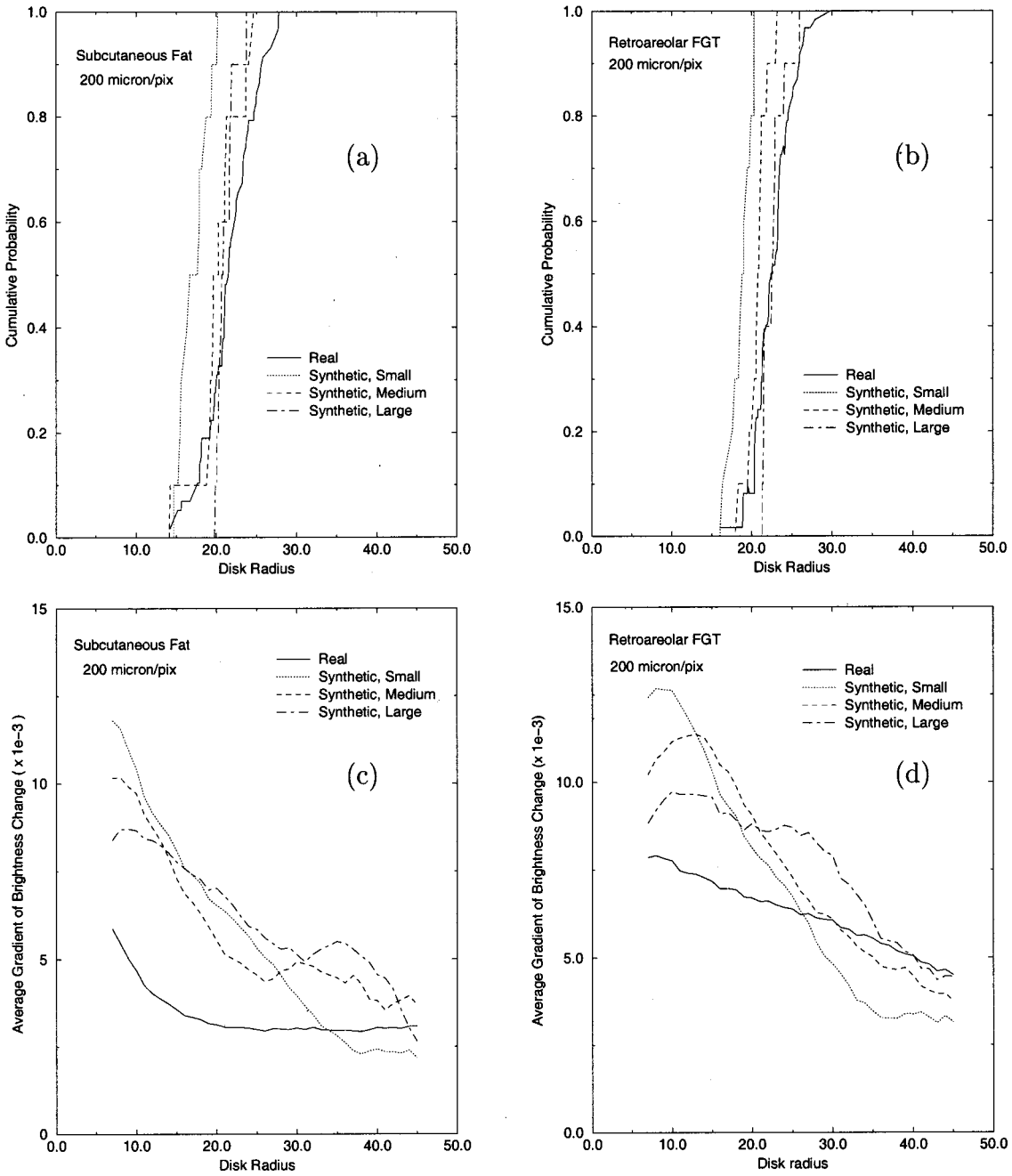

FIG. 6. Comparison between clinical and synthetic images using size analysis: CDFs of the mean features for (a) subcutaneous adipose tissue and (b) retroareolar fibroglandular tissue; average histograms for (c) subcutaneous adipose tissue and (d) retroareolar fibroglandular tissue.

7–10 show the results for subcutaneous adipose tissue, retromammary adipose tissue, retroareolar glandular and dense glandular tissue regions, respectively. The abscissa in these graphs is not a continuous variable, but indicates the size range of the simulated breast anatomic structures (adipose compartments). Texture features are labeled by different symbols: circle=average structure size, diamond=texture energy, and triangle=fractal dimension.

The graphs labeled (a) in Figs. 7–10 show the values of histogram correlation. Higher values indicate better correlation, with a maximum possible value of unity. A boot-strap analysis of the histogram correlation (averaged over all tissue types, texture measures, and synthetic structure sizes) was used to calculate a standard deviation of approximately 0.003. In the CDF difference graphs [labeled (b)], the maximum difference between the CDFs is equal to one. Thus, the lower the value, the better the agreement between the clinical and synthetic textures. The CDF difference is translated into a significance level, as explained previously, on the vertical axis at the right of the graphs. Values of the CDF difference corresponding to significance levels of 1%, 5%, and 20% are shown.

It can be seen that the size analysis (represented by circles) shows better agreement between the clinical and syn-

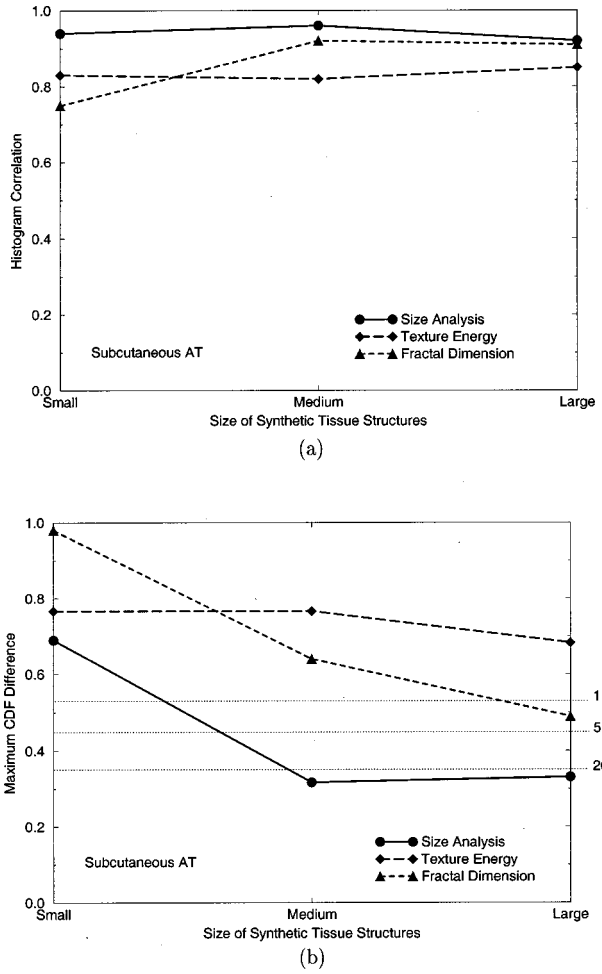

FIG. 7. Summary of (a) histogram correlation and (b) the results of the KS test, for subcutaneous adipose tissue. The data are presented for three ranges of size of synthetic tissue structures and for the three texture analysis methods.

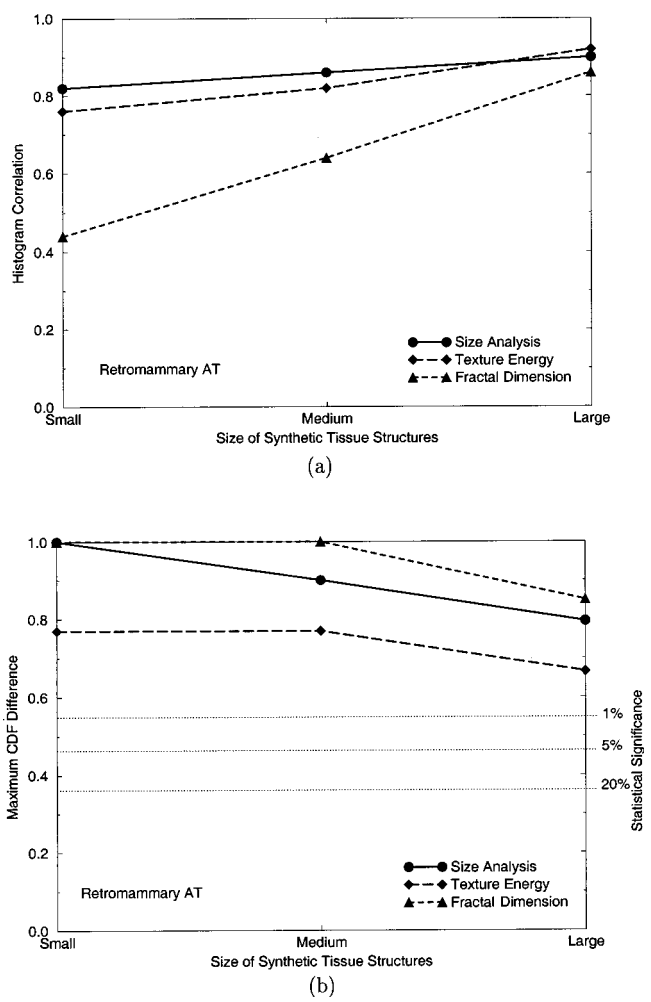

FIG. 8. Summary of (a) histogram correlation and (b) the results of the KS test, for retromammary adipose tissue. The data are presented for three ranges of size of synthetic tissue structures and for the three texture analysis methods.

thetic texture, than the other two features. This is expected as the size of the radiolucencies is related to the size of the adipose compartments, while the other features are more sensitive to local, small scale structure. The current version of our model does not include fine, local tissue detail.

Results for the retroareolar glandular tissue regions are shown in Fig. 9. Most of the feature values are concentrated very close to one in the histogram correlation graph, indicating good agreement between the simulation and the real mammograms. Also, the CDF difference for the retroareolar glandular tissue are lower that for the other tissue regions. The dense glandular region shows similarly good agreement (Fig. 10). By comparison, both glandular tissue regions are simulated better than the adipose regions (Figs. 7 and 8).

A repeated comparison between synthetic and real mammograms at a resolution of 400 μm/pixel (not plotted) showed that the simulation results were not affected significantly by the change of resolution. In addition, from Figs. 7 to 10 one can see that the analyzed features are sensitive to the size of simulated anatomical structures and that the agreement between the synthetic and clinical mammograms depends upon our selection of the simulated structure size. A partial analysis of a larger set of synthetic mammograms was also performed and no significant changes in the comparison with the set of real images were found.

The agreement between distributions of mean texture features suggests that the synthetic images sufficiently well simulate mean features of the clinical texture. Similarity between the averaged histograms of real and synthetic texture features means that our synthetic images can simulate the range of features seen over a large group of mammograms, not necessarily matching the feature distribution of any particular mammogram.

By varying the parameters which control the sizes of breast tissue model elements, we were able to match the average statistical properties of clinical mammograms for all tissue types except the retromammary fat. The best match for clinical mammogram texture was achieved for the simulated compartments with radii of 4–13.3 mm (''Medium'' and ''Large,'' as labeled in Table I in the accompanying paper[1]) in predominantly adipose tissue region, and with radii of 2.7–5.33 mm (''Large'') and 1.3–2.7 mm (''Small'') in the retroareolar and dense FGT region, respectively. These pa-

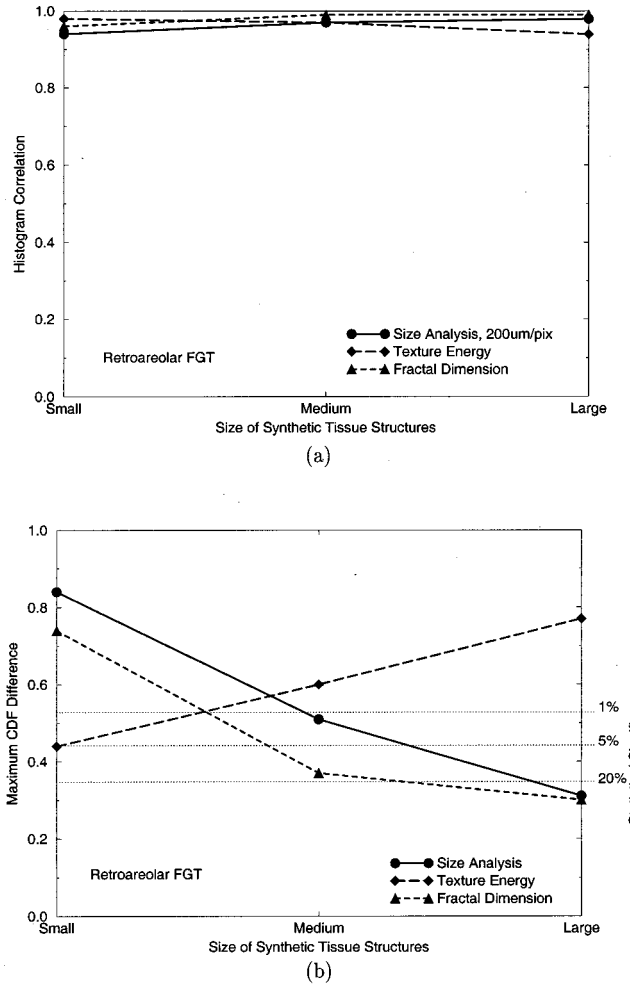

FIG. 9. Summary of (a) histogram correlation and (b) the results of the KS test, for retroareolar fibroglandular tissue. The data are presented for three ranges of size of synthetic tissue structures and for the three texture analysis methods.

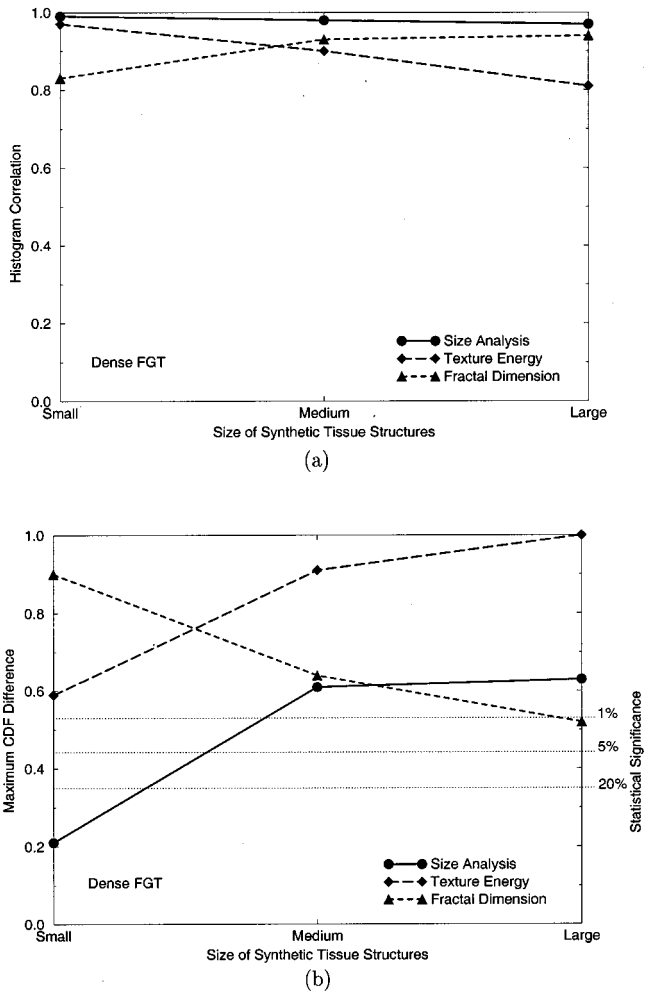

FIG. 10. Summary of (a) histogram correlation and (b) the results of the KS test, for dense fibroglandular tissue. The data are presented for three ranges of size of synthetic tissue structures and for the three texture analysis methods.

rameters were chosen because of high histogram correlation values and low maximum CDF difference; the corresponding distribution of the size analysis features for real and synthetic images cannot be distinguished at the 5% level. A difference between the retroareolar and dense FGT region is expected, since the retroareolar region contains more fat clustered in larger compartments than in dense regions. Further understanding of the clinical retromammary adipose tissue structure is needed to improve the simulation.

IV. CONCLUSIONS

Evaluation of the synthetic mammograms was performed by texture analysis and comparison with normal clinical mammograms from the MIAS database. By varying the distribution of tissue structures in the model we have been able to match some of the statistical properties of clinical mammograms. Quantitatively, the synthetic mammograms have a similar distribution of the values averaged over a large number of mammograms for several texture features, namely the average size of image objects, the texture energy, and the fractal dimension. The analysis of mammogram object size is closely related to the size analysis of the medium scale phantom elements, simulating adipose compartments in the breast, which are responsible for generation of the synthetic mammogram texture. This was the first such use of the morphological analysis of lucent mammogram regions, representing the projections of adipose compartments. Previous applications of morphology focused on the mammogram regions brighter than their surroundings, such as the fibrous structures and microcalcifications. The texture energy and the fractal dimension are more sensitive to the local variation of pixel intensities due to the small scale breast tissue structures.

Our model clearly captures the coarse tissue structures of the breast for all the tissue subregions except the retromammary fat. This exception is likely due to the fact that currently the retromammary adipose tissue is simulated in the same way as the subcutaneous tissue, although it is possible that the amount of fibroglandular tissue differs in these two regions. The model is less capable of capturing the small scale structures of the breast, e.g., blood and lymph vessels and fine tissue detail, which affect fine texture and give or-

ganized structure familiar to radiologists. The overly geometric appearance of the borders between the AT and FGT regions in the synthetic mammograms can be improved by small, random variations in the position of the borders in the compressed tissue model. Differences between the synthetic and clinical images are more evident for texture measures that emphasize smaller spatial scales, in agreement with the qualitative visual assessment. We expect that the introduction of detailed tissue structures in our breast model will enhance the local variations of synthetic mammograms and the variations in feature distribution needed to better match clinical images.

ACKNOWLEDGMENTS

This work was supported in part by NSF Grant No. IRI-9504363 and by US Army Breast Cancer Research Program Grant Nos. DAMD 17-96-1-6128 and DAMD 17-98-1-8169.

APPENDIX: BACKGROUND ON TEXTURE ANALYSIS

1. Mathematical morphology

Morphological operations are performed on an image, $f(x,y)$, using a second set of pixels, S, called the structuring element. The basic morphological operations, are defined by:[13]

$$\begin{aligned}&\text{Dilation: } (f\oplus S)(x,y)=\max\{f(x+x',y+y')|(x',y')\in S\},\\&\text{Erosion: } (f\ominus S)(x,y)=\min\{f(x-x',y-y')|(x',y')\in S\},\\&\text{Opening: } f\bigcirc S=(f\ominus S)\oplus S,\\&\text{Closing: } f\bullet S=(f\oplus S)\ominus S.\end{aligned}\tag{A1}$$

2. Texture energy analysis

2D convolution masks for texture energy analysis are derived using 1D masks with different number of zero-crossings, designed to detect different texture properties. For example, the five-element 1D masks are:[9]

$$\begin{aligned}&\text{Level: } L5=[1\ \ 4\ \ 6\ \ 4\ \ 1],\\&\text{Edge: } E5=[-1\ \ -2\ \ 0\ \ 2\ \ 1],\\&\text{Spot: } S5=[-1\ \ 0\ \ 2\ \ 0\ \ -1],\\&\text{Wave: } W5=[-1\ \ 2\ \ 0\ \ -2\ \ 1],\\&\text{Ripple: } R5=[1\ \ -4\ \ 6\ \ -4\ \ 1].\end{aligned}$$

The most often used 2D masks in mammogram analysis are 5×5 ''level'' and ''ripple'' masks, obtained by the product of the corresponding 1D masks ($R5R5=R5^{T}R5$ and $L5L5=L5^{T}L5$):

$$\text{R5R5}=\begin{bmatrix}1&-4&6&-4&1\\-4&16&-24&16&-4\\6&-24&36&-24&6\\-4&16&-24&16&-4\\1&-4&6&-4&1\end{bmatrix},$$

$$\text{L5L5}=\begin{bmatrix}1&4&6&4&1\\4&16&24&16&4\\6&24&36&24&6\\4&16&24&16&4\\1&4&6&4&1\end{bmatrix}.\tag{A2}$$

3. Fractal analysis

Image fractal dimension is usually defined using the area of the image surface. When the scale, ϵ, is increased (which corresponds to decreasing the resolution) the area of a fractal surface, $A(\epsilon)$, decreases. The fractal dimension, D, is related to the slope of decreasing area on a log–log plot, as

$$\log A(\epsilon)=\log\text{const}+(2-D)\log\epsilon.\tag{A3}$$

There are several algorithms for computing fractal dimension based upon box counting, image power spectrum, or iterated function systems.[26] Any attempt at measuring fractal dimension must deal with the fact that self-similarity of real images holds only over a limited range of scales, due to actual structure and limitations of the imaging process.[27]

[a]Author to whom correspondence should be addressed; electronic mail: Andrew.Maidment@mail.tju.edu

[1] P. R. Bakic, M. Albert, D. Brzakovic, and A. D. A. Maidment, ''Mammogram synthesis using 3D simulation. I. Breast tissue model and image acquisition simulation,'' Med. Phys. **29**, 2131–2139 (2002).
[2] F. O. Bochud, C. K. Abbey, and M. P. Eckstein, ''Statistical texture synthesis of mammographic images with clustered lumpy backgrounds,'' Opt. Express **4**, 33–43 (1999).
[3] J. J. Heine, S. R. Deans, R. P. Velthuizen, and L. P. Clarke, ''On the statistical nature of mammograms,'' Med. Phys. **26**, 2254–2265 (1999).
[4] P. Taylor, R. Owens, and D. Ingram, ''Simulated mammography using synthetic 3D breasts,'' in *Digital Mammography Nijmegen, 1998*, Proceedings of the Fourth International Workshop on Digital Mammography, Nijmegen, The Netherlands, June 1998, edited by N. Karssemeijer, M. Thijssen, J. Hendriks, and L. van Erning (Kluwer Academic, Dordrecht, 1998), pp. 283–290.
[5] P. R. Bakic, ''Breast tissue description and modeling in mammography,'' Ph.D. dissertation, Lehigh University, Bethlehem, PA, 2000.
[6] P. R. Bakic, M. Albert, and A. D. A. Maidment, ''Evaluation of breast ductal networks using ramification matrices,'' in *IWDM 2002*, Proceedings of the Sixth International Workshop on Digital Mammography (Bremen, Germany, June 2002).
[7] J. Suckling *et al.*, ''The Mammographic Image Analysis Society Digital Mammogram Database,'' in *Digital Mammography*, Proceedings of the Second International Workshop on Digital Mammography, York, England, 10–12 July 1994, edited by A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns (Elsevier Science, Amsterdam, 1994), pp. 375–378 (http://www.wiau.man.ac.uk/services/MIAS/MIASmini.html).
[8] J. Serra, *Image Analysis and Mathematical Morphology* (Academic, London, 1982).
[9] K. I. Laws, ''Textured Image Segmentation,'' Ph.D. dissertation, University of Southern California, Los Angeles, CA, 1980.
[10] S. Peleg, J. Naor, R. Hartley, and D. Avnir, ''Multiple resolution texture analysis and classification,'' IEEE Trans. Pattern Anal. Mach. Intell. **6**, 518–523 (1984).

[11] G. Matheron, *Random Sets and Integral Geometry* (Wiley, New York, 1975).
[12] R. M. Haralick, S. R. Sternberg, and X. Zhuang, ''Image analysis using mathematical morphology,'' IEEE Trans. Pattern Anal. Mach. Intell. **9**, 532–550 (1987).
[13] S. Behrens and J. Dengler, ''Analyzing the structure of medical images with morphological size distributions,'' Proceedings of the Tenth International Conference on Pattern Recognition, Atlantic City, NJ, June 1990, pp. 886–890.
[14] P. I. Miller and S. M. Astley, ''Classification of breast tissue by texture analysis,'' Image Vis. Comput. **10**, 277–282 (1992).
[15] M. Unser, ''Local linear transforms for texture measurements,'' Signal Process. **11**, 61–79 (1986).
[16] W. P. Kegelmeyer, Jr. and M. C. Allmen, ''Dense feature maps for detection of calcifications,'' in *Digital Mammography,* Proceedings of the Second International Workshop on Digital Mammography, York, England, 10–12 July 1994, edited by A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns (Elsevier Science, Amsterdam, 1994), pp. 3–12.
[17] N. Vujovic and D. Brzakovic, ''Establishing correspondence between control points in pairs of mammographic images,'' IEEE Trans. Image Process. **6**, 1388–1399 (1997).
[18] M. Sallam and K. Bowyer, ''Registration time sequences of mammograms using a two-dimensional image unwarping technique,'' in *Digital Mammography,* Proceedings of the Second International Workshop on Digital Mammography, York, England, 10–12 July 1994, edited by A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns (Elsevier Science, Amsterdam, 1994), pp. 121–130.
[19] C. B. Caldwell, S. J. Stapleton, D. W. Holdsworth, R. A. Jong, W. J. Weiser, G. Cooke, and M. J. Yaffe, ''Characterization of mammographic parenchymal pattern by fractal dimension,'' Phys. Med. Biol. **35**, 235–247 (1990).
[20] F. Lefebvre, H. Benali, R. Gilles, E. Kahn, and R. Di Paola, ''A fractal approach to the segmentation of microcalcifications in digital mammograms,'' Med. Phys. **22**, 381–390 (1995).
[21] H. Li, K. J. R. Liu, and S.-C. B. Lo, ''Fractal modeling and segmentation for the enhancement of microcalcifications in digital mammograms,'' IEEE Trans. Med. Imaging **16**, 785–798 (1997).
[22] X. Liu, G. Zhang, and M. D. Fox, ''Fractal description and classification of breast tumors,'' Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1991, Vol. 13, pp. 112–113.
[23] J. N. Wolfe, ''Risk for breast cancer development determined by mammographic parenchymal patterns,'' Cancer **37**, 2486–2492 (1976).
[24] A. G. Frodesen, O. Skjeggestad, and H. Tofte, *Probability and Statistics in Particle Physics* (Universitetsforlaget, Oslo, 1979).
[25] G. K. Kanji, *100 Statistical Tests* (SAGE, London, 1999).
[26] M. Barnsley, *Fractals Everywhere* (Academic, Boston, 1988).
[27] A. I. Penn and M. H. Loew, ''Estimating fractal dimensions with fractal interpolation function models,'' IEEE Trans. Med. Imaging **16**, 930–937 (1997).