Noise Equalization for Detection of Microcalcification Clusters in Direct Digital Mammogram Images

Kristin J. McLoughlin, Philip J. Bones*, *Senior Member, IEEE*, and Nico Karssemeijer, *Member, IEEE*

Abstract—Equalizing image noise is shown to be an important step in the automatic detection of microcalcifications in digital mammography. This study extends a well established film-screen noise equalization scheme developed by Veldkamp *et al.* for application to full-field digital mammogram (FFDM) images. A simple noise model is determined based on the assumption that quantum noise is dominant in direct digital X-ray imaging. Estimation of the noise as a function of the gray level is improved by calculating the noise statistics using a truncated distribution method. Experimental support for the quantum noise assumption is presented for a set of step wedge phantom images. Performance of the noise equalization technique is also tested as a preprocessing stage to a microcalcification detection scheme. It is shown that the square root model based approach which FFDM allows leads to a robust estimation of the high frequency image noise. This provides better microcalcification detection performance when compared to the film-screen noise equalization method developed by Veldkamp. Substantially better results are obtained than when noise equalization is omitted. A database of 124 direct digital mammogram images containing 28 microcalcification clusters was used for evaluation of the method.

Index Terms—Cancer, computer-aided diagnosis, noise measurement, X-ray imaging.

I. INTRODUCTION

IN many western countries breast cancer is the most common form of cancer among women. The World Health Organization estimates that more than 385 000 women worldwide die of the disease each year. Clear evidence shows that early diagnosis and treatment of breast cancer can significantly increase the chance of survival for patients [1]. One of the most important radiological signs for early detection of breast cancer is the presence and appearance of microcalcifications. These are small regions of elevated intensity against the varying background density of the X-ray mammogram, which often appear in clusters [2]. Much attention has been directed toward developing an automated detection scheme for the detection of microcalcification clusters. It has been shown in digitized film-screen mammograms that pixels sized 100 μm or less give sufficient resolution to represent the individual shapes of microcalcifications [3]. However, the presence of intensity dependent noise is a significant factor in microcalcification detection [4]. In this paper, we address the problem of how to reduce the effect of intensity dependent noise in direct digital mammograms comparing our findings with a method previously developed for digitized film screen mammograms. By incorporating a new model for the intensity dependence of the noise, the new method is shown to be more robust than the previous method when applied to digital mammograms. We also demonstrate that equalizing the noise in direct digital mammograms has a significant influence on the success of microcalcification detection.

Most microcalcification detection methods described in the literature were developed for digitized film-screen mammograms. Many of these methods use some form of noise-dependent adaptive threshold that is locally determined. Nishikawa *et al.* used a global gray level threshold as an initial processing stage followed by a local adaptive thresholding technique [5]. Chan *et al.* used a local gray level threshold, which varies with the standard deviation of the surrounding pixel values [6]. Neiber *et al.* used a local threshold, dependent on the difference between local maximum and mean gray levels [7]. Adaptive thresholds that are determined locally in small image regions are influenced by image structures like lines and edges as well as the image noise. Therefore, in regions with a lot of image structure, thresholds are unlikely to be adjusted optimally for the noise level and the detection performance may deteriorate. Other approaches have been based on adaptive noise suppression and hybrid wavelet adaptive enhancement methods [8]–[10]. A disadvantage of these latter methods is that they suppress the noise (or enhance the signal) locally without taking the signal dependency of the noise into account.

Karssemeijer developed a model in which images are rescaled to equalize the noise [4], a method that was later improved upon by Veldkamp *et al.* [11]. This method transforms the input image in which noise depends strongly on the intensity into an image with a homogeneous noise level. A number of studies have successfully applied this approach in microcalcification detection [12]–[14]. Netsch *et al.* used a method based on Laplacian scale-space representation of the mammogram and applied noise equalization [12]. Maitournam *et al.* used splines to model the trend in each mammogram, which was then subtracted from the original image. The resultant noisy image was thresholded after variance equalization [13]. Poissonnier and Brady compared two noise compensation techniques; the first was a transformation providing a linear relationship between grayscale and film density, while the second was noise equalization. They

concluded that noise equalization is a superior approach as it is adaptive to varying imaging conditions [14]. However, the method of Veldkamp *et al.* has difficulty estimating the noise in the brightest (most dense) areas of a digitized film-screen mammogram, because the pixels that contribute may become very few in number. The new model-based approach described below is simpler and avoids these problems.

To date, film-screen mammography has been the most common and effective technique for the detection of breast cancer. However, the film-screen system has several technical limitations which can reduce the accuracy of breast cancer diagnosis [15]. It is believed that full-field digital mammography (FFDM) has the potential to increase the sensitivity of mammography screening and emerge as a film-screen mammography replacement [16]. Both modalities suffer from the presence of image noise that is intensity dependent. In digitized film-screen mammograms, the relationship between intensity and noise variance is nonlinear and varies significantly from image to image. We show here, however, that for FFDM a theoretically predicted model for the relationship can be employed to good effect. Initially we investigate the validity of the theoretical model using step wedge phantom images obtained with a General Electric (GE) Senographe system. An equalization method based on the model is then tested on a dataset of 124 FFDM images, as a preprocessing stage to a microcalcification detection scheme developed by Karssemeijer [17]. Three different preprocessing techniques are compared based on the performance of the microcalcification detection scheme. The first uses the new model-based method, the second uses the method of Veldkamp *et al.*, and the third omits noise equalization altogether. Free-response operating characteristic (FROC) curves are used for evaluation of the detection performance.

II. FFDM NOISE

All radiological images contain random fluctuations due to the statistics of X-ray quantum absorption. This noise can limit the reliability of detection of small or subtle structures. In digitized film-screen images, other sources, due to the digitizer, the structure of the fluorescent screen and the granularity of the film emulsion, contribute to the noise. An ideal imaging system would be "X-ray quantum limited," meaning that X-ray quantum noise is the dominant noise source present. Generally, existing mammographic film-screen systems are not quantum limited due to the above mentioned additional noise components [15].

In an FFDM system, a linear relationship exists between gray level and exposure as is shown in Fig. 1. To generate the plots a 5.0-cm-thick polymethyl methacrylate block was imaged with three combinations of target and filter material and a range of exposures. The mean pixel value was calculated for a 4 cm^2 region-of-interest in the center of the image. Quantum noise is due to the finite number of detected X-ray photons and the signal-to-noise ratio due to this source alone varies as the square root of the total number of photons detected [15]. Therefore, if quantum noise is assumed to be the dominant noise source present, a square root model will provide an accurate estimate
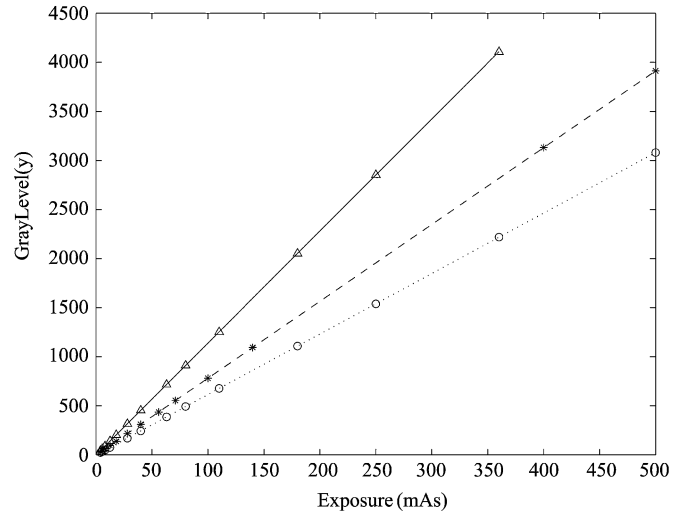

Fig. 1. Plots of pixel value (12-bit number) versus exposure in mAs for a 5.0-cm PMMA block at 28 kVp and three target/filter combinations: △: rhodium/rhodium; ○: molybdenum/molybdenum; and ∗: molybdenum/rhodium.

of the noise with respect to gray level. We exploit this model to perform noise equalization in a robust manner.

In this paper, the method described by Veldkamp *et al.* is extended for use on direct digital mammograms. The dependence of the noise on the gray level y is removed by rescaling the local contrast using the standard deviation of the local contrast $\sigma_c(y)$ as a measure for the noise. Local contrast c_i at site i is defined for this purpose as

$$c_i = y_i - \frac{1}{N} \sum_{j \in \delta_i} y_j \qquad (1)$$

where δ_i represents a neighborhood of pixels centered on i of size N. Equation (1) takes the form of a simple high-pass spatial filter, therefore the standard deviation of the local contrast $\sigma_c(y)$ provides a measure of the high frequency image noise as a function of the gray level. To obtain $\sigma_c(y)$ for digitized film-screen mammograms, the grayscale is divided into nonoverlapping but adjacent bins numbered $k = 1, 2, \ldots, K$. The kth bin has center gray value y_k. A histogram of local contrast values is formed for all pixels with gray levels within the kth bin and this is normalized to form an estimate of the probability density function (pdf) $f(c|k)$. The value of $\sigma_c(k)$ is then calculated from $f(c \mid k)$, assuming $f(c \mid k)$ to be Gaussian. This is repeated for all of the grayscale bins. Interpolation of the estimated $\sigma_c(k)$ values provides an estimate of the continuous function $\sigma_c(y)$. The mean of the local contrast with respect to gray level $\mu_c(y)$ can also be obtained in a similar fashion. Taking the signal dependency of both $\sigma_c(y)$ and $\mu_c(y)$ into consideration, the normalized local contrast at site i is calculated using

$$c_i' = \frac{(c_i - \mu_c(y_i))}{\sigma_c(y_i)}. \qquad (2)$$

This process removes the noise dependency on the gray level, hence the description "noise equalization." No additional information (obtained from phantom recordings, for example) is in-
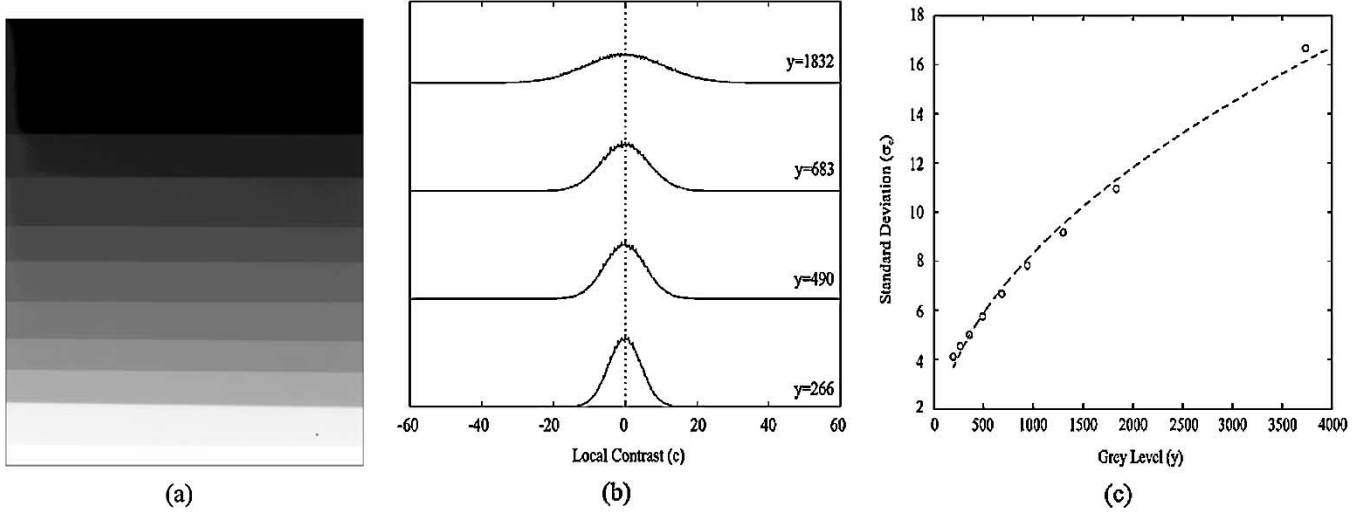
Fig. 2. (a) An example phantom image (27 kVp, 140 mAs, target material = molybdenum, filter material = rhodium). (b) Histograms of local contrast calculated from the phantom image for four different bins. The vertical axis for each histogram represents the relative frequency of occurrence. (c) Square root fit $\sigma_c^{\mathrm{mod}}(y)$ "-," to the $\sigma_c(k)$ data "o."

volved in the method, which makes the approach robust and independent of film type and film development characteristics.

To apply noise equalization to FFDM images, we incorporate the square root model into the method described above. In this case, to obtain the continuous function $\sigma_c(y)$ (measure of the high-frequency noise) the grayscale is again divided into nonoverlapping bins. After computing the local contrast using (1), the pdf $f(c \mid k)$ is estimated and $\sigma_c(k)$ calculated for the kth bin as before. However, a square root function fit through the set $\{\sigma_c(k)\}$ is used to estimate $\sigma_c(y)$. We define the local contrast deviation square root model

$$\sigma_c^{\mathrm{mod}}(y) = my^{1/2} + d \tag{3}$$

where m and d are constants determined by minimizing the mean squared error between $\{\sigma_c^{\mathrm{mod}}(y_k)\}$ and $\{\sigma_c(k)\}$. The main advantages of the model-based approach over Veldkamp's method are: 1) the estimation of the relationship between noise variance and intensity remains reliable in those ranges of intensity for which the population of pixels in a particular mammogram is relatively low; 2) the specific choice of bins is much less critical; and 3) less computational effort is required.

The following sections investigate the validity of the square root noise model using both a direct digital step wedge phantom and FFDM images. The use of the new method in microcalcification detection is also demonstrated.

III. PHANTOM IMAGES

In this section, the square root noise model is tested on a set of direct digital step wedge phantom images. Demonstrating that the noise model is accurate for images with no background structure provides initial validation of the assumed quantum noise dominance.

A. Acquisition

A single step wedge phantom was created by partially overlapping nine sheets of 5 mm and 10 mm thick perspex. Twelve direct digital images were obtained from this phantom

using the GE Senographe system at the Nijmegen Medical Center. The original 14 bits per pixel GE images were linearly rescaled to 12 bits. Each image was 2294×1914 pixels in size and had a 100-μm resolution. The acquisition parameters for each phantom image were different, with the range chosen to coincide with typical settings for direct digital mammograms (see Section III-C). An example phantom image is shown in Fig. 2(a).

B. Noise Model Application

Implementation of the square root noise model requires c to be binned with respect to gray level y. The obvious method of binning is to place all pixels lying inside one step of the wedge into a single bin, so the number of bins K corresponds to the number of steps in the phantom. This is justified by the fact that pixels associated with a given step are expected to possess similar gray levels. The local contrast pdf $f(c \mid k)$ can then be estimated. Fig. 2(b) shows four local contrast distributions (corresponding to four different bins) calculated from the phantom image in Fig. 2(a). The local contrast values were obtained using a 9×9 window for $\delta_i (N = 81)$. Also shown in this figure are the corresponding mean gray levels for each bin. Fig. 2(b) indicates that each $\mu_c(k)$ remains very close to zero as expected from symmetry considerations.

The standard deviation $\sigma_c(k)$ is estimated from $f(c \mid k)$, assuming $f(c \mid k)$ to be Gaussian. The best fit continuous function $\sigma_c^{\mathrm{mod}}(y)$ is then estimated using (3). Fig. 2(c) shows the estimates $\{\sigma_c(k)\}$ and the corresponding $\sigma_c^{\mathrm{mod}}(y)$ obtained for the phantom image in Fig. 2(a). It is clear from this figure that a square root function provides a good fit to the binned data. Note that y, c, d, $\sigma_c(k)$, and $\sigma_c^{\mathrm{mod}}(y)$ are all measured in terms of gray levels, i.e., 0 to 4095 for 12-bit data.

The phantom images used in the study were observed to have a small number of artifacts with very high magnitude contrast compared with their surrounding pixels. The artifacts were determined to be due to imperfections in the material used to make the phantom. The presence of these artifacts biased the estimation of $\sigma_c(k)$ significantly. To ameliorate this problem the
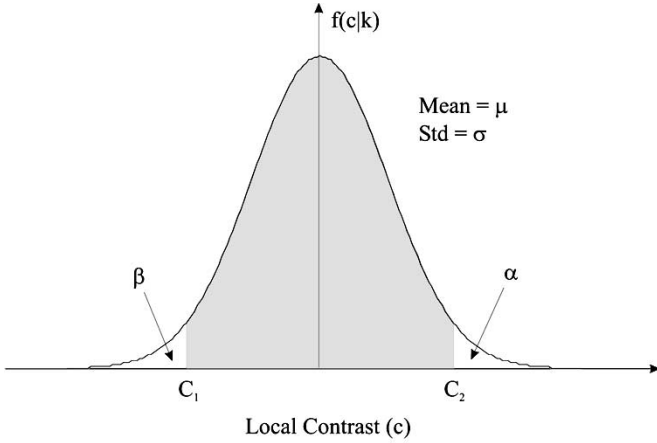
Fig. 3. Gaussian local contrast distribution ($\text{mean} = \mu$, $\text{standard deviation} = \sigma$) for a single bin with corresponding truncation points (C_1 and C_2) marked.

contrast distributions were truncated to exclude extreme values. The population parameter $\sigma_c(k)$ was then estimated based on the truncated distribution as follows.

Fig. 3 shows an original Gaussian contrast distribution $f(c\,|\,k)$ with mean μ and standard deviation σ. The truncation points labeled C_1 and C_2 are chosen to be symmetric around $c = 0$ and to satisfy

$$\int_{C_1}^{C_2} f(c\,|\,k)\,dc = L \tag{4}$$

where L is a predefined fraction of the area under $f(c\,|\,k)$ (the shaded area in Fig. 3). The parameter α represents the proportion of the population above C_2 while β represents the proportion of the population below C_1. Outliers caused by artifacts within the image will occur in the regions represented by α and β. It is assumed that these outliers occur relatively infrequently and, therefore, will not heavily influence the calculated proportions.

By utilizing the cumulative distribution function (Φ), and assuming the contrast distribution is Gaussian outside the truncation points, α and β can be written as

$$\alpha = 1 - \Phi\left(\frac{C_2 - \mu}{\sigma}\right) \tag{5}$$

$$\beta = \Phi\left(\frac{C_1 - \mu}{\sigma}\right). \tag{6}$$

These equations can be rearranged for μ and σ to give

$$\sigma = \frac{C_2 - C_1}{\Phi^{-1}(1 - \alpha) - \Phi^{-1}(\beta)} \tag{7}$$

and

$$\mu = C_1 - \Phi^{-1}(\beta)\sigma \tag{8}$$

which provides a means of calculating the statistics of the nontruncated local contrast distribution based on the known truncated data, provided the proportions α and β are known. The $\sigma_c(k)$ values shown in Fig. 2(c) were calculated using this truncated distribution method with parameter $L = 0.95$.

C. Results

The noise model technique described in Section III-B was applied to each of the 12 phantom images. Shown in Table I are the acquisition parameters for each image, which were chosen to coincide with typical settings for direct digital mammograms. In some phantom images, the intensity for one or more of the thinnest and/or thickest steps was clearly in saturation, so these were excluded from the calculations.

To provide a measure of how well the noise model fits the data the correlation coefficient r between $\sqrt{y_k}$ and $\sigma_c(k)$ was calculated for each phantom image, with a correlation coefficient of 1 representing a perfect fit. For the phantom images the value of r ranged from 0.993 to 0.999 indicating that a square root noise model is an appropriate representation of the image data. This measure, however, does not take into account any linear scaling of $\sigma_c^{\text{mod}}(y)$, therefore an additional accuracy measure, the relative mean squared error e, was calculated as

$$e = \frac{\sum_{k=1}^{K} \left(\sigma_c(k) - \sigma_c^{\text{mod}}(y_k)\right)^2}{\sum_{k=1}^{K} \left(\sigma_c^{\text{mod}}(y_k)\right)^2}. \tag{9}$$

A relative mean squared error of 0 indicates a perfect fit to the data. The maximum value of e for all 12 phantom images is 3.2×10^{-3} providing validation that $\sigma_c^{\text{mod}}(y)$ is an accurate measure of the high frequency noise.

In a strictly quantum-limited system the noise present at zero exposure (zero gray level) would be zero. This is represented by the noise model parameter d. Table I shows that the estimated values for d vary slightly around zero for different image acquisition conditions. In order to determine a representative noise value for zero exposure, we obtained an image from the GE Senographe system with a 4 mm aluminum plate placed in front of the detector. The contrast standard deviation $\sigma_c(0)$ was calculated to be 0.633, which is similar in magnitude to the estimated d values. When the remaining noise model parameter m was recalculated for the 12 phantom images with d fixed at 0.633, the value of the error e necessarily increased, with the maximum being 3.4×10^{-3}. This is still an acceptably small error value, suggesting that the noise model estimation could be reduced to finding a single parameter. Both the 1-parameter and 2-parameter methods are used in Section IV-E.

IV. FFDM IMAGES

In this section, noise equalization is applied to a dataset of FFDM images. The equalization technique is implemented as a preprocessing stage to a microcalcification computer-aided detection (CAD) scheme developed by Karssemeijer [17]. Three different preprocessing techniques are compared based on the performance of the microcalcification detection scheme. The first preprocessing technique implements the square root noise equalization model, while the second technique implements Veldkamps's contrast normalization method [11]. These two techniques are compared with the results obtained when noise equalization is omitted.

A. Acquisition

A database was constructed by selecting raw FFDM images obtained with the GE Senographe system. The dataset consists

TABLE I

STEP WEDGE PHANTOM IMAGE ACQUISITION PARAMETERS: TUBE VOLTAGE (KVP), ELECTRON DOSE (MAS), ENTRANCE DOSE (MGY), TARGET MATERIAL, FILTER MATERIAL. THE ENTRANCE DOSE WAS MEASURED BY ENTRANCE SURFACE AIR KERMA AT 5 CM ABOVE THE BUCKY. PARAMETERS m AND d ARE OBTAINED FROM (3). THE RELATIVE MEAN SQUARED ERROR e DETERMINES HOW WELL THE NOISE MODEL FITS THE DATA

Image	kVp	mAs	mGy	Target	Filter	m	d	$e(\times 10^{-3})$
1	26	180	14.9	Molybdenum	Molybdenum	0.253	0.415	3.200
2	26	280	23.1	Molybdenum	Molybdenum	0.239	0.707	1.100
3	31	71	8.7	Rhodium	Rhodium	0.255	0.530	0.431
4	31	200	24.4	Rhodium	Rhodium	0.283	-0.151	0.275
5	27	140	11.1	Molybdenum	Rhodium	0.265	-0.011	1.300
6	27	100	7.9	Molybdenum	Rhodium	0.253	0.275	1.300
7	30	110	11.2	Molybdenum	Rhodium	0.250	0.496	0.257
8	30	180	18.3	Molybdenum	Rhodium	0.269	0.051	0.179
9	30	180	18.7	Rhodium	Rhodium	0.277	-0.200	0.075
10	30	125	12.9	Rhodium	Rhodium	0.272	-0.090	0.134
11	28	71	7.3	Molybdenum	Molybdenum	0.236	0.631	1.300
12	28	125	12.9	Molybdenum	Molybdenum	0.258	0.155	1.500

of 124 mammograms composed of both craniocaudal and mediolateral oblique views from 33 patients (eight mammograms were not obtainable). Twenty eight of the mammograms contained a single microcalcification cluster (either benign or malignant). Labeling of the calcification clusters was performed by an expert radiologist who provided both the pathology and location of each true cluster. All other mammograms used in the study were classed as 'normal' for the purposes of the study as they were free of clustered microcalcifications on careful inspection, despite some showing other abnormalities, such as areas of increased density or parenchymal distortion. Each mammogram was 2294×1914 pixels in size and had a 100-μm resolution. The original 14 bits per pixel images were linearly rescaled to 12 bits.

B. Noise Model Application

To obtain $\sigma_c(k)$ for the FFDM images the grayscale was divided into nonoverlapping adjacent bins numbered $k = 1, 2, \ldots, 10$ with each bin built up of an approximately equal number of pixels. As for the phantom images, $\sigma_c(k)$ was calculated from $f(c \mid k)$ using the truncated distribution method and $\mu_c(k)$ was set to zero for all bins. The best fit continuous function $\sigma_c^{\mathrm{mod}}(y)$ was calculated using (3). Taking the signal dependency of $\sigma_c^{\mathrm{mod}}(y)$ into consideration, the local contrast feature c_i was normalized by

$$c_i' = \frac{c_i}{\sigma_c^{\mathrm{mod}}(y_i)} \qquad (10)$$

where c_i' represents the normalized local contrast at site i. This is very similar to (2), however, in this case $\mu_c(y)$ has been set to zero for all y.

It was found necessary to exclude the pixels nearest the breast-air boundary in estimating the noise model parameters. Therefore, the periphery of the breast region was eroded to a depth of 100 pixels, approximately corresponding to the extent of the uncompressed portion of the breast.

C. CAD Scheme

The method that is used for detection of microcalcifications in digital mammograms is based on the use of Bayesian techniques

and application of a Markov random field model, in which the latter models the fact that microcalcifications occur in clusters [17]. Each mammogram image is segmented into three categories: connective tissue, calcification and background. Starting from an initial segmentation, the labeling is optimized by applying an iterative rule for updating pixel labels.

The detection scheme uses three different features for representing the image data: the output of a line/edge detector and the normalized local contrast at two different spatial resolutions [17]. The use of a line/edge detector is necessary because both thin lines and regions with strong gradients may easily give rise to false-positive detections. The lower resolution local contrast is derived from c_i' by smoothing the normalized local contrast image using a 3×3 uniform filter kernel.

During the detection process, pixel labels are iteratively updated by maximizing their probability, given the image data in the small neighborhood of site i and given the current estimate for the rest of the labeling. The *a priori* probability of the labels is modeled by a Markov random field. For instance, a pixel is more likely to be part of a calcification if there are other calcifications in the neighborhood. The Markov random field model parameters mentioned in [17] were used for implementation of the CAD scheme.

D. Performance Evaluation

Four FFDM images from the database were selected at random to estimate the acquisition dependent portion of the CAD scheme. Each of these images was manually segmented into three categories: microcalcification, connective tissue and background. The conditional pdfs were estimated for each category using the segmented images. Due to the small size of the image database, separate testing and training datasets were not used. For the evaluation of the detection performance, the number of true and false positive clusters was determined for each mammogram as the sensitivity of microcalcification detection was varied. A FROC curve was constructed by plotting the true positive fraction as a function of the mean number of false positive clusters per image [18]. Labeling of the true clusters was done according to an expert radiologist who provided both the pathology and the location of the
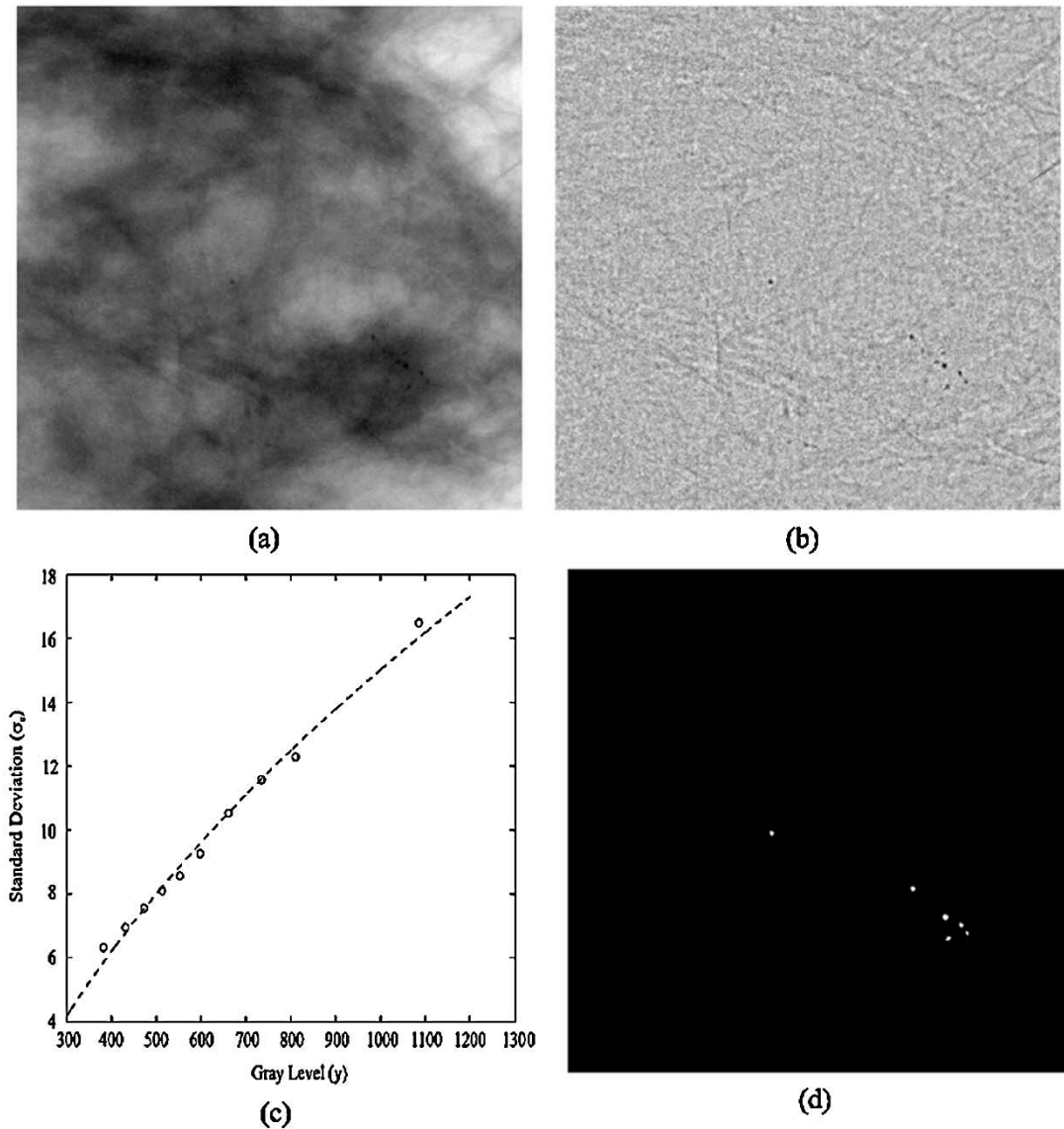
Fig. 4. (a) Part of a direct digital raw mammogram image showing a single microcalcification cluster. (b) Local contrast image corresponding to (a). (c) Square root fit $\sigma_c^{\mathrm{mod}}(y)$ "-," to the $\sigma_c(k)$ data "o" estimated for the complete mammogram except for near the breast-air boundary. (d) Calcification segments corresponding to (a).

true clusters. For counting the number of true positives, the calcification segmented images were clustered using mode cluster analysis [19]. True positive clusters were regarded as detected if two or more calcifications were found in the ground truth marked by the expert radiologist. All other clusters found were considered to be false positives. No verification of the detection of individual calcifications was performed.

E. Results

Fig. 4(a) shows part of a raw direct digital mammogram image containing a single microcalcification cluster. The corresponding local contrast image shown in Fig. 4(b) highlights the microcalcifications present by producing extreme negative local contrast values at their locations. A plot of $\sigma_c(k)$ versus gray level for the complete mammogram (but excluding the pixels near the breast-air boundary) can be seen in Fig. 4(c) along with the best fit square root noise model $\sigma_c^{\mathrm{mod}}(y)$. For the mammogram in Fig. 4(a) the relative mean squared error

(e) of the noise model was calculated to be 7.5×10^{-4}. The maximum value of e for all mammogram images in the dataset was 3.1×10^{-3}. This small error value confirms that the square root noise model is accurate for application to direct digital mammogram images.

Fig. 4(d) shows the corresponding microcalcification CAD output employing equalization based on the square root noise model. For this particular example both Veldkamp's noise equalization scheme and the square root noise equalization technique assist in the correct detection of the microcalcification cluster, however when noise compensation is omitted the cluster goes undetected.

The FROC curves shown in Fig. 5 evaluate the detection performance of the CAD system utilizing different preprocessing noise equalization schemes. The lowest curve represents the results obtained when noise equalization is omitted. This provides clear evidence that high frequency noise compensation is important for the detection of microcalcifications. An improvement is
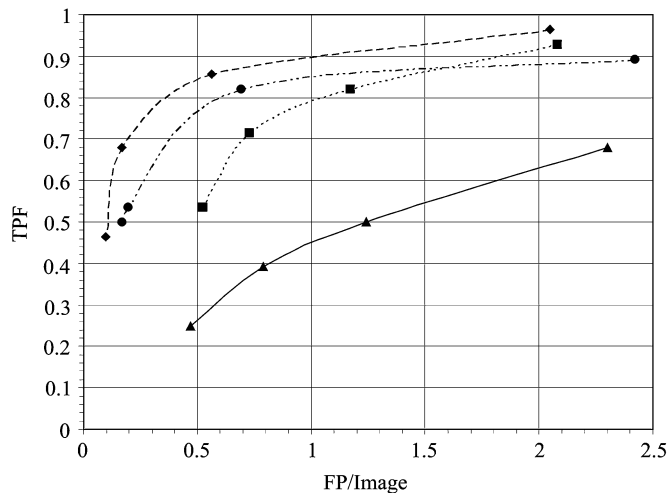
Fig. 5. FROC curves showing the CAD performance using different preprocessing conditions. ◇: square root noise model with an estimated d parameter; ○: square root noise model with a fixed d parameter; □: Veldkamp's noise equalization scheme; and △: when noise compensation is omitted.

shown in detection performance when using Veldkamp's noise equalization scheme described in [11], however the best performance is obtained when implementing the square root noise model. The CAD performance for both a fixed and varying square root noise model parameter d provides the top two FROC curves in Fig. 5. These two curves indicate that fixing d does not greatly reduce the sensitivity of the CAD system. When comparing the square root noise model with Veldkamp's equalization scheme, at a true positive fraction of 0.92 the number of false positives per image decreases from 1.8 to 1.2. These results demonstrate that the square root model based approach which FFDM allows leads to a more robust estimation of the high frequency image noise and thereby better microcalcification detection. It must be acknowledged that the same images were used for training and testing the detection scheme.

V. Discussion and Conclusion

The use of a theoretical model for the dependency of noise variance on intensity in FFDM allows a straightforward and accurate approach to noise equalization. The phantom image study shows that the square root noise model provides a good representation of the variation in contrast standard deviation with intensity. Either a 1-parameter or 2-parameter model was found to be adequate, but slightly superior performance in microcalcification detection was shown using the 2-parameter variant of the method. Care is needed to make sure that the system manufacturer has not applied some form of postprocessing to the mammograms which may invalidate the model; "raw" images should be used.

We found in the FFDM images that there was a significant deviation from the model in the peripheral area of the breast, where the tissue is uncompressed. We argue that in the peripheral area of the breast the image is not photon limited and noise processes other than quantum noise may play a significant part. Furthermore, recent experiments performed by researchers in Nijmegen show that irregularity of the skin surface, amplified

near the breast edge because of the small angle between the skin surface and the incident X-ray beam, contribute significantly to the standard deviation of the contrast there. More work is needed to fully explain the observations.

A consequence of excluding pixels near the breast edge from the noise estimation is that the maximum mean squared error found for the FFDM images was slightly smaller than that found for the phantom images. The exclusions meant that the range of exposures for the FFDM images was less than for the phantom images, therefore, the range of exposures over which the error in the noise model fit was estimated was also less, making it easier to achieve a good fit.

We found that artifacts within the phantom images caused overestimation of $\sigma_c(k)$, the contrast standard deviation for the kth bin. This estimation was improved by truncating the local contrast distributions to exclude these outliers and allowing for the truncation in the estimation. Although the likelihood of similar artifacts in actual mammograms is low, the same method has been found to give good results when applied to FFDM images. Our experience suggests that allowing for the existence of a small population of outliers in any image will increase the robustness of the noise equalization method over the range of images encountered.

The new noise equalization technique is applicable in combination with FFDM detector systems other than the amorphous silicon system used in this study. For example, slot-scan systems also provide raw images in which pixel values are proportional to exposure and in which image noise is dominated by quantum noise. Applicability to computer radiography (CR) images needs to be investigated, however, because in CR imaging the fixed pattern noise of the screens must be dealt with.

We have demonstrated that noise equalization, performed with the use of a square root model, produces more accurate detection of microcalcifications in FFDM than either no noise equalization or the more complex method proposed by Veldkamp *et al.* for digitized film-screen mammograms. An acceptably high sensitivity to microcalcifications was shown over a test set of 124 FFDMs with a relatively low number of false positive detections.

References

[1] S. A. Feig, "Decreased breast cancer mortality through mammographic screening: Results of a clinical trial," *Radiology*, vol. 167, no. 3, pp. 659–665, June 1988.

[2] R. Highnam and M. Brady, *Mammographic Image Analysis*, 1st ed, Dordrecht, The Netherlands: Kluwer Academic, 1999.

[3] N. Karssemeijer, J. T. Frieling, and J. H. Hendriks, "Spatial resolution in digital mammography," *Investigat. Radiol.*, vol. 28, pp. 413–419, May 1993.

[4] N. Karssemeijer, "Adaptive noise equalization and detection of microcalcification clusters in mammography," *Int. J. Pattern Recogn. Artificial Intell.*, vol. 7, no. 6, pp. 1357–1375, May 1993.

[5] R. M. Nishikawa, M. L. Giger, K. Doi, C. J. Vyborny, and R. A. Schmidt, "Computer-aided detection of clustered microcalcification: An improved method for grouping detected signals," *Med. Phys.*, vol. 20, no. 6, pp. 1661–1666, Nov. 1993.

[6] H. P. Chan, K. Doi, C. J. Vyborny, R. A. Schmidt, C. E. Metz, K. L. Lam, T. Ogura, Y. Wu, and H. Macmahon, "Improvement in radiologist's detection of clustered microcalcifications on mammograms," *Investigat. Radiol.*, vol. 25, no. 10, pp. 1102–1110, Oct. 1990.

[7] H. Neiber, T. Muller, and R. Stotzka, "Local contrast enhancement for the detection of microcalcifications," in *Proc. 5th Int. Workshop Digital Mammography*, M. J. Yaffee, Ed., Nijmegen, The Netherlands, 2000, pp. 598–604.

[8] R. N. Strickland and H. I. Hahn, "Wavelet transforms for detecting micro-calcifications in mammograms," *IEEE Trans. Med. Imag.*, vol. 15, pp. 218–228, Apr. 1996.

[9] W. Qian, M. Kallergi, L. P. Clarke, H. D. Li, P. Venugopal, D. Song, and R. A. Clark, "Tree structured wavelet transform segmentation of microcalcifications in digital mammograms," *Med. Phys.*, vol. 22, pp. 1247–1254, Aug. 1995.

[10] W. Qian, M. Kallergi, L. P. Clarke, and R. A. Clark, "Tree-structured nonlinear filters in digital mammography," *IEEE Trans. Med. Imag.*, vol. 13, pp. 25–36, Mar. 1994.

[11] W. Veldkamp and N. Karssemeijer, "Normalization of local contrast in mammograms," *IEEE Trans. Med. Imag.*, vol. 19, pp. 731–738, July 2000.

[12] T. Netseh and H. O. Peitgen, "Scale-space signatures for the detection of clustered microcalcifications in digital mammograms," *IEEE Trans. Med. Imag.*, vol. 18, pp. 774–786, Sept. 1999.

[13] A. Maitournam, C. Graffigne, and A. Stauss, "Modeling of digital mammograms using bicubic spline function and additive noise," in *Proc. 4th Int. Workshop Digital Mammography*, N. Karssemeijer, M. A. O. Thijssen, J. H. Hendrilcs, and L. J. vanErning, Eds., Nijmegen, The Netherlands, 1998, pp. 177–184.

[14] M. Poissonnier and M. Brady, "Noise equalization, film-screen artifacts, and density representation," in *Proc. 5th Int. Workshop Digital Mammography*, M. J. Yaffee, Ed., Toronto, ON, Canada, 2000, pp. 605–611.

[15] M. J. Yaffee, "Digital Mammography," in *Handbook of Medical Imaging: Physics and Psychophysics*, J. Beutel, H. L. Kundel, and R. L. VanMetter, Eds. Bellingham, WA: S.P.I.E., 2000, vol. 1, pp. 329–372.

[16] R. E. Hendriks, J. M. Lewin, P. Isaacs, C. K. Kuni, C. J. D'orsi, L. Moss, A. Karellas, G. A. Sisney, and G. R. Cutter, "Interim clinical evaluation of FPDM in a screening cohort: Comparison with screen film mammography in 4 965 exams," in *Proc. 5th Int. Workshop Digital Mammography*, M. J. Yaffee, Ed., Toronto, ON, Canada, 2000, pp. 468–474.

[17] N. Karssemeijer, "Stochastic model for automated detection of calcifications in digital mammograms," *Image Vis. Computing*, vol. 10, no. 6, pp. 369–375, Aug. 1992.

[18] D. B. Kopans, *Breast Imaging*, 2nd ed. Philadelphia, PA: Lippincott-Raven, 1998.

[19] B. S. Everitt, *Cluster Analysis*, 3rd ed, London, U.K.: Halsted, 1993.