Computer-aided detection of clustered microcalcifications in multiscale bilateral filtering regularized reconstructed digital breast tomosynthesis volume

Ravi K. Samala,[a] Heang-Ping Chan, Yao Lu, Lubomir Hadjiiski, and Jun Wei
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-5842

Berkman Sahiner
Center for Devices and Radiological Health, U.S. Food and Drug Administration, Maryland 20993

Mark A. Helvie
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-5842

(Received 31 July 2013; revised 18 December 2013; accepted for publication 18 December 2013; published 10 January 2014)

Purpose: Develop a computer-aided detection (CADe) system for clustered microcalcifications in digital breast tomosynthesis (DBT) volume enhanced with multiscale bilateral filtering (MSBF) regularization.
Methods: With Institutional Review Board approval and written informed consent, two-view DBT of 154 breasts, of which 116 had biopsy-proven microcalcification (MC) clusters and 38 were free of MCs, was imaged with a General Electric GEN2 prototype DBT system. The DBT volumes were reconstructed with MSBF-regularized simultaneous algebraic reconstruction technique (SART) that was designed to enhance MCs and reduce background noise while preserving the quality of other tissue structures. The contrast-to-noise ratio (CNR) of MCs was further improved with enhancement-modulated calcification response (EMCR) preprocessing, which combined multiscale Hessian response to enhance MCs by shape and bandpass filtering to remove the low-frequency structured background. MC candidates were then located in the EMCR volume using iterative thresholding and segmented by adaptive region growing. Two sets of potential MC objects, cluster centroid objects and MC seed objects, were generated and the CNR of each object was calculated. The number of candidates in each set was controlled based on the breast volume. Dynamic clustering around the centroid objects grouped the MC candidates to form clusters. Adaptive criteria were designed to reduce false positive (FP) clusters based on the size, CNR values and the number of MCs in the cluster, cluster shape, and cluster based maximum intensity projection. Free-response receiver operating characteristic (FROC) and jackknife alternative FROC (JAFROC) analyses were used to assess the performance and compare with that of a previous study.
Results: Unpaired two-tailed t-test showed a significant increase ($p < 0.0001$) in the ratio of CNRs for MCs with and without MSBF regularization compared to similar ratios for FPs. For view-based detection, a sensitivity of 85% was achieved at an FP rate of 2.16 per DBT volume. For case-based detection, a sensitivity of 85% was achieved at an FP rate of 0.85 per DBT volume. JAFROC analysis showed a significant improvement in the performance of the current CADe system compared to that of our previous system ($p = 0.003$).
Conclusions: MBSF regularized SART reconstruction enhances MCs. The enhancement in the signals, in combination with properly designed adaptive threshold criteria, effective MC feature analysis, and false positive reduction techniques, leads to a significant improvement in the detection of clustered MCs in DBT. © *2014 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4860955]

Key words: digital breast tomosynthesis, computer-aided detection, microcalcification, multiscale enhancement, regularized reconstruction

1. INTRODUCTION

A recent study[1] of 205 radiologists interpreting over one million screening mammograms found the median sensitivity to be 83.8% with a median recall rate of 9.3% requiring diagnostic workup for potential abnormal findings. Diagnostic workup and biopsy of false positives (FP) increase health care costs and patient anxiety. Digital breast tomosynthesis (DBT) is a new quasi three-dimensional (3D) breast imaging modality. DBT has the potential to reduce false positives due to overlapping fibroglandular tissue and reduce false negatives caused by camouflaging of lesions by overlying structures. Several preliminary clinical studies have found an improvement in mass detection and a reduction in recall rates using DBT in comparison to digital mammography (DM).[2–9] Comparison of mass visibility between DBT and DM showed

consistent results in favor of DBT.[3,5,10,11] DBT, thus, is advantageous for breast mass detection and diagnosis, especially in breasts with heterogeneously dense parenchyma.[2–9]

In addition to improving the sensitivity and specificity for mass detection, the ability of DBT to detect microcalcification clusters (MCs), which may be the only sign of early breast cancer, is an important consideration for clinical application of DBT as a standalone screening modality. Although DBT improves the conspicuity of soft tissue lesions compared to DM by reducing breast parenchymal complexity, it poses challenges for detection of MCs. These challenges include: (a) an MC may be split among many slices, (b) blurring due to reconstruction from multiple angular projections, (c) blurring due to potential patient motion, and (d) focal spot blur in DBT systems with continuous x-ray source motion.[12] The detectability of subtle signals is also strongly affected by the tomographic angle and the angular distribution of the projections of the DBT system.[13–17] Preliminary results from studies comparing the detectability of MCs in breasts between DBT and DM varied. Some studies[2,18] indicate that MC detection in DBT is inferior to that of DM, while others found that DBT is comparable[3] or even superior[19] to DM. The different results may be attributed to the differences in the aforementioned DBT system design parameters, reconstruction techniques, as well as other human factors such as experience in DBT reading, reader fatigue caused by the large number of slices per DBT volume, and lower visual quality and fewer numbers of microcalcifications on each slice and their morphology. With the success of computer-aided detection (CADe) for detection of MCs in DM, it is expected that CADe may also play an important role in improving MC detection in DBT.

Limited preliminary studies on CADe for MCs in DBT have been performed using small datasets. The detection was performed in projection views (PVs), the reconstructed slices, or in the reconstructed volume (RV). Peters *et al.*[20] detected calcifications on a small set of DBT. A bandpass filter based wavelet kernel was used to separate the potential calcification candidates from the background on the PVs. A feature map was generated for each PV image and the correspondence between two-dimensional (2D) and 3D locations determined by the DBT acquisition geometry was used as a criterion to identify the calcifications. Park *et al.*[21] applied a 2D CADe algorithm developed for digitized screen-film mammograms (SFM) to the PV and the reconstructed DBT slices. Reiser *et al.*[22] developed an algorithm to detect MCs in PV images to avoid the dependence of the CADe performance on the reconstruction algorithm. Van Schie and Karssemeijer[23] estimated a nonuniform noise model from each individual DBT RV which was used for normalization of the local contrast feature. Potential microcalcifications were detected by thresholding of the local contrast feature and the microcalcification candidates within 5 mm radius were grouped to form MCs.

We are developing a CADe system for microcalcification cluster detection in DBT. In a previous study,[24] our CADe system achieved a sensitivity of 85% at an FP rate of 3.4 per DBT volume. Our laboratory recently developed a regularized simultaneous algebraic reconstruction technique (SART) that utilizes multiscale bilateral filtering (MSBF) during the iterative reconstruction to reduce noise and enhance the signal strength of microcalcifications.[25] In this study, we investigated the detection of microcalcifications in the DBT volume reconstructed by the MSBF regularized SART. The CADe system was trained to exploit the enhanced signal properties and the overall performance was compared to the previous results.[24]

2. METHODS AND MATERIALS

A General Electric GEN2 prototype system was used for DBT imaging. The system uses a step-and-shoot design, a stationary flat panel CsI/a:Si detector with a pixel pitch of 0.1 mm × 0.1 mm, Rh anode, and Rh filter, and acquires 21 PVs in 3° increments in a total tomographic angle of 60°. DBT reconstruction was performed using SART with multiscale bilateral regularization. All DBT volumes were reconstructed at 1 mm slice spacing with an in-plane resolution of 0.1 mm × 0.1 mm. The flow diagram in Fig. 1 can be broadly divided into four processes: (a) regularized reconstruction with enhancement (Sec. 2.B), (b) preprocessing to

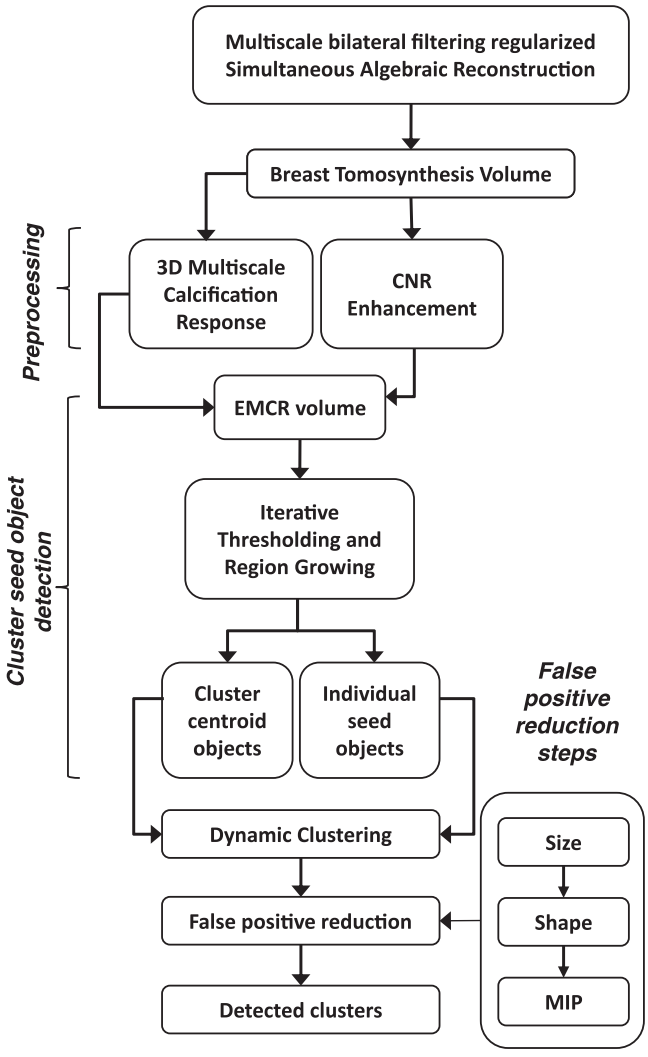


FIG. 1. Schematic of our CADe system for microcalcification cluster detection in DBT.

enhance the signals and suppress the structured background (Sec. 2.C), (c) detection and segmentation of cluster seed objects and signal candidates (Sec. 2.D), (d) dynamic clustering for cluster formation and false positive reduction to refine the detected clusters (Sec. 2.E).

2.A. Data set

With approval of Institutional Review Board (IRB) and written informed consent, 154 subjects were recruited from the Breast Imaging Division, Department of Radiology at the University of Michigan Health System. Craniocaudal (CC) and mediolateral oblique (MLO) views of DBT were acquired from the breast recommended for biopsy of a suspicious lesion. Of the 154 breasts, 116 had biopsy proven MCs (34 malignant, 120 benign) and 38 were free of MCs. The dataset was split into two independent subsets for training and validation of the CADe system. The training set consisted of 64 cases (127 views – one view was lost due to technical problem) with MCs and the remaining 52 MC cases and 38 cases free of MC (total 180 views) were used for testing. A Mammography Quality Standards Act (MQSA) approved radiologist marked the location of the biopsied MC with a 3D bounding box based on all clinical information available. For the training set of 127 views and the test set of 104 views, 124 and 100 biopsy-proven MCs were marked, respectively. For three breasts in the training set and four breasts in the test set, the MC was visible only in one of the DBT views.

2.B. Multiscale bilateral filtering regularized tomosynthesis reconstruction

We have been investigating regularized SART to improve the contrast-to-noise ratio (CNR) of microcalcifications in DBT for both human and machine detection.[25–27] Our early studies indicated that selective diffusion regularization method is superior to total variation regularization in terms of the CNR of microcalcifications. However, fine fibrous structures and mass spicules may be blurred and very subtle microcalcifications may be inadvertently smoothed as noise if they failed the selective diffusion criterion that depends on the local gradient of the signals. Recently, we developed a new MSBF regularized SART that can enhance the high frequency structures such as microcalcifications and suppress noise without smoothing the spicules and fibrous tissue. The details of the technique have been described in Ref. 25. In brief, regularization is performed at the end of each SART iteration. A reconstructed DBT slice is first separated into low and high frequency bands using Laplacian pyramid decomposition.[28] Bilateral filtering is applied to the high frequency bands that contain the noise and microcalcifications. By properly choosing the domain and range filter parameters, the bilateral filter[29] can selectively enhance the microcalcifications while suppressing the noise. The DBT slice is then reconstructed from all levels of the Laplacian pyramid. The DBT volume is used as an input to the CADe system. In our discussion, a DBT volume denotes the collection of all slices reconstructed from the 21 PVs acquired for a DBT scan in a given view (CC or MLO view).

2.C. Preprocessing

The purpose of preprocessing in a CADe system is to enhance the lesion of interest and suppress the structured anatomical background to improve signal detection. It is observed that microcalcifications are compact, dense objects and have higher CNR compared to soft tissue structures of comparable sizes. Both of these properties are exploited using shape information from second order derivatives (Sec. 2.C.1) and contrast information from bandpass filtering (Sec. 2.C.2).

2.C.1. Multiscale calcification response

As described in our previous work,[24] for enhancement of microcalcifications, the input DBT volume is processed by 3D multiscale filtering and a calcification response is derived from the Hessian matrices at each voxel. The same process is applied to the regularized reconstructed DBT volume.

Second order spatial differential operators have been used widely to extract features based on edges. The Hessian matrix is a symmetric matrix consisting of second order partial derivatives of the local structure. The three eigenvalues of the Hessian matrix characterize the shape of an object at that spatial location. To reduce the variation in the second derivatives due to noise, the DBT volume $I(x, y, z)$ is first convolved with a 3D Gaussian smoothing function at a scale σ:

$$s_\sigma(x, y, z) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\{-(x^2 + y^2 + z^2)/2\sigma^2\}, \quad (1)$$

which results in a smoothed volume $f(x, y, z)$. For every voxel in $f(x, y, z)$, the Hessian matrix is given by

$$H_\sigma(x, y, z) = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}. \quad (2)$$

The eigen decomposition of the H_σ gives the principal axes $(\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3)$ of the local structure. The corresponding eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ can be effectively used to determine the shape of the object. The response function (r_σ) is designed to have high values if the object is an approximately spherical structure at scale σ:

$$r_\sigma = \begin{cases} \lambda_3^2/\lambda_1 & \text{if } \lambda_1 \le \lambda_2 \le \lambda_3 \le 0 \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of the Hessian matrix at (x, y, z) and scale σ. The response value is large if the scale matches the size of the object. Because the sizes of microcalcifications are variable and the size of a potential object at a given location is not known, this process is performed at multiple scales σ_i, $i = 1, \ldots, n$, to cover a range of possible sizes. The multiscale calcification response (MCR) is then defined as $E(x, y, z)$ at a given location over different scales:

$$E(x, y, z) = \frac{r_{\sigma_{i^*}}}{\sigma_{i^*}}, \quad \text{where } i^* = \arg\max_i\{r_{\sigma_i}\}. \quad (4)$$

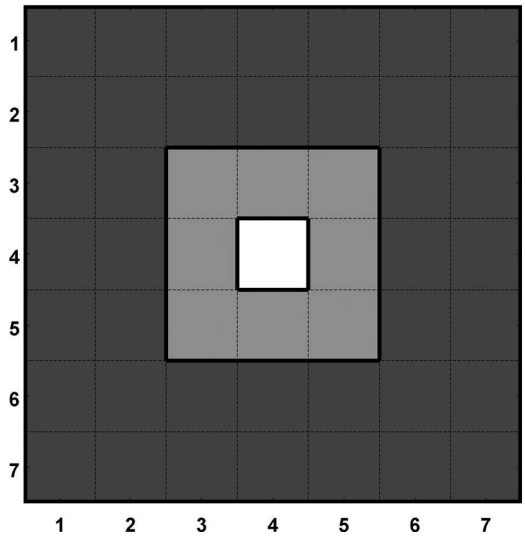


FIG. 2. Filter $F(x, y)$ used for CNR enhancement, which is composed of three box filters with kernel sizes, $M_1 = 7, M_2 = 3, M_3 = 1$.

With this approach for microcalcification detection, both the object intensity and the shape information are taken into consideration, and high response can be obtained from microcalcifications of various sizes.

2.C.2. CNR enhancement

Detection of microcalcifications is affected significantly by the presence of anatomical structured noise. We implement a 2D bandpass filter to suppress the structured background and simultaneously enhance the signal, thereby enhancing the CNR. The CNR enhancement filter is constructed from a combination of three filters (F_1, F_2, F_3), each with specific purpose. Filters F_1 and F_2 with kernel sizes of $M_1 \times M_1$ and $M_2 \times M_2$, respectively, are combined to obtain an estimate of the background and filter F_3 with a kernel size of $M_3 \times M_3$ is used to enhance the signal. Each filter kernel weight is $1/M_i^2$ in a $M_i \times M_i$ box and 0 elsewhere. The combined bandpass filter $F(x, y)$ selectively enhances the signal and simultaneously suppresses the background. The 2D filter is convolved with the original input DBT volume slice-by-slice, resulting in a CNR enhanced volume (C). An example of the bandpass filter kernel with $M_1 = 7, M_2 = 3, M_3 = 1$ is shown in Fig. 2:

$$F(x, y) = F_3(x, y) - \frac{1}{M_1^2 - M_2^2} \times \left[M_1^2 F_1(x, y) - M_2^2 F_2(x, y)\right], \quad (5)$$

where $M_1 > M_2 \geq M_3$;

$$C(x, y, z_j) = I(x, y, z_j) \otimes F(x, y), \quad (6)$$

where $j = 1, \ldots, K$, and K is the total number of slices in the volume.

2.C.3. EMCR volume

An enhancement-modulated calcification response (EMCR) volume is generated from the CNR enhanced and multiscale calcification response volumes. In the EMCR volume, subtle signals are further emphasized because true microcalcifications should have relatively high voxel values in both volumes. We previously weighted the MCR with the CNR enhanced volume:

$$\mathrm{EMCR}_1(x, y, z) = E(x, y, z) * C(x, y, z). \quad (7)$$

A more flexible weighting method is considered in this study:

$$\mathrm{EMCR}_2(x, y, z) = \exp\{2[w_1 \log(E(x, y, z)) + w_2 \log(C(x, y, z))]\}, \quad (8)$$

where (w_1, w_2) are adjustable weights that can selectively emphasize one of the volumes. Figure 3 illustrates different stages of preprocessing steps for an MC that was biopsy-proven to be ductal carcinoma *in situ* (DCIS).

2.D. Cluster and object seed detection

The aforementioned preprocessing steps produce a new volume in which the potential microcalcifications are enhanced and the low frequency background is removed. The subsequent steps are prescreening, segmentation, quantification of the signal characteristics, false positive reduction, and eventually forming clusters.

2.D.1. Seed object detection

The voxel values in the EMCR volume are CNR-weighted calcification response that indicates the presence of a microcalcification. The stronger the response value, the higher the likelihood. The N highest response objects are identified as the cluster centroid objects that will be used as starting points for forming clusters. The top N_s objects are identified as the individual seed objects that will be used as seed points for segmentation of the individual objects and as the potential members of a cluster. The cluster centroid object set is a subset of the individual seed object set.

High response objects are found in the EMCR volume by an iterative process that combines thresholding and region growing. The initial EMCR threshold is first chosen to be high enough to detect only few objects. Voxels above the threshold are grouped into 3D objects using 26-connectivity.[30] The threshold is then reduced iteratively until the preset numbers of N centroid objects and top N_s seed objects are found. For each detected 3D object, the maximum EMCR voxel value of the object is used as a score to rank the top N and top N_s signals. The coordinates of the voxel with the maximum EMCR value are used to identify the 3D location of the object.

The numbers of candidates for the cluster seeds N and the individual calcifications N_s are chosen adaptively according to the breast volume in the DBT view. Based on analysis of the breast volumes of the training cases and the number of prescreening candidates to include at this stage, the breast volumes are partitioned into four brackets for the N and N_s values as shown in Fig. 4. In the largest volume range of greater than 2000×10^3 mm^3, the N and N_s values are chosen to be 100 and 800, respectively. The values are reduced to 80%, 50%, and 40% for the three smaller volume ranges. The purpose of

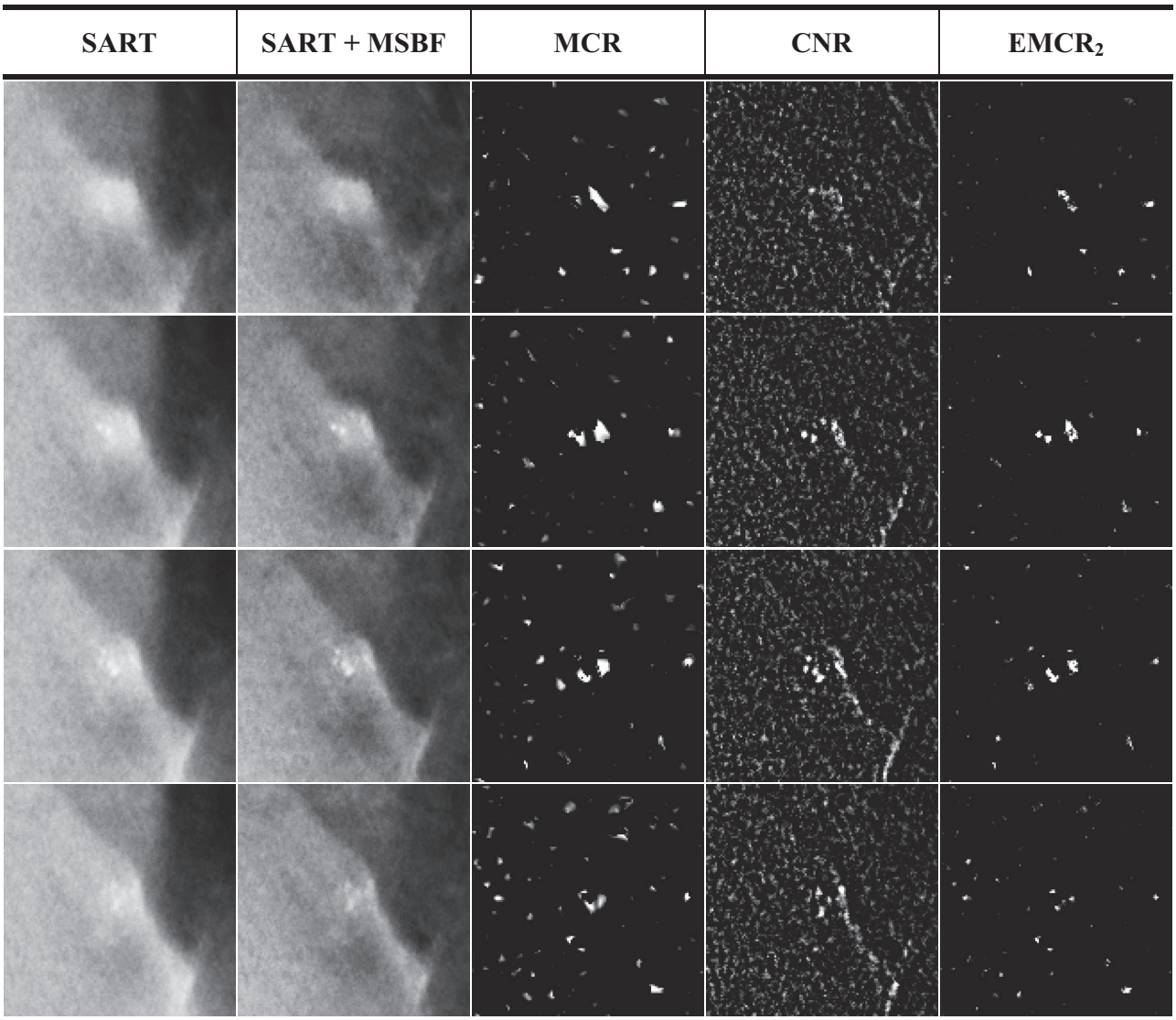


FIG. 3. Different stages of CADe preprocessing module (Horizontally: SART-Simultaneous algebraic reconstruction technique, MSBF-Multiscale bilateral filtering regularization, MCR-Multiscale calcification response, CNR-CNR enhanced volume, $EMCR_2$-Enhancement-modulated calcification response). (Vertically: slices 24–27). The case shown is biopsy-proven to be ductal carcinoma *in situ*. The ROI shown is 1.3 × 1.3 mm (130 × 130 pixels).

stratification of the volume range is to approximately maintain the number of seed objects per unit breast volume, which improves the robustness of the CADe system for breasts of different sizes.

2.D.2. CNR calculation and segmentation

For each seed object determined in the previous step, object segmentation is performed in the CNR enhanced volume. Local foreground (FG) and background (BG) information in the CNR enhanced volume is estimated for the seed object in a 5.1 × 5.1 × 5 mm box centered at the geometric center of the object. Starting from the center location of the seed object, region growing based on 26-connectivity and a CNR criterion is used to determine if a neighboring voxel is a part of the object. For a given voxel i, the CNR_i is calculated as

$$\mathrm{CNR}_i = \frac{V_i - \mu_{\mathrm{BG}}}{\sigma_{\mathrm{BG}}}, \tag{9}$$

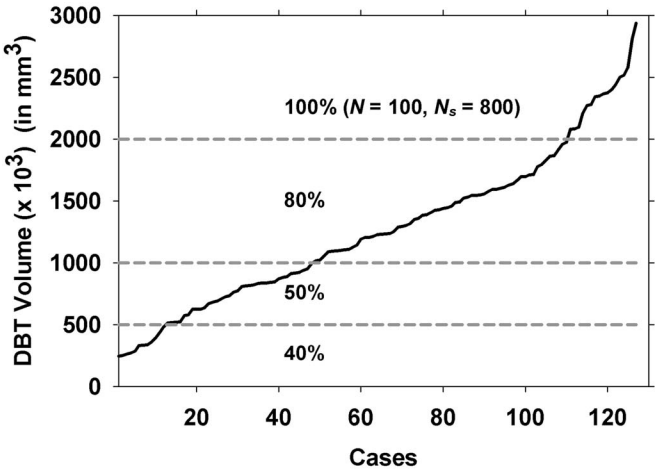


FIG. 4. The dependence of the numbers of top N and N_s objects based on the breast volume in the DBT views of the training set.

where V_i is the voxel value in the CNR-enhanced volume, μ_{BG} and σ_{BG} are the mean and standard deviation estimated in the 5.1 × 5.1 × 5 mm box excluding all voxels that are determined to belong to potential objects in the previous seed object detection step. Object growth and segmentation are terminated if no more neighboring voxel has CNR greater than a chosen threshold. A CNR threshold of 3.0 is used as in our previous study. The CNR of the object, CNR_{obj}, is defined as the maximum CNR among all voxels of the grown object:

$$\mathrm{CNR}_{obj} = \frac{V_{max} - \mu_{BG}}{\sigma_{BG}}. \quad (10)$$

Feature extraction is subsequently performed on the segmented object. Any object connected more than 12 mm in the depth direction is considered artifact and removed. Artifacts generated from surgical clips and large dense calcifications are removed through this criterion. Size and eccentricity of each object are calculated. Based on the training set, criteria are chosen such that objects with number of voxels >30 voxels and eccentricity >3 are considered large benign calcifications or FPs and are eliminated from the N_s object list.

2.D.3. Adaptive CNR thresholds

The CNR_{obj} values and number of microcalcifications within an MC vary over a wide range. To analyze the characteristics of MCs, a subset of 20 cases (40 DBT volumes) were randomly selected from the training set and the locations of the individual microcalcifications in the cluster were manually marked by an experienced medical physicist within the bounding box of the cluster identified by the radiologist.

Figure 5(a) shows the number of manually marked microcalcifications in the 40 clusters and the subset of those included in the top N_s objects at the prescreening stage. Figure 5(b) shows the mean and standard deviation of the CNR_{obj} values for the manually marked microcalcifications within the top N_s objects in each cluster and the mean and standard deviation of the CNR_{obj} values for all top N_s objects excluding those overlapped with the manually marked microcalcifications, which would be predominantly FPs. It can be seen that there are large variations in the distributions of the CNR_{obj} values among the different DBT volumes except that the mean CNR_{obj} for the FPs is relatively consistent. In order to compensate for such a variation, thresholds are estimated adaptively for the individual volumes. These thresholds are used to classify the detected objects as noise, subtle, moderate, or strong. For each DBT volume, based on the distribution of the CNR_{obj} values of the N_s objects, four thresholds are generated as

$$T_t = \mu + k_t * \sigma_{CNR}, \quad (11)$$

where μ is the mean, σ_{CNR} is the standard deviation of the CNR_{obj} distribution, and T_t with $t = 0, \ldots, 3$ are the adaptive CNR thresholds. T_t are determined automatically from the CNR distributions of the candidate objects in the breast volume being analyzed and the k_t values. The k_t values and the usage of each of the four thresholds are described in Sec. 2.E.

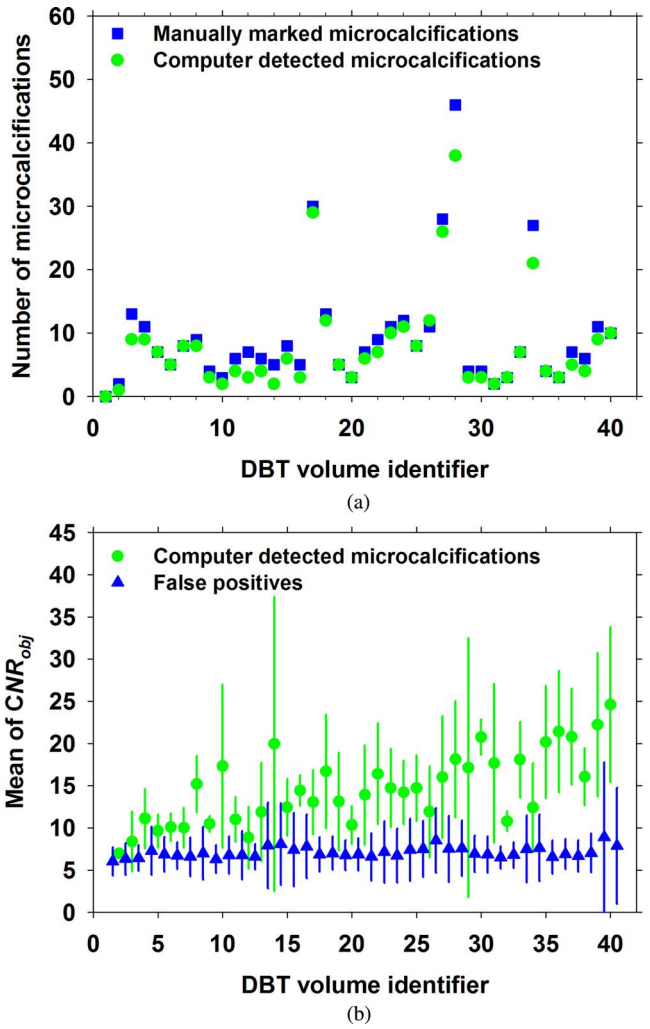


FIG. 5. (a) Number of microcalcifications within each radiologist provided 3D box for a subset of 20 cases (40 DBT volumes) from the training set. Squares: manually marked microcalcifications. Circles: the subset of top N_s objects that overlapped with manually marked microcalcifications. The DBT volume numbers followed the same ordering as in Fig. 5(b). (b) Mean and standard deviation of the CNR_{obj} values in the same 40 DBT volumes. Circles: the top N_s objects that overlapped with manually marked microcalcifications. Triangles: all N_s objects in the DBT volume excluding those overlapped with the manually marked MCs. The DBT volume numbers were ordered by sorting the mean CNR_{obj} values of the true microcalcifications (circles) in the CC-view DBT and assigned them the odd numbers, and the MLO view of the same case was ordered next to its CC view (even numbers). The DBT volume numbers of the false positives (triangles) followed the same order as the true microcalcifications but were offset by 0.5 to avoid overlap of the error bars (±1 standard deviation). Note: The cluster in DBT volume 1 was not visible, as determined by the radiologist so there was no CNR_{obj} value.

2.E. Cluster detection and false positive reduction

2.E.1. Dynamic clustering

The top N cluster seeds and the top N_s objects are separately ranked based on their CNR_{obj} values. A dynamic clustering process groups the objects into clusters as follows. Starting from the highest ranked cluster seed, clustering starts if the given seed has a CNR value greater than the T_0 threshold. This is the initial cluster centroid. From the cluster centroid object, it searches for cluster members from the N and N_s

lists. The objects within a distance of 5 mm and greater than the T_1 threshold are added to the cluster sequentially. During the addition of each individual object to the cluster, the cluster centroid location is dynamically updated based on the collective geometric center of the objects already included as cluster members. This cluster size is, therefore, not limited to 10 mm in diameter and can potentially grow to cover larger clusters. Dynamic clustering extends the cluster size adaptively, which will reduce the chance of grouping widely scattered calcifications and FPs to be a cluster compared to the use of a fixed and large radius for clustering. Any object that has been included as a member of a cluster will be eliminated from the lists. A cluster stops growing when no more objects satisfy the criteria. The next highest ranked cluster seed on the N list will then be used to grow the next cluster. The dynamic clustering process ends when the N list is exhausted. The k_0 and k_1 values were experimentally chosen to be 0.2 and −0.1, respectively, based on the training set.

2.E.2. False positive reduction

The cluster formation stage only considers the distance from the centroid object and the CNR values of the objects. FP clusters can be formed from individual FPs in close proximity that originate from different sources. Some examples of FP clusters are shown in Fig 6. The following FP reduction steps make use of the features of microcalcifications and clusters to distinguish the true and FP clusters.

Size, CNR, and number of objects in cluster. Object-based features such as size (SZ_t), CNR, number of objects (c_t), within each cluster are combined to form cluster features. Three rules are designed as shown in Table I to identify clusters of a range of subtlety. The adaptive CNR thresholds T_t are determined by Eq. (11) and the k_t values. The k_2 and k_3 values were chosen experimentally to be 0.5 and 0.6, respectively, from the training set. The SZ_t values (t = 1, 2, 3) were experimentally selected to be 5, 8, and 14 voxels, representing subtle, medium, and obvious MCs, respectively.

Because the breast boundaries and the vessels and connective tissue in the breast periphery cause strong enhancement in tomosynthesis reconstruction and multiscale bilateral regularization, more stringent criteria are used to reduce FPs near the breast peripheral region. Thus, for clusters with its centroid falling within 6 mm of the breast boundary, only rules 1 and 2 will be applied.

Cluster shape. The shape of the bounding box of a cluster is used to estimate the distribution of the calcifications in the cluster. A long and narrow shaped box may indicate FPs caused by artifacts of high density objects in the breast such as metallic markers or benign calcifications. Given the bounding box of a cluster, the cluster is eliminated as FP if the smaller dimension of the box is smaller than a narrow dimension threshold of 3 mm and the rectangularity of the box is greater than 7.

Maximum intensity projection (MIP). Some of the curved or branching tissue or vascular structures appear as multiple disconnected objects in 3D and form a cluster. To detect this type of FP clusters, MIP of a cluster in the CNR enhanced domain is projected along the depth direction to the (x, y) plane. The CNR based segmentation (see Section 2.D.2) is then performed for every object in the MIP plane. If the objects are connected on the MIP plane and the number of objects in the cluster is reduced to below the threshold, then the cluster is classified as FP.

2.F. Performance analysis

Rank-sensitivity plot. To optimize the parameter selection during the training stage, a curve similar to a free response receiver operating characteristic (FROC) plot is used to evaluate the detection performance of cluster seed objects at the prescreening stage. The x-axis of the rank-sensitivity plot is the rank threshold R of the highest rank objects based on the EMCR score that are kept as cluster seeds at prescreening. The y-axis is the fraction (F) of DBT volumes that have at least one cluster seed within the bounding box of the true cluster at the rank threshold R, defined as

$$F(R) = \frac{\sum_{\text{all volumes}} V_n}{V}, \tag{12}$$

where $V_n = 1$ if the true cluster in a given DBT volume contains at least one cluster seed object, and $V_n = 0$ otherwise. $\sum_{\text{all volumes}} V_n$ is the number of volumes that the true cluster is included in the R highest rank objects and V is the total number of volumes containing true clusters in the data set. Therefore, the rank-sensitivity plot at a given R is the percentage of true clusters that contain at least one cluster seed when the top R objects are kept in the DBT volume at the pre-screening stage. The rank-sensitivity plot is useful for optimization of the performance of the CADe system at the prescreening stage. It is important to maximize the sensitivity at this stage because missed clusters will never be recovered at later stages.

FROC. The performance of the CADe system was assessed using FROC analysis. The highest CNR value among the signals within each cluster, i.e., the CNR score, is used as the decision variable for detection. The FROC curve is generated by varying the CNR threshold over a range from the highest to the lowest sensitivity and counting the clusters that exceeds each threshold. A cluster is classified as TP if the location of the cluster centroid lies within the 3D bounding box marked by the radiologist and FP otherwise. FROC curves for the test set are reported in two ways. The view-based analysis treats the same cluster imaged in the two views as independent targets while case-based analysis treats the same cluster in the two views as the same target such that detection of the cluster in either one or both views is considered to be a TP detection.

JAFROC. Jackknife alternative FROC (JAFROC) analysis[31] is used to measure the performance difference between the previous study and the current method as two different treatments within a multicase scenario. Each view has only one biopsy-proven TP cluster, in case of multiple computer-detected clusters with centroid lying within the radiologists' bounding box, only one TP is counted for FROC

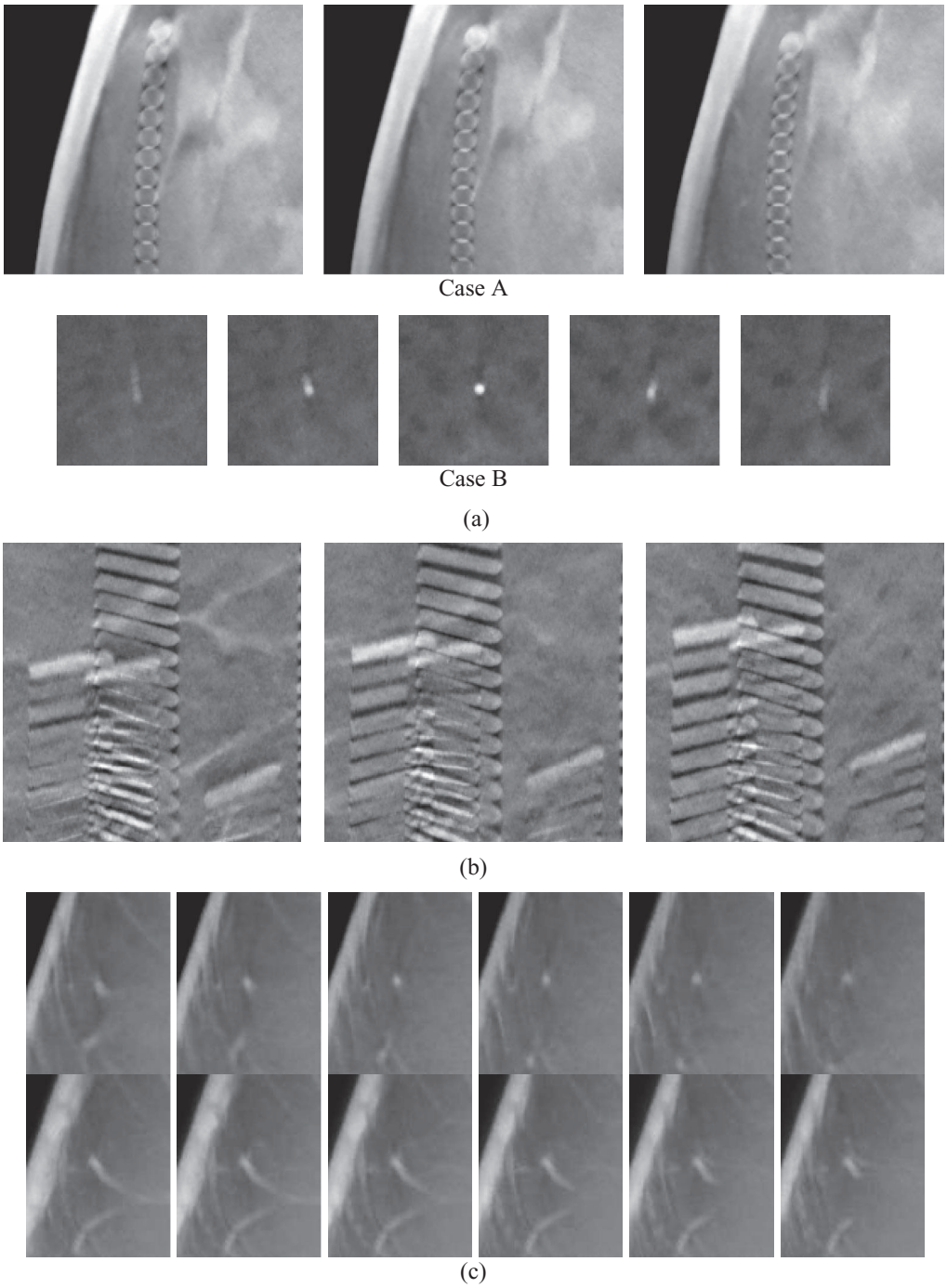


FIG. 6. (a) Examples of FPs from in-plane artifacts of metallic marker (Case A, slice numbers 43–45) and benign calcification (Case B, slice numbers 4–8). The FPs were eliminated by the cluster shape criterion. (b) Example of FPs from in-plane artifacts from surgical clips (slice number 39, 41, and 43). The resulting FP cluster was eliminated using the decision rules in Table I. (c) Example of FPs from branching tissue structures (slice number 2934) upper row: DBT slices, lower row: maximum intensity projection of the corresponding slices using a running 11-slice window. The resulting FP cluster in (c) was eliminated by the MIP analysis.

TABLE I. Decision rules based on features extracted from candidate signals in a cluster.

	Cluster containing microcalcifications with characteristics of	Rule for each object in a cluster	Number of objects satisfying the rule within a cluster	Criterion
1	Low CNR but larger in number	$\geq SZ_1$ and $\geq T_1$	c_1	if $(c_1 \geq 4)$
2	Medium and high CNR	$\geq SZ_2$ and $\geq T_2$	c_2	if $(c_2 \geq 3)$
3	Mixed low CNR and high CNR but few in number	$\geq SZ_3$ and $\geq T_3$	c_3	if $(c_3 = 1)$ and $(c_1 \geq 2)$

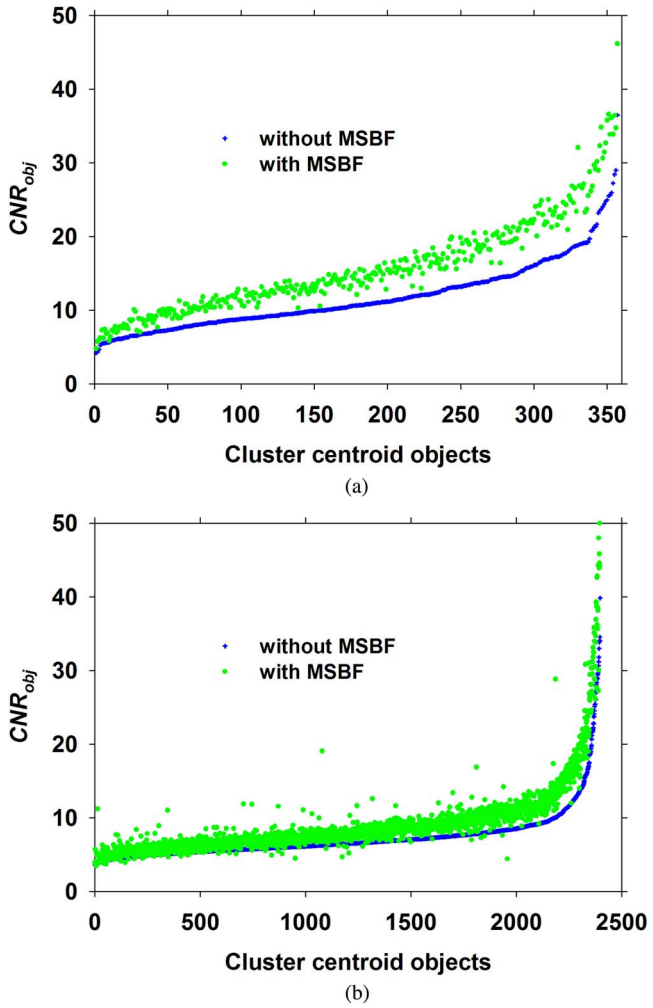


FIG. 7. Illustration of CNR_{obj} in DBT with and without MSBF regularization. (a) The subset of N cluster centroid objects that overlapped with the manually marked true microcalcifications, (b) all other N cluster centroid objects excluding those overlapping with the true microcalcifications in (a) for a subset of 20 cases (40 DBT volumes) from the training set. The CNR_{obj} for the microcalcifications (or other objects) in the DBT volumes reconstructed without MSBF were sorted by magnitude and the CNR_{obj} of the corresponding microcalcifications (or other objects) in the DBT volumes reconstructed with MSBF were plotted in the same order.

and JAFROC analyses. JAFROC analysis is applied to the view-based FROC curves, which provides the figure-of-merit (FOM) of each curve and the significance of the differences between the two curves being compared.

3. RESULTS

3.A. Effect of multiscale bilateral regularization

To investigate if there is a gain in the CNR of microcalcifications relative to that of FPs in the DBT volumes by regularized SART, the CNR_{obj} values of the top N cluster centroid objects that overlapped with the manually marked true microcalcification locations in 20 training cases (40 DBT volumes) when the MSBF-SART reconstructed volumes were used as input were compared to those when the conventional SART reconstructed volumes were used as input, as shown in Fig. 7(a). The same comparison for the top N objects detected at prescreening excluding the manually marked microcalcification locations, which would be predominantly FPs due to noise and other structures, was shown in Fig. 7(b). It can be seen that both the true microcalcifications and the other objects were enhanced by the MSBF regularized reconstruction. However, the improvement in the CNR_{obj} for the true signals was, on average, higher than that for the other objects. We calculated the ratio of the CNR_{obj} values with and without MSBF regularization for each individual calcification [objects in Fig. 7(a)] and FPs [objects in Fig. 7(b)] and compared the distributions of the two groups of CNR ratios by unpaired two-tailed t-test. The analysis shows that the CNR ratios of the true calcifications were significantly higher ($p < 0.0001$) than those of the FPs. The results indicate that MSBF can differentially enhance the true microcalcifications more than the FPs, thereby improving the classification between true and false positives.

3.B. Effect of EMCR formulation

The parameters for Hessian multiscale response filter and CNR enhancement were optimized during our previous work.[24] Based on the training data, the Gaussian scales for MCR volume generation were fixed at 0.25, 0.30, and 0.40 mm and the kernel sizes (M_1, M_2, M_3) for CNR enhancement were fixed at (7, 3, 1). For EMCR volume generation, the weights were chosen to be $w_1 = 0.4$ and $w_2 = 0.6$, for the weighted $EMCR_2$ volume in Eq. (8). The rank-sensitivity plots are compared in Fig. 8 for the two alternative EMCRs defined in Eqs. (7) and (8). A trend can be observed that more true clusters would contain a larger number of cluster seed objects in some range of R if $EMCR_2$ was used instead of $EMCR_1$. Alternatively, for some range of sensitivities, the $EMCR_2$ curve required a smaller R than the $EMCR_1$ curve, indicating that the cluster seeds in the biopsy-proven clusters

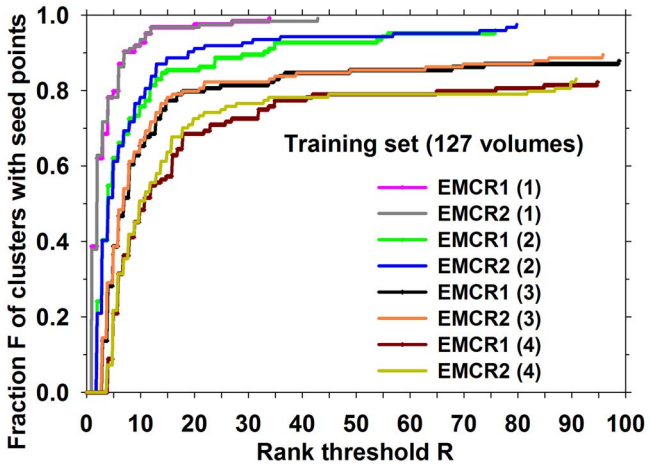


FIG. 8. Rank-sensitivity plots comparing two EMCRs for detection of cluster seeds at the prescreening stage. The two EMCRs were compared for the fractions of true clusters containing at least 1, 2, 3, and 4 cluster seeds, denoted as 1, 2, 3, 4 in the legend, respectively, when the top R objects are kept. The fraction of clusters was relative to the total of 124 clusters in the 127 training volumes.

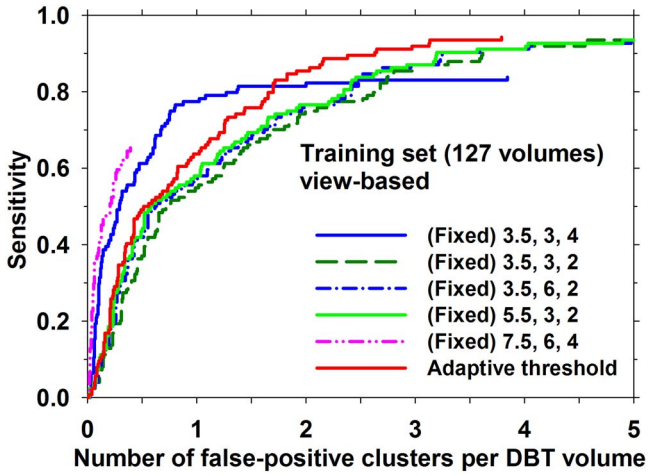


FIG. 9. Comparison of FROC curves for fixed and adaptive CNR thresholds. For the fixed threshold curves, the values in the legend indicate the CNR threshold, the size threshold, and the number of signals in a cluster used for the detection. Refer to Sec. 2.E.2 for size and signal count parameters related to the adaptive threshold approach. The sensitivity was relative to the total of 124 clusters in the 127 training volumes.

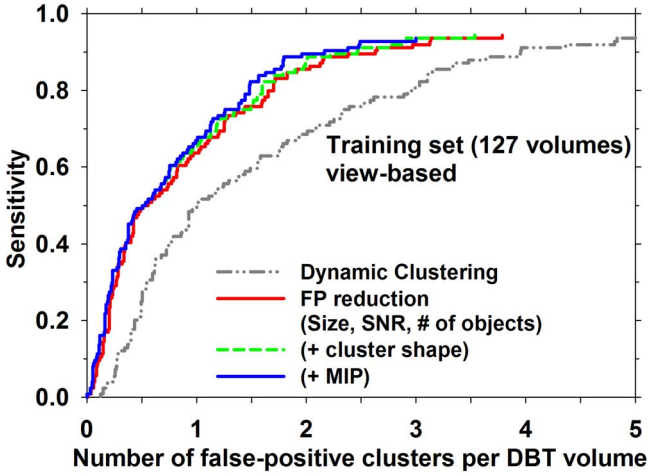


FIG. 10. FROC curves for the training set at different stages of false positive reduction. In the legend, "+ cluster shape" indicates the addition of the cluster shape criterion to the three criteria: size, CNR, and the number of objects in the cluster, "+ MIP" indicates the addition of the MIP criterion to the four preceding criteria. The sensitivity was relative to the total of 124 clusters in the 127 training volumes.

were ranked higher when the weighted $EMCR_2$ volume was used. $EMCR_2$ was, therefore, chosen for the current CADe system.

3.C. Fixed CNR threshold vs adaptive CNR threshold

In our previous work, a fixed CNR threshold was used across all volumes for formation of the clusters, contrary to the adaptive CNR threshold approach described in Sec. 2.D.3. Figure 9 compares the FROC curves for the fixed and adaptive CNR threshold methods for the training set. The detection results were analyzed after the first FP reduction step using size, CNR, and number of objects within a cluster.

3.D. Effect of false positive reduction

FROC curves are compared for the training set at different FP reduction stages, as summarized in Fig. 10 and Table II. The FP rate for 85% sensitivity was initially at 3.15 FPs per DBT volume and reduced to 1.70 FPs per DBT volume after all the FP reduction steps.

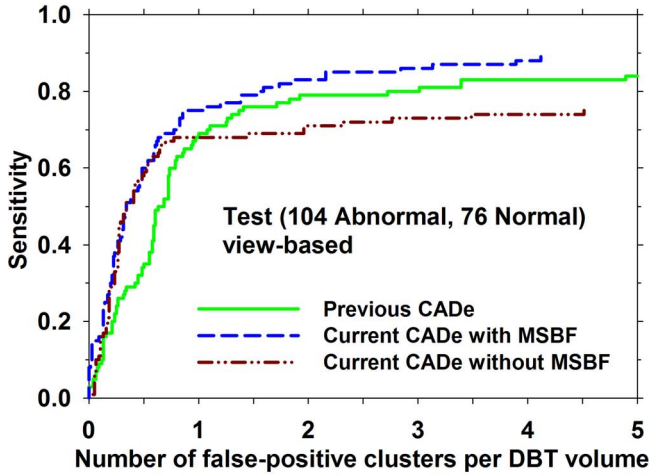
3.E. Validation

The CADe system with the trained parameters and decision rules was applied to the test set. The view-based and case-based test FROC curves, respectively, of the current system with and without the MSBF regularization are compared to the corresponding curve obtained from the CADe system in our previous study[24] in Figs. 11(a) and 11(b). The sensitivity was assessed from 52 cases (104 views) with MCs and the average number of FPs per DBT volume was estimated from 38 cases (76 views) free of MCs. Comparing the previous CADe system without MSBF and the current CADe system with MSBF, at 85% view-based sensitivity, the FP rate reduced from 5.42 to 2.16 FPs per DBT volume; at 85% case-based sensitivity, the FP rate decreased from 2.72 to 0.85 FPs per DBT volume. It may be noted that the performance of 85% view-based sensitivity at 5.42 FPs per DBT volume from the previous version of our CADe system was obtained from the subset of cases matched to the current test set. The previously reported results[24] were obtained from 72 breasts (144 views) with MCs and the 38 breasts (76 views) free of MCs, of which 20 breasts with MCs were included in our current training set so that they were excluded from the current test set. JAFROC analysis of the view-based FROC curves shows a statistically significant improvement in the FOM with a p-value < 0.05 as shown in Table III.

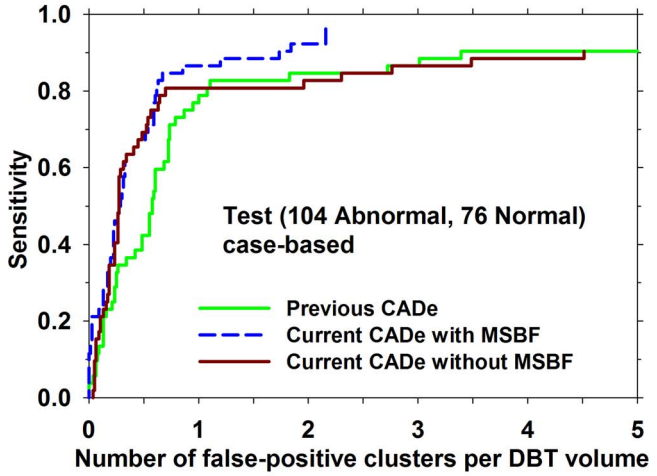
To evaluate the improvement in the detectability of the MCs due to MSBF regularization, we applied the current CADe system to the DBT reconstructed with SART without regularization. It can be seen from Fig. 11(a) that, for about 70% of the clusters that are detected at relatively high CNR thresholds (the region of the curve at low FP rates), the detection sensitivities in the DBTs with and without MSBF are comparable, and both are better than the curve from the previous CADe system. However, without MSBF enhancement, the very subtle clusters cannot be detected by the current

TABLE II. FROC estimated FPs at cluster formation and different FP reduction stages for the training set.

	FPs per DBT volume at 85% sensitivity
Initial clusters	3.15
FP reduction (size, CNR, number of objects in cluster)	1.91
FP reduction (cluster shape)	1.90
FP reduction (MIP)	1.71



(a)



(b)

FIG. 11. FROC curves for the test set of 180 DBT volumes from 52 breasts with MCs and 38 breasts free of MCs obtained from the current and previous methods. The sensitivity was estimated relative to the total of 100 clusters in the 104 DBT volumes of the 52 cases with MCs. The FP rates were estimated from the 76 DBT volumes free of MCs. (a) View-based FROC curves: the MC in each view was counted as an independent target. (b) Case-based FROC curves: the MC in a breast was considered as detected if it was detected in either one or both views. The FP rate per case can be derived by multiplying the FP rate per DBT volume with the number of DBT volumes per case.

CADe system. This gain in sensitivity by MSBF surpasses what could be achieved even when the CADe system was designed to trade-off specificity for sensitivity as was done in our previous CADe system.

TABLE III. Improvement in the FROC curves for the test set between the preliminary and the current study by JAFROC analysis. The difference in the Figure-of-Merit (FOM) between the two treatments is statistically significant with a p-value < 0.05.

Treatment	FOM	95% Confidence interval
Previous CADe	0.49	(0.42, 0.55)
Current CADe with MSBF	0.58	(0.52, 0.65)
p-value	0.003	

4. DISCUSSION

The SART with multiscale bilateral regularization enhances the high frequency calcification signals (Fig. 3) and smoothens the noise without blurring the mass margins and spiculations.[25,27] The microcalcifications in the DBT volumes by regularized SART have higher CNR than those in the conventional SART. The increased CNR is important for the separation of TP (true-positive) and FP microcalcifications, which contributed substantially to the improvement in the detection accuracy of our current CADe system.

The regularized DBT volume was further transformed to generate the MCR and CNR enhanced volumes. The response function of the Hessian analysis was optimized to obtain high MCR values for spherical objects. The CNR enhanced volume highlighted the high contrast structures while suppressing the background. Combination of the MCR and the CNR enhanced volumes generated an EMCR volume which takes advantage of the enhanced structure and contrast information. The generation of EMCR can be considered to be an FP suppression process at the signal level. As seen in Fig. 8, the new EMCR improved the ranking of the true cluster seeds within the set of potential cluster seeds. The higher ranking or larger number of seeds in the cluster provides a greater chance that a cluster would be formed in the clustering stage. $EMCR_2$ is, therefore, chosen in the current CADe system.

The multiscale bilateral regularization increases the CNR of the true microcalcifications relative to the FP objects, which will lead to higher EMCR ranking, on average, of the true microcalcifications in the set of top N_s candidates. The increase in the ranking of the true signals will increase the chance that the individual microcalcifications will be captured as cluster members earlier in the clustering process to form a true cluster.

The overall impact of MSBF regularization on the detectability of subtle MCs is well demonstrated in the FROC curves with and without MSBF compared in Fig. 11. Without MSBF enhancement, the CADe system cannot differentiate the subtle microcalcifications and noise effectively. Our previous CADe system had to make compromise in order to detect the subtle clusters but also retain a larger number of FPs at all sensitivities. With MSBF and redesigning the CADe system to exploit the increased separation between the microcalcifications and FPs, the FP rate is reduced at a given sensitivity, or alternatively, the sensitivity is increased at a given FP rate, compared to that achieved with the previous CADe system. The new adaptive strategies can accommodate microcalcifications that have reasonably good CNR with and without MSBF, as observed from the consistent performance in the region of the curves at low FP rates. For the very subtle MCs that are close to noise, they will be lost without regularization if the FP rate has to be kept low.

The variation of the CNR_{obj} of the detected objects at prescreening depends on the complexity of the breast parenchyma; dense breasts have larger variations than fatty breasts due to the fibroglandular structures. The regularized reconstruction enhances the conspicuity of TP signals as well as fibrous structures, although the enhancement of true signals

is greater than that of FPs [Figs. 7(a) and 7(b)]. The use of a low fixed CNR threshold across volumes to increase sensitivity would result in a high FP rate, while a high fixed threshold to reduce the FP rate might lead to low sensitivity. In order to adapt to the CNR variations across breast volumes, we designed an adaptive CNR threshold method in which the CNR threshold is calculated for each volume based on the distribution of the CNR_{obj} of the segmented N_s objects. Figure 9 shows the effect of the fixed CNR threshold, the microcalcification size, and the number of microcalcifications in the cluster on the FROC curves. With an increase in the threshold for the number of microcalcifications, the sensitivity increased steeply at the low FP range but it leveled off and could not reach as high sensitivity as the other conditions. The adaptive CNR threshold method could reach a higher sensitivity in the mid FP range and it is more robust against changes in the CNR distribution from different DBT volumes. The adaptive CNR threshold method was, therefore, implemented in our current CAD system.

Based on the analysis of the TP and FP clusters on the training set, several major types of detected objects were identified, (a) true microcalcifications, (b) arterial calcifications, (c) bending or branching fibrous tissue, ducts, or vascular structures, (d) inter- or in-plane artifacts due to metal clips or dense benign calcifications, and (e) random noise. Some of the true calcifications that were considered clinically insignificant and were not marked by the radiologist were counted as FP signals. The decision rules based on size, CNR, and number of signals in a cluster (Table I) effectively reduced the FPs either at the individual object or cluster levels. Figure 10 indicates that the FP rate was reduced from 3.15 to 1.91 FPs per DBT volume at 85% sensitivity. The shape of the cluster is related to the distribution of the objects within the cluster. Clusters that are too elongated can be an indication of arterial calcifications, the shadows of metal clips, or edges of fibrous tissue or other structures. A rectangularity criterion was imposed on the shape of the bounding box to reduce some of the FPs. Vessels and fibrous structures can also extend across slices and be detected as separated objects due to noise in the images. Because of the blurring in the depth direction in DBT, the vessels and fibrous tissues cannot be traced as a continuous structure in 3D. Instead, we used the maximum intensity projection of the cluster to the plane parallel to the DBT slices to obtain a planar image. The connectivity of the detected objects and their shape are assessed in the CNR enhanced domain to eliminate extended connected structures, reducing the FP rate to 1.71 per DBT volume for the training set. The effectiveness of these criteria depends on the parenchymal structures of the breast.

The current CAD system showed a significant improvement in the figure-of-merit of the FROC curve from 0.49 to 0.58 compared to our previous CAD system (Table III). The current system is different from the previous system in several ways. The microcalcifications are first enhanced by the MSBF regularized SART. The EMCR formulation is improved. The fixed CNR threshold method is replaced by an adaptive CNR threshold method. The numbers of cluster seeds and individual seeds are adaptively determined based on the reconstructed breast volume. The decision rules that combine the CNR threshold, the signal size threshold, and the number of signals in a cluster for FP reduction are redesigned. A maximum intensity projection method is added to identify FPs due to extended structures detected as points in 3D. The collective effects of these improvements significantly increase the sensitivity and specificity of the CAD system. The reduction in FPs will facilitate correspondence analysis between clusters found in CC and MLO views. We will investigate methods to register the DBT volumes in CC and MLO views and combine the detected cluster information from the two views to further reduce FPs.

The current CADe system is designed to be adaptive to the CNRs of microcalcifications in the input DBT volume. The difference in the CNRs may be caused by many factors such as differences in the reconstruction techniques, regularization, image acquisition parameters, and the detector and system design characteristics of the DBT systems, in addition to the natural variations in the properties of the lesions in the patient population. Without testing the CADe system with the wide varieties of DBTs, it is not known whether the system is robust against these factors. However, some hints may be observed from the analysis of the CNR_{obj} in our dataset, which contains microcalcifications with CNR_{obj} over a range of 5–50 [Fig. 7(a)] and mean CNR_{obj} over a range of about 5–25 [Fig. 5(b)]. The comparison of the FROC curves in Fig. 11 for DBT with and without MSBF shows that the performance of the current CADe system is comparable for about 70% of the MCs with and without MSBF. This indicates that the CADe system is adaptive to a relatively wide range of CNRs. However, for DBTs in which the CNRs of the microcalcifications are low relative to the FPs due to any factors mentioned above, the sensitivity of the CADe system will be low without trading off the specificity, which is expected for any CAD systems or human readers. It is, therefore, important to acquire and reconstruct DBT with good image quality for detection of subtle microcalcifications. This study demonstrates that MSBF regularization is an effective approach to enhancing the CNR of subtle microcalcifications, which is expected to increase their chance of being detected by this or other CADe systems. Evaluation of the performance of our CADe system for DBTs of different image quality is out of the scope of the current study.

Detection of breast cancer at early stage can reduce the mortality rate and DBT is poised to surpass DM in sensitivity and specificity.[2–11,19] However, the time and effort required to search for microcalcifications in the large DBT volume and the inevitable oversight under the workflow demand is a major concern of replacing DM with DBT. Slab view and synthesis of a 2D image from the DBT volume to create a DM-like image are potential approaches that can facilitate the visual search of microcalcifications by radiologists. A CADe system that can efficiently search for subtle lesions in the DBT volume will further assist radiologists and utilize the 3D information. It will be of interest to develop CADe methods for detection of microcalcifications in the synthesized DM when synthesized DMs become commonly available, to compare the detectability of microcalcifications between the

reconstructed DBT volumes, the synthesized DMs, and the original projection views without reconstruction, and to investigate if combining the information from any of these approaches may improve detection accuracy. Ultimately, the utility of CADe for assisting radiologists in DBT interpretation will need to be investigated.

5. CONCLUSION

This study shows that the detection performance of our CADe system for clustered MCs in DBT has significantly improved over our previous work. The multiscale bilateral regularized SART reconstruction of DBT increased the CNR of the microcalcifications, which, in combination with the improved design of the computer vision techniques at the various stages of the CADe system, especially the adaptive CNR thresholds and FP reduction methods, improved the differentiation of true and false positives. Further investigations will be needed to reduce FPs caused by the in-plane and interplane artifacts from high density objects such as metal clips and large benign calcifications and to improve the sensitivity.

ACKNOWLEDGMENTS

This work is supported by USPHS Grant No. RO1 CA151443. The digital breast tomosynthesis system was developed by the GE Global Research Group, with input and some revisions from the University of Michigan investigators, through the Biomedical Research Partnership (USPHS Grant No. CA91713, PI: Paul Carson, Ph.D.) collaboration. The content of this paper does not necessarily reflect the position of the funding agencies and no official endorsement of any equipment and product of any companies mentioned should be inferred.

[a]Author to whom correspondence should be addressed. Electronic mail: rsamala@umich.edu; Telephone: (734) 647-8556; Fax: (734) 615-5513.

[1]J. G. Elmore, S. L. Jackson, L. Abraham, D. L. Miglioretti, P. A. Carney, B. M. Geller, B. C. Yankaskas, K. Kerlikowske, T. Onega, R. D. Rosenberg, E. A. Sickles, and D. S. M. Buist, "Variability in interpretive performance at screening mammography and radiologists' characteristics associated with accuracy," Radiology **253**, 641–651 (2009).

[2]S. P. Poplack, T. D. Tosteson, C. A. Kogel, and H. M. Nagy, "Digital breast tomosynthesis: Initial experience in 98 women with abnormal digital screening mammography," Am. J. Roentgenol. **189**, 616–623 (2007).

[3]I. Andersson, D. M. Ikeda, S. Zackrisson, M. Ruschin, T. Svahn, P. Timberg, and A. Tingberg, "Breast tomosynthesis and digital mammography: A comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings," Eur. Radiol. **18**, 2817–2825 (2008).

[4]P. Skaane, A. I. Bandos, R. Gullien, E. B. Eben, U. Ekseth, U. Haakenaasen, M. Izadi, I. N. Jebsen, G. Jahr, and M. Krager, "Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program," Radiology **267**, 47–56 (2013).

[5]T. M. Svahn, D. P. Chakraborty, D. Ikeda, S. Zackrisson, Y. Do, S. Mattsson, and I. Andersson, "Breast tomosynthesis and digital mammography: A comparison of diagnostic accuracy," Br. J. Radiol. **85**, e1074–e1082 (2012).

[6]D. Gur, A. I. Bandos, H. E. Rockette, M. L. Zuley, J. H. Sumkin, D. M. Chough, and C. M. Hakim, "Localized detection and classification of abnormalities on FFDM and tomosynthesis examinations rated under an FROC paradigm," Am. J. Roentgenol. **196**, 737–741 (2011).

[7]G. Gennaro, R. E. Hendrick, P. Ruppel, R. Chersevani, C. di Maggio, M. La Grassa, L. Pescarini, I. Polico, A. Proietti, and E. Baldan, "Performance comparison of single-view digital breast tomosynthesis plus single-view digital mammography with two-view digital mammography," Eur. Radiol. **23**, 664–672 (2013).

[8]E. A. Rafferty, J. M. Park, L. E. Philpotts, S. P. Poplack, J. H. Sumkin, E. F. Halpern, and L. T. Niklason, "Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: Results of a multicenter, multireader trial," Radiology **266**, 104–113 (2013).

[9]M. G. Wallis, E. Moa, F. Zanca, K. Leifland, and M. Danielsson, "Two-view and single-view tomosynthesis versus full-field digital mammography: High-resolution x-ray imaging observer study," Radiology **262**, 788–796 (2012).

[10]P. Skaane, R. Gullien, H. Bjørndal, E. B. Eben, U. Ekseth, U. Haakenaasen, G. Jahr, I. N. Jebsen, and M. Krager, "Digital breast tomosynthesis (DBT): Initial experience in a clinical setting," Acta Radiol. **53**, 524–529 (2012).

[11]P. Timberg, M. Båth, I. Andersson, S. Mattsson, A. Tingberg, and M. Ruschin, "Visibility of microcalcification clusters and masses in breast tomosynthesis image volumes and digital mammography: A 4AFC human observer study," Med. Phys. **39**, 2431–2437 (2012).

[12]X. Qian, A. Tucker, E. Gidcumb, J. Shan, G. Yang, X. Calderon-Colon, S. Sultana, J. Lu, O. Zhou, D. Spronk, F. Sprenger, Y. Zhang, D. Kennedy, T. Farbizio, and Z. Jing, "High resolution stationary digital breast tomosynthesis using distributed carbon nanotube x-ray source array," Med. Phys. **39**, 2090–2099 (2012).

[13]Y. Lu, H.-P. Chan, J. Wei, M. M. Goodsitt, P. L. Carson, L. Hadjiiski, A. Schmitz, J. W. Eberhard, and B. E. H. Claus, "Image quality of microcalcifications in digital breast tomosynthesis: Effects of projection-view distributions," Med. Phys. **38**, 5703–5712 (2011).

[14]H.-P. Chan, M. M. Goodsitt, M. A. Helvie, C. Paramagul, C. H. Neal, P. L. Carson, M. A. Roubidoux, M. Noroozian, and A. V. Nees, "Digital breast tomosynthesis (DBT): Observer performance study of microcalcification cluster detection in breast phantom DBT acquired with variable tomographic angles, angular increments, and number of projections," RSNA Program Book VSBR41-08 (2012).

[15]Y. H. Hu, B. Zhao, and W. Zhao, "Image artifacts in digital breast tomosynthesis: Investigation of the effects of system geometry and reconstruction parameters using a linear system approach," Med. Phys. **35**, 5242–5252 (2008).

[16]I. Sechopoulos and C. Ghetti, "Optimization of the acquisition geometry in digital tomosynthesis of the breast," Med. Phys. **36**, 1199–1207 (2009).

[17]I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," Med. Phys. **37**, 1591–1600 (2010).

[18]M. L. Spangler, M. L. Zuley, J. H. Sumkin, G. Abrams, M. A. Ganott, C. Hakim, R. Perrin, D. M. Chough, R. Shah, and D. Gur, "Detection and classification of calcifications on digital breast tomosynthesis and 2D digital mammography: A comparison," Am. J. Roentgenol. **196**, 320–324 (2011).

[19]D. Kopans, S. Gavenonis, E. Halpern, and R. Moore, "Calcifications in the breast and digital breast tomosynthesis," Breast J. **17**, 638–644 (2011).

[20]G. Peters, S. Muller, S. Bernard, R. Iordache, F. Wheeler, and I. Bloch, "Reconstruction-independent 3D CAD for calcification detection in digital breast tomosynthesis using fuzzy particles," Prog. Pattern Recogn., Image Anal. Appl. **3773**, 400–408 (2005).

[21]S. C. Park, B. Zheng, X. H. Wang, and D. Gur, "Applying a 2D based CAD scheme for detecting micro-calcification clusters using digital breast tomosynthesis images: An assessment," Proc. SPIE **6915**, 6915071–6915078 (2008).

[22]I. Reiser, R. M. Nishikawa, A. V. Edwards, D. B. Kopans, R. A. Schmidt, J. Papaioannou, and R. H. Moore, "Automated detection of microcalcification clusters for digital breast tomosynthesis using projection data only: A preliminary study," Med. Phys. **35**, 1486–1493 (2008).

[23]G. van Schie and N. Karssemeijer, "Noise model for microcalcification detection in reconstructed tomosynthesis slices," Proc. SPIE **7260**, 72600M-1–72600M-8 (2009).

[24]B. Sahiner, H.-P. Chan, L. M. Hadjiiski, M. A. Helvie, J. Wei, C. Zhou, and Y. Lu, "Computer-aided detection of clustered microcalcifications in digital breast tomosynthesis: A 3D approach," Med. Phys. **39**, 28–39 (2012).

[25]Y. Lu, H.-P. Chan, J. Wei, R. K. Samala, L. Hadjiiski, and P. L. Carson, "Multiscale bilateral regularization in digital breast tomosynthesis (DBT)," *RSNA Program Book SSE* (2012), Vol. 22.

[26]Y. Lu, H.-P. Chan, J. Wei, and L. M. Hadjiiski, "Selective-diffusion regularization for enhancement of microcalcifications in digital breast tomosynthesis reconstruction," Med. Phys. **37**, 6003–6014 (2010).
[27]Y. Lu, H.-P. Chan, J. Wei, L. Hadjiiski, and C. Zhou, "Multiscale regularized reconstruction for enhancing microcalcification in digital breast tomosynthesis," Proc. SPIE **8313**, 831322-1–831322-9 (2012).
[28]P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. Commun. **COM-31**, 337–345 (1983).
[29]C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the Sixth International Conference on Computer Vision* (IEEE, 1998), pp. 839–846.
[30]A. Huang, H.-M. Liu, C.-W. Lee, C.-Y. Yang, and Y.-M. Tsang, "On concise 3-D simple point characterizations: A marching cubes paradigm,"IEEE Trans. Med. Imaging **28**, 43–51 (2009).
[31]D. P. Chakraborty, "Validation and statistical power comparison of methods for analyzing free-response observer performance studies," Acad. Radiol. **15**, 1554–1566 (2008).