Restoration of Low-Dose Digital Breast Tomosynthesis

Lucas R. Borges[1,3,*], Lucio Azzari[2,3], Predrag R. Bakic[4], Andrew D. A. Maidment[4], Marcelo A. C. Vieira[1], Alessandro Foi[2,3]

[1]Department of Electrical and Computer Engineering, University of São Paulo, Brazil
[2]Laboratory of Signal Processing, Tampere University of Technology, Finland
[3]Noiseless Imaging Ltd, Tampere, Finland
[4]Department of Radiology, University of Pennsylvania, USA

[*]corresponding author; e-mail: lucas.rodrigues.borges@usp.br.

In breast cancer screening, the radiation dose must be kept to the minimum necessary to achieve the desired diagnostic objective, thus minimizing risks associated with cancer induction. However, decreasing the radiation dose also degrades the image quality. In this work we restore digital breast tomosynthesis (DBT) projections acquired at low radiation doses with the goal of achieving a quality comparable to that obtained from current standard full-dose imaging protocols. A multiframe denoising algorithm was applied to low-dose projections, which are filtered jointly. Furthermore, a weighted average was used to inject a varying portion of the noisy signal back into the denoised one, in order to attain a signal-to-noise ratio comparable to that of standard full-dose projections. The entire restoration framework leverages a signal-dependent noise model with quantum gain which varies both upon the projection angle and on the pixel position. A clinical DBT system and a 3D anthropomorphic breast phantom were used to validate the proposed method, both on DBT projections and slices from the 3D reconstructed volume. The framework is shown to attain the standard full-dose image quality from data acquired at 50% lower radiation dose, whereas progressive loss of relevant details compromises the image quality if the dosage is further decreased.

Keywords: denoising, digital breast tomosynthesis, dose reduction, variance stabilization

1 Introduction

Digital breast tomosynthesis (DBT) is becoming a major clinical tool for breast cancer screening. In a DBT exam, discrete 2D projections are acquired within a limited angle arc around the breast. The 2D projections are then processed to reconstruct slices of a 3D volume of

the breast, reducing limitations related to tissue overlap, commonly found in conventional 2D mammography.

In DBT, as in other x-ray imaging modalities, the radiation dose must be kept at the minimum necessary to achieve the desired diagnostic objective [1], thus minimizing the risks associated with cancer induction [2, 3]. However, decreasing the radiation dose also degrades the image quality, which can affect the radiologists' performance when giving a diagnosis [4, 5, 6]. Thus, the trade-off between radiation dose and image quality must be carefully balanced.

A number of works have shown the potential of denoising at improving image quality in medical imaging, *e.g.*, for computed tomography (CT) [7, 8] and magnetic resonance (MRI) [9, 10, 11]. However, there are very few works exploring denoising for DBT. In [12], the authors proposed the application of a patch-based denoising algorithm to raw DBT projections. Another work investigated the application of denoising to the slices of reconstructed DBT volumes [13]. In recent work [14], our group has described a noise model for DBT projections that accounts for specific characteristics such as spatially-varying quantum gain, signal-dependent noise, electronic noise and pixel offset. Further, we have adopted this detailed noise model to propose an efficient denoising pipeline for DBT raw projections [15]. However, two main aspects limits the clinical usability of DBT images denoised by the abovementioned methods. First, the denoising introduces bias, typically in the form of smearing of singularities, which can be especially problematic for medical images. Second, denoised images lack the granular texture-like appearance typical of clinical images, and therefore are perceived as odd by the trained radiologists.

In this context, we here present a quantitative assessment of the above issues and we propose the combination of noisy and denoised low-dose DBT projections to recover the noise properties and visual appearance of a standard full-dose DBT image. Furthermore, the image combination decreases the bias introduced by the denoising process. The image combination is performed through a weighted average with weighting coefficients calculated based on the target noise characteristics. Other contributions of this work include the careful pre-processing of the quantum gain map, the use of a multiframe denoising algorithm, the description of a metric that evaluates bias and residual noise separately and the extensive validation performed not only on the projections, but also on the reconstructed slices.

2 Observation model

Let $z : X \times \Theta \to \mathbb{R}^+$ be an ensemble of DBT projections, where $X \subset \mathbb{Z}^2$ is the 2D set of pixel spatial coordinates spanning each projection and $\Theta \subset (-\pi, \pi]$ is a discrete set of projection angles. We model the observed pixels $z(x, \theta)$, $x \in X$, $\theta \in \Theta$, by the affine-variance model [14] which is based on the Poisson-Gaussian statistics of the system [16, 17] and links the expectation and variance of z as

$$\mathrm{E}\left\{z(x,\theta)\,|y(x,\theta)\right\} = y(x,\theta) + \tau\,, \tag{1}$$

$$\mathrm{var}\left\{z(x,\theta)\,|y(x,\theta)\right\} = \lambda(x,\theta)\,y(x,\theta) + \sigma_E^2\,. \tag{2}$$

where $y > 0$ is a (unknown) noise-free signal proportional to the energy of the x-rays reaching the detector, $\tau > 0$ is a constant signal offset, $\lambda : X \times \Theta \to \mathbb{R}^+$ is the linear coefficient of the noise variance function, which can be attributed to the quantum efficiency and gain in the image formation, and $\sigma_E^2 > 0$ is a constant component, representing the variance of the signal-independent electronic noise. Both τ and σ_E^2 are constant and common to each pixel, whereas λ varies depending on the spatial coordinate x and angle θ. Hence, each pixel $z(x, \theta)$ may follow a different signal-dependent noise model.

Now, let us consider DBT projections z_γ acquired using a lower current-time product, and therefore resulting in lower radiation dose, where $0 < \gamma < 1$ denotes the dose-reduction

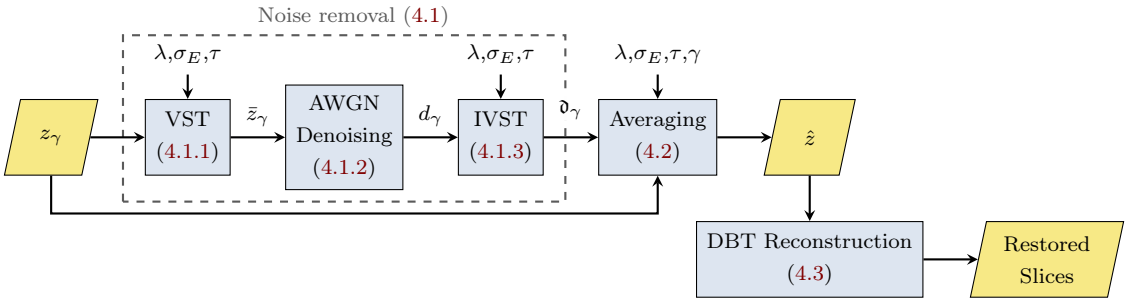


Figure 1: Flowchart of the proposed restoration pipeline. The input z_γ is an ensemble of low-dose DBT projections, whereas $\hat{z}$ is an approximation of the standard full-dose DBT projections; $\hat{z}$ is fed into a DBT reconstruction software to obtain 3D-volume slices. The numbers in parentheses indicate the corresponding subsections with details.

factor. Because the dose reduction is achieved exclusively by reducing the current-time product, reducing the dose by a factor γ is equivalent to scaling the noise-free signal y by the same factor [18]. Hence, the expectation and variance of z_γ are

$$\mathrm{E}\left\{z_\gamma(x,\theta)\,|y(x,\theta)\right\} = \gamma y(x,\theta) + \tau\,, \tag{3}$$

$$\mathrm{var}\left\{z_\gamma(x,\theta)\,|y(x,\theta)\right\} = \gamma\lambda(x,\theta)\,y(x,\theta) + \sigma_E^2\,. \tag{4}$$

Note that the signal-to-noise ratio (SNR) of z_γ is smaller than the SNR of z. The SNR can be computed as the ratio between the squared expectation and variance of $(z_\gamma(x,\theta) - \tau)/\gamma$, which equals $\frac{\gamma y}{\lambda + \sigma_E^2 \gamma^{-1} y^{-1}}$ and which goes to zero with γy. This is a property common to various radiation-based imaging systems.

In this work we restrict the modeling of the image-formation process of the projections to the above equations, without attempting to express y as, *e.g.*, the Radon transform of the 3D volume to be reconstructed. Accordingly, we treat the DBT reconstruction process as a black box and therefore we do not present an observation model for the reconstructed slices of the 3D volume. As a matter of fact (as we detail in Section 4.3), we utilize a commercial DBT reconstruction software whose algorithm is confidential and not available to us.

3 Problem formulation

The primary goal of this work is to obtain an approximation $\hat{z}$ of the standard full-dose projections z by processing the low-dose projections z_γ. Specifically, we desire that, in terms of expectation and variance, $\hat{z}$ matches the model (1)-(2). As a secondary goal we also want that the slices of the 3D volume reconstructed from the processed projections $\hat{z}$ are consistent with those reconstructed from the full-dose z. The achievement of these goals is the premise for considering the proposed approach as a potential pre-processing stage in DBT so to obtain comparable image quality from lower-dose acquisition.

4 Methods

Figure 1 shows the proposed pipeline used to produce an approximation of the 3D-volume slices of the standard full-dose DBT. We first use the lower dose acquisition z_γ to compute the approximation $\hat{z}$, which is eventually fed into a DBT reconstruction software to obtain the estimated 3D-volume slices. In this section we introduce and describe its main steps.

4.1 Noise removal

Since most of the filters available off-the-shelf are designed and optimized for the additive white Gaussian noise (AWGN) model, we adopt a variance-stabilization framework in order to enable the use of those available filters and thus make our pipeline more versatile. The typical variance-stabilization framework consists of three main steps: the noisy signal is first processed by applying a variance stabilizing transformation (VST) which produces a signal that can be treated as corrupted by AWGN; then, any filter for AWGN can be used to effectively attenuate the noise; finally, an inverse VST is applied to return the signal to its original range. The advantage of this framework is that it is independent of the particular choice of AWGN filter, so that any practitioner can easily adopt it in their system with their filter of choice. The VST framework is thus widely adopted and is shown to provide state-of-the-art restoration performance in various applications [19, 20, 21, 22, 23, 24, 25].

We detail next the adopted VST framework, which is appropriate for the observation model (3)-(4).

4.1.1 Variance stabilizing transformation (VST)

We stabilize the noise variance (4) of the set z_γ of low-dose projections using the generalized Anscombe VST [26]:

$$\bar{z}_\gamma(x,\theta) = \begin{cases} 2\sqrt{T\left[z_\gamma(x,\theta)\right] + 3/8} & T\left[z_\gamma(x,\theta)\right] \geq -3/8 \\ 0 & T\left[z_\gamma(x,\theta)\right] < -3/8\,, \end{cases} \tag{5}$$

where $T\left[z_\gamma(x,\theta)\right] = \lambda^{-1}(x,\theta)\, z_\gamma(x,\theta) + \lambda^{-2}(x,\theta)\, \sigma_E^2 - \lambda^{-1}(x,\theta)\, \tau$. Under a Poisson-Gaussian modeling of $z_\gamma(x,\theta)$, it can be shown that $\bar{z}_\gamma(x,\theta)$ is asymptotically Gaussian for large $y(x,\theta)$ and that in particular [26, 20]

$$\mathrm{E}\left\{\bar{z}_\gamma(x,\theta)\,|y(x,\theta)\right\} = 2\sqrt{T\left[\mathrm{E}\left\{z_\gamma(x,\theta)\,|y(x,\theta)\right\}\right] + 1/8 + \mathcal{O}(y^{-1}(x,\theta))}\,, \tag{6}$$

$$\mathrm{var}\left\{\bar{z}_\gamma(x,\theta)\,|y(x,\theta)\right\} = 1 + \mathcal{O}\big(y^{-2}(x,\theta)\big)\ . \tag{7}$$

In practice, based on (7), the noise corrupting $\bar{z}_\gamma(x,\theta)$ is treated as standard Gaussian $\mathcal{N}(0,1)$ in the denoising. However, comparing (6) with (5), we note a discrepancy of $1/4 + \mathcal{O}(y^{-1}(x,\theta))$, which corresponds to bias introduced when taking expectations inside and outside of the non-linear square root function. This bias is ignored when assuming zero-mean noise in $\bar{z}_\gamma(x,\theta)$, and will thus be addressed and compensated after denoising in Section 4.1.3.

4.1.2 AWGN denoising: RF3D

Projections of an object (in this case a breast) at adjacent angles feature various similarities which can be exploited for denoising. Noting that the similarity between adjacent projections is comparable to the similarity between consecutive frames of a video, we consider the set of noisy projections as a *de-facto* video, and we use a video-denoising algorithm, namely RF3D [27], to denoise the ensemble of noisy projections $\bar{z}_\gamma$.

In this context, the first step of RF3D consists of subdividing $\bar{z}_\gamma$ into a multitude of partially overlapping spatio-angular subvolumes: for each square block within a projection, the algorithm finds the most similar blocks in adjacent projections, which are thus stacked along the angular dimension to build a subvolume. Each such subvolume is characterized by spatial *intra-block* regularity inherited by the local content of each projection, and by angular *inter-block* regularity due to the similarity between stacked blocks. Each subvolume is then decorrelated by a 3D transform and filtered by shrinkage of the transform spectrum. This step suppresses noise

because the spectrum is sparse by virtue of the two forms of regularity within the subvolume: most of the spectrum coefficients have low magnitude and contain mostly noise, whereas the signal of interest is captured by few large-magnitude coefficients. The shrinkage threshold is proportional to the noise standard deviation in $\bar{z}_\gamma$, which is unitary as a result of the VST (5). Next, the filtered spectrum is inverse transformed, providing an estimate of the clean subvolume. Because of the overlap between subvolumes, we have multiple estimates for each pixel in the set of projections; therefore RF3D concludes with an aggregation step where these multiple estimates are averaged together using adaptive weights proportional to the sparsity of the respective subvolume spectra. In this way, RF3D produces an approximation $d_\gamma = \Phi[\bar{z}_\gamma]$ of the expectation $\mathrm{E}\{\bar{z}_\gamma|y\}$.

We refer the reader to [27] for further details and note that an implementation of this algorithm is publicly available.

4.1.3 Inverse VST

The denoised $d_\gamma = \Phi[\bar{z}_\gamma]$ is returned to its original intensity range by the exact unbiased inverse of the generalized Anscombe transform [20], producing a set of filtered projections $\mathfrak{d}_\gamma$ which approximates $\mathrm{E}\{z_\gamma|y\}$:

$$\mathfrak{d}_\gamma(x,\theta) = \lambda(x,\theta)\left(\frac{1}{4}d_\gamma^2(x,\theta) + \frac{1}{4}\sqrt{\frac{3}{2}}d_\gamma^{-1}(x,\theta) - \frac{11}{8}d_\gamma^{-2}(x,\theta) + \right. \\ \left. + \frac{5}{8}\sqrt{\frac{3}{2}}d_\gamma^{-3}(x,\theta) - \frac{1}{8} - \frac{\sigma_E^2}{\lambda^2(x,\theta)}\right) + \tau. \tag{8}$$

Specifically, this form of inversion compensates the bias discussed in Section 4.1.1, so that if $d_\gamma = \mathrm{E}\{\bar{z}_\gamma|y\}$ then $\mathfrak{d}_\gamma = \mathrm{E}\{z_\gamma|y\}$. Note that this form of inversion is publicly available for download at [20].

4.2 Noisy-denoised weighted average

As specified in Section 3, the aim of this work is to obtain a set of projections $\hat{z}$ such that

$$\mathrm{E}\{\hat{z}(x,\theta)\,|y(x,\theta)\} = \mathrm{E}\{z(x,\theta)\,|y(x,\theta)\} = y(x,\theta) + \tau, \tag{9}$$
$$\mathrm{var}\{\hat{z}(x,\theta)\,|y(x,\theta)\} = \mathrm{var}\{z(x,\theta)\,|y(x,\theta)\} = \lambda(x,\theta)\,y(x,\theta) + \sigma_E^2. \tag{10}$$

To this end, we adopt a weighted average of the form

$$\hat{z}(x,\theta) = w_\gamma(x,\theta)\,(z_\gamma(x,\theta) - \tau) + \bar{w}_\gamma(x,\theta)\,(\mathfrak{d}_\gamma(x,\theta) - \tau) + \tau, \tag{11}$$

where w_γ and $\bar{w}_\gamma$ are weights that dependent both on the pixel position x and on the projection angle θ. Assuming that the denoising is ideally successful, *i.e.* $\mathfrak{d}_\gamma(x,\theta) = \mathrm{E}\{z_\gamma(x,\theta)\,|y(x,\theta)\}$, it trivially follows from (3) that setting

$$\bar{w}_\gamma(x,\theta) = \gamma^{-1} - w_\gamma(x,\theta) \tag{12}$$

yields (9). This substitution also results in $\mathrm{var}\{\hat{z}(x,\theta)\,|y(x,\theta)\} = w_\gamma^2(x,\theta)\,\mathrm{var}\{z_\gamma(x,\theta)\,|y(x,\theta)\}$ for any choice of $w_\gamma(x,\theta)$. Hence, to obtain (10), it suffices to set

$$w_\gamma(x,\theta) = \sqrt{\frac{\mathrm{var}\{z(x,\theta)\,|y(x,\theta)\}}{\mathrm{var}\{z_\gamma(x,\theta)\,|y(x,\theta)\}}} = \sqrt{\frac{\lambda y(x,\theta) + \sigma_E^2}{\gamma\lambda(x,\theta)\,y(x,\theta) + \sigma_E^2}}\,. \tag{13}$$

Note that the noise-free signal y is not available. Thus, in practice, we replace it in (13) by the estimate $\hat{y}(x,\theta) = (\mathfrak{d}_\gamma(x,\theta) - \tau)/\gamma$.

4.3 DBT reconstruction

Projections were reconstructed into DBT slices using the commercially available reconstruction software Briona Standard v4.0 (Real Time Tomography, Villanova, PA). A general overview of this software is given in [28].

5 Materials & Experimental setup

The validation of the proposed method is performed by comparing the restored low-dose acquisitions to the standard full-dose acquisitions, both before and after the reconstruction process.

5.1 Materials

The clinical acquisitions used for this work were acquired by a Selenia Dimensions (Hologic, Bedford, MA) DBT system at the Hospital of the University of Pennsylvania. The system is equipped with an amorphous selenium (a-Se) detector that performs direct conversion of x-rays, therefore reporting minimally correlated noise [14]. The detector has a pixel pitch of 70 μm, with 2×2 binning, thus resulting in a 140 μm pixel pitch in the projections. The acquisition geometry consists of 15 projections, obtained at approximately uniform steps within a 15° arc around the breast, *i.e.* $\Theta = \{\theta_1 = -7.5°, \ldots, \theta_{15} = 7.5°\}$.

To allow repeated exposures of the same object, a 3D anthropomorphic physical breast phantom was used as the imaging object. The phantom was prototyped by CIRS, Inc. (Reston, VA), under the license from the University of Pennsylvania [29, 30]. It consists of six slabs, each containing simulated anatomical structures manufactured using materials that mimic the physical properties of the human breast tissue. The phantom simulates a 450 ml breast, compressed to 5 cm, with 17% volumetric breast density (excluding the skin).

In addition to the healthy breast anatomy, individual pieces of calcium oxalate (99%, Alfa Aesar, Ward Hill, MA) with different diameters were placed between slabs of the phantom to simulate microcalcifications. These microcalcifications were positioned in clusters, which are a common indicative of cancer development. Figure 2 shows an image of the phantom slabs. The red arrows indicate the location of the microcalcification clusters.

To estimate the parameters τ, $\lambda(x, \theta)$, and σ_E, calibration images were acquired using a 4-cm thick polymethyl methacrylate (PMMA) uniform calibration phantom. This phantom is commonly used for uniformity tests by the medical physicists, and was available at the Hospital of the University of Pennsylvania.

The image acquisition was divided into two parts, both conducted on the same day, therefore avoiding any changes on the system calibration.

During the first part of the acquisition, images were acquired using the 3D anthropomorphic physical breast phantom. The phantom was positioned close to the posterior edge of the breast support; the compression plate was lowered to touch the material; at the central projection angle ($\theta_8 = 0°$), the phantom slabs were parallel to the detector and perpendicular to the x-ray beam axis.

The automatic exposure control (AEC) was used for defining the standard full-dose radiographic parameters for this phantom: the machine estimates the optimal peak kilovoltage (kV_p) and current-time product (mAs) using information such as breast thickness and estimated density. The obtained parameters were 31 kV_p and 60 mAs.

The system was then used in manual mode for all the remaining acquisitions.

To obtain 100% (standard full dose), 50% ($\gamma = 0.5$), 25% ($\gamma = 0.25$), 15% ($\gamma = 0.15$), and 5% ($\gamma = 0.05$) of the standard full dose while maintaining the peak kilovoltage fixed at 31 kV_p,

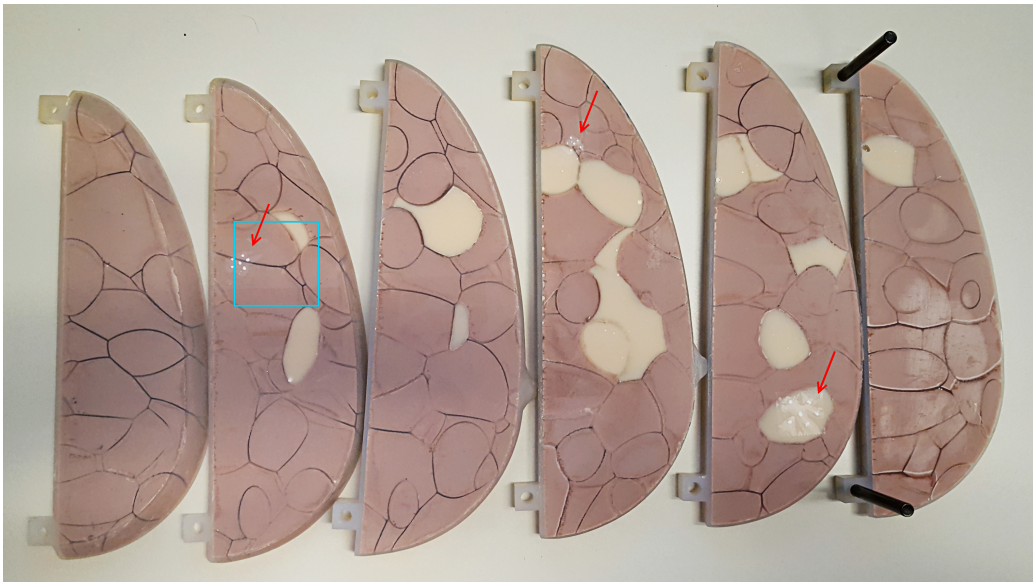


Figure 2: Picture of the 3D anthropomorphic breast phantom used throughout the experiments. The red arrows indicate the position of the microcalcification clusters; the cyan square indicates the location of the region of interest shown in Figure 6.

we set the current-time product to 60 mAs, 30 mAs, 15 mAs, 9 mAs, and 3 mAs, respectively. Ten realizations were acquired at each current-time product.

A separate set of ten full-dose (60 mAs) acquisitions of the phantom were averaged together to produce virtually noise-free projections, which we treat as ground-truth projections for the objective validation. These ground-truth projections were also used to reconstruct the ground-truth slices.

During the second part of the acquisition, calibration images were acquired using the uniform PMMA block. Ten realizations were acquired at each of the kV_p and mAs combinations adopted for the anthropomorphic phantom.

5.2 Metrics

Equations (1) and (3) show that projections acquired at different dose levels have different mean values. Thus, to allow the comparison against the ground-truth (obtained from standard full-dose acquisitions), all the processed projections are first mapped to the intensity range of the ground-truth. This is done by fitting affine functions which map pixel values from each of the low-dose projections to the corresponding pixel values in the corresponding ground-truth projection, as adopted in [31]. The same technique was applied to the reconstructed slices. This was done to allow direct comparison between images obtained from acquisitions at different dose levels and to make the error metrics robust with respect to systematic differences in pixel intensities.

The main metric used for validation is the mean normalized squared error (MNSE), which we compute as follows: first, for every pixel, we compute the normalized quadratic error as the squared difference from the ground-truth divided by the local contrast, defined as the variance of the ground-truth over a window of size 64×64 pixels; second, the MNSE is obtained as the average of the pixelwise normalized quadratic errors on the breast over the entire set of projections or slices. We adopted a normalized version of the MSE due to the high-dynamic range of the images. The normalization guarantees that regions with significantly higher or lower contrast are equally accounted by the metric. Furthermore, we decompose the MNSE into variance and squared bias portions. The rationale of separately analyzing variance and

Figure 3: Central ground-truth projection, corresponding binary mask, and central projection superposed by the binary mask. The nipple region has been magnified for better visual assessment.

bias within our pipeline is that the variance is informative about the similarity between the noise in restored and standard full-dose images, whereas the bias assesses the impact of the pipeline to the underlying signal and is indicative of systematic distortions such as smearing of details.

These metrics are calculated exclusively on the pixels inside the breast, ignoring the background. The inner region was selected by binary masks, generated by comparing the intensity of the ground-truth pixels to a threshold, followed by a morphological erosion operator; Figure 3 shows an example of these masks. The threshold has been chosen manually from the image intensity histogram, where background and foreground are well separated from each other.

We also measure the in-plane resolvability of singularities using the full width at half maximum (FWHM) of a microcalcification, as in [32]. The line profile of a selected microcalcification is taken from each of the ten realizations of the reconstructed in-plane slice. Next, the average profile is fitted using the sum of a Gaussian function, which represents the microcalcification, and a linear function, which represents the local background. The FWHM is then computed as

$$\mathrm{FWHM} = 2r\sigma\sqrt{2\ln 2}\,, \tag{14}$$

where σ is the standard deviation of the fitted Gaussian and r is the pixel size of the detector. Higher values of FWHM from a same structure indicate loss of resolution.

5.3 Projection alignment

Since each projection is slightly shifted vertically with respect to its adjacent projections (due to the acquisition geometry), before denoising we first align vertically each projection $z_\gamma(\cdot,\theta)$, $\theta\in\Theta$, to the first one $z_\gamma(\cdot,\theta_1)$. This improves the denoising performance of RF3D because it increases the chances of successful block matching within a smaller search region and under stronger noise. We compute each vertical shift by finding the maximum of the vertical cross-correlation between two projections. After denoising, the projections are shifted back to their original position.

5.4 Parameter estimation

The noise parameters have been estimated as described in [14]. In particular, the signal offset τ has been computed as described by the National Health Service Breast Screening Program (NHSBSP) in [33]. The signal-independent noise variance σ_E^2 has been estimated using flat regions taken from each calibration set: since the noise variance (2) is affine, σ_E^2 is the intercept of

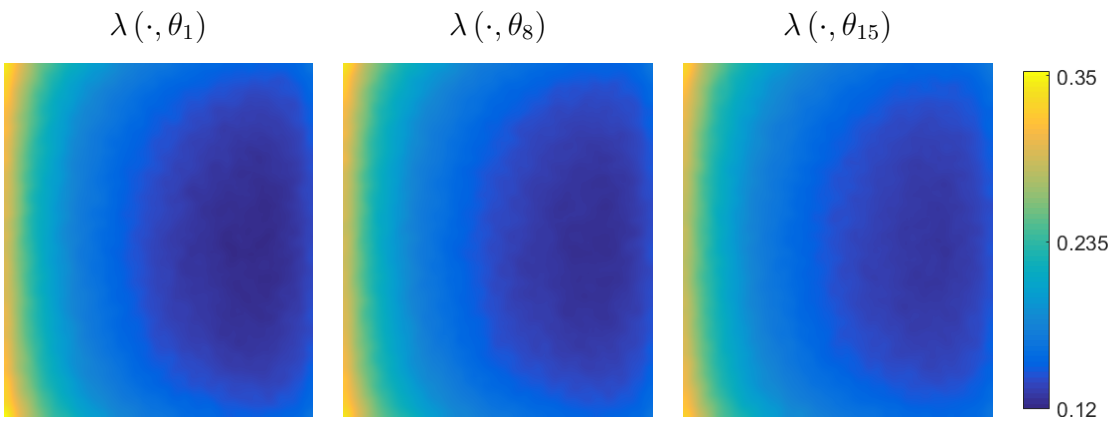


Figure 4: Coefficient maps $\lambda(\cdot, \theta)$ of the linear component of the signal-dependent noise, visualized for the first (θ_1), middle (θ_8), and last (θ_{15}) angles in the set of projections.

the line fitting the mean-variance scatterplot computed from those flat regions. The coefficient $\lambda(x, \theta)$ for the linear signal-dependent part of the noise has been instead estimated as

$$\lambda(x, \theta) = S\left[\frac{\hat{\sigma}^2(x, \theta) - \sigma_E^2}{\hat{\mu}(x, \theta) - \tau}\right], \tag{15}$$

where $\hat{\mu}(x, \theta)$ and $\hat{\sigma}^2(x, \theta)$ are, respectively, estimates of the local signal mean and noise variances computed on a 32×32 sliding window over the calibration sets and S is a quadratic polynomial smoother applied for regularization. In Figure 4 we show the 1st, 7th, and 15th angular plane of λ estimated from the full-dose acquisition.

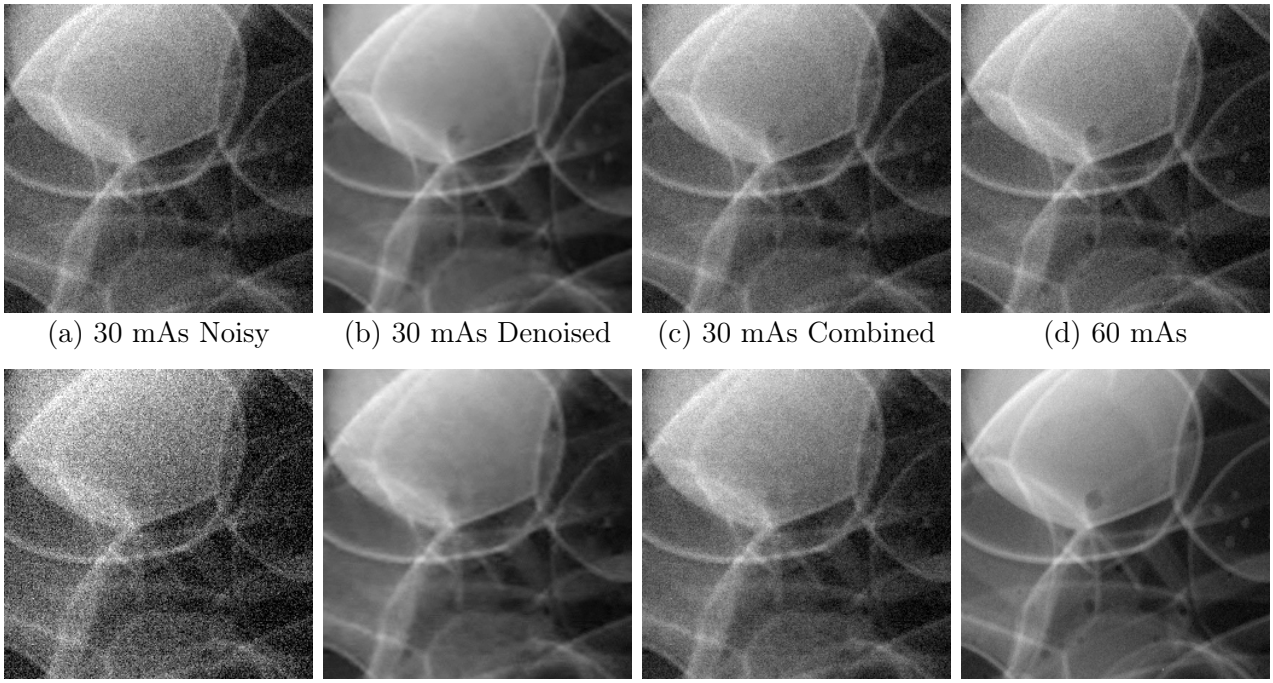
5.5 Implementation

The pipeline was validated using our MATLAB single-threaded CPU-based implementation. The total processing time was 17.3 (±0.3) minutes per DBT set (15 projections) on a 2.8-GHz Intel Xeon E3-1505M CPU. Note that most of the computational time arises from the RF3D denoising, as other processing steps are trivial in terms of computational cost.

6 Results

Figure 5 shows a magnified region of interest (ROI) with 256×256 pixels, for visual evaluation of the central projection at different noise levels. Note that the selected ROI contains a cluster of microcalcifications located on the right side, and that these microcalcifications are the hardest region of the dataset to recover. We study the behaviour of the proposed method on this ROI in order to consider the worst case scenario.

Figure 5(a) and Figure 5(e) show the ROIs taken from the input low-dose images at 30 mAs and 9 mAs, respectively. As expected, they are noisier than the standard full-dose acquisition in Figure 5(d). Figure 5(b) and Figure 5(f) show the ROIs taken from the denoised images. Note that Figure 5(b) and Figure 5(f) are denoised and do not contain any of the noise seen in the standard full-dose acquisition in Figure 5(d). Figure 5(c) and Figure 5(g) are the resulting approximations of the standard full-dose acquisition in Figure 5(d). A comparison between Figure 5(c) and Figure 5(d) shows that the combined image at 30 mAs (50% of the standard full dose) achieves visually similar results to the standard full-dose image. Meanwhile, the comparison between Figure 5(g) and Figure 5(d) (9 mAs, 15% of the standard full dose) shows that some details were lost during the restoration process, which indicates the limitations of the proposed method.



(a) 30 mAs Noisy (b) 30 mAs Denoised (c) 30 mAs Combined (d) 60 mAs

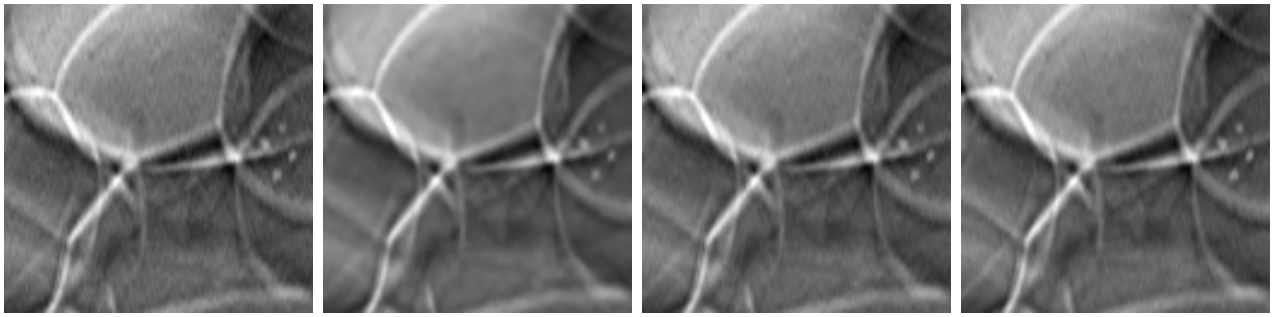
(e) 9 mAs Noisy (f) 9 mAs Denoised (g) 9 mAs Combined (h) Ground-truth

Figure 5: Magnified region of interest taken from the central projection (before reconstruction). For completeness, we also show the ROI from the ground-truth projection, although it plays no role for the visual validation: the goal is to approximate the 60 mAs full-dose image with the combined low-dose images.

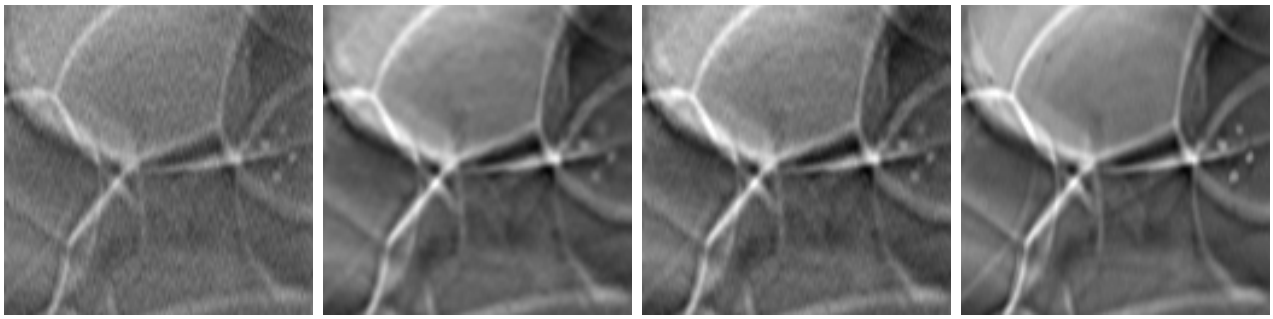
Reconstructed DBT slices, parallel to the breast support, are obtained by feeding each noisy or processed set of projections to the reconstruction software. Note that no denoising or image combination is applied after reconstruction. Figure 6 shows a ROI with the same microcalcifications shown in Figure 5, but now taken from reconstructed slices. This 256×256 ROI is located in the phantom as illustrated in Figure 2. Figure 6(a) and Figure 6(e) show the ROI taken from the slices reconstructed from the noisy projections at 30 mAs and 9 mAs, respectively. Noteworthy, Figure 6(e) features lower contrast with respect to Figure 6(a). We hypothesize that these differences come from a change of behavior on the reconstruction software, due to the high levels of noise. The denoised and combined images shown in Figure 6(f) and Figure 6(g), respectively, yield improved contrast comparable to the contrast seen on the standard full-dose image shown in Figure 6(d). The comparison between Figure 6(c) and Figure 6(d) provide visual evidence that the pipeline was able to recover the quality of a standard full-dose image with minimal loss of sharpness, even when analyzed on the reconstructed domain. Some differences in image sharpness can be seen by analyzing the top calcification of the cluster, which has a sharper aspect in Figure 6(d) compared to Figure 6(c). On the other hand, if the pipeline is used to recover images acquired at extremely low dose there can be significant loss of sharpness of the calcifications, as in the case of Figure 6(g).

Cross-sections of the top calcification seen in Figure 6 is shown in Figure 7. The profiles show that the denoising of the projections degrades the contrast of the microcalcification within the DBT slice. However, the contrast is partly recovered after the weighted averaging. In particular, the DBT slices reconstructed from combined projections at 50% (30 mAs) yielded a profile similar to that from standard full dose.

Table 1 shows the MNSE as well as the decomposition into bias and variance, calculated both on the projections and reconstructed slices. Note that the combination of denoised and noisy images at half-dose (30 mAs) yielded images with similar bias and noise variance as the standard full-dose image. At lower dose levels (15 mAs, 9 mAs, 3 mAs), the approximations



(a) 30 mAs Noisy (b) 30 mAs Denoised (c) 30 mAs Combined (d) 60 mAs



(e) 9 mAs Noisy (f) 9 mAs Denoised (g) 9 mAs Combined (h) Ground-truth

Figure 6: Magnified region of interest taken from the in-plane slice (after reconstruction). For completeness, we also show the ROI from the reconstructed ground-truth slice, although it plays no role for the visual validation: the goal is to approximate the 60 mAs full-dose image with the combined low-dose images.

become coarse and the differences between expected and measured errors become more different.

Note that, as desired, the MNSE reports the same value as the sum between the bias squared and the variance. Furthermore, the overall errors increase as the radiation dose decrease. Individual analysis at each dose level show that the bias squared increases with the denoising process, while the variance decreases and achieve values close to zero. The combination between noisy and denoised images yield images with lower bias and higher variance compared to the denoised images. The goal of this work is to obtain combined images with errors of bias and variance comparable to the standard full-dose image, both on the projections and reconstructed slices.

The resolvability of singularities was analyzed using the FWHM of a microcalcification, as shown in Table 2. At 50% (30 mAs) dose, slices reconstructed from any of the three sets of projections feature FWHM values comparable to that achieved at the standard full dose (60 mAs). This indicates that the microcalcification was preserved during the entire pipeline. At further lower doses, the FWHM values progressively exhibit a significant deviation from the standard full dose (60 mAs). It is interesting to observe that the denoising *per se* does not degrade the FWHM while making the fitting more precise.

7 Discussion

Analysis of the noisy data prior to the restoration process shows that acquisitions at lower radiation levels indeed yielded images with lower quality. The differences can be visually appreciated in Figure 5 and Figure 6. The results presented in Table 1 and Table 2 also support this assertion. A simple analysis of the noise variance on the projections, presented in Table 1, shows that the errors due to noise approximately double when the radiation dose is decreased to half, which highlights that Poisson noise is a dominant degradation.

Table 1: Mean Normalized Square Error (MNSE) measured from projections (*i.e.* before DBT reconstruction) and slices (after DBT reconstruction). The MNSE is decomposed into squared bias and variance.

Projections						
mAs	**Metric**	**Noisy**	**Denoised**	**Combined**	**Standard Full Dose (60 mAs)**	
30	**Bias^2**	1.99%	3.69%	2.15%	**Bias^2**	1.97%
	Variance	40.72%	1.27%	21.16%		
	MNSE	42.72% ± 0.29%	4.96% ± 0.02%	23.30% ± 0.14%		
15	**Bias^2**	2.17%	4.14%	2.72%		
	Variance	83.93%	2.15%	22.58%	**Variance**	19.01%
	MNSE	86.10% ± 0.37%	6.29% ± 0.06%	25.31% ± 0.14%		
9	**Bias^2**	2.28%	4.47%	3.22%		
	Variance	147.08%	2.92%	23.58%		
	MNSE	149.36% ± 0.35%	7.39% ± 0.09%	26.79% ± 0.17%	**MNSE**	20.98% ± 0.11%
3	**Bias^2**	3.65%	7.05%	5.99%		
	Variance	488.55%	8.68%	29.33%		
	MNSE	492.20% ± 0.97%	15.74% ± 0.22%	35.32% ± 0.24%		

Slices						
mAs	**Metric**	**Noisy**	**Denoised**	**Combined**	**Standard Full Dose (60 mAs)**	
30	**Bias^2**	0.18%	0.49%	0.17%	**Bias^2**	0.12%
	Variance	1.89%	0.50%	1.27%		
	MNSE	2.07% ± 0.03%	0.99% ± 0.02%	1.44% ± 0.02%		
15	**Bias^2**	0.27%	0.72%	0.35%		
	Variance	3.77%	0.88%	1.76%	**Variance**	0.94%
	MNSE	4.04% ± 0.04%	1.60% ± 0.03%	2.11% ± 0.03%		
9	**Bias^2**	35.35%	1.02%	0.58%		
	Variance	2.73%	1.27%	2.25%		
	MNSE	38.09% ± 0.36%	2.30% ± 0.02%	2.83% ± 0.04%	**MNSE**	1.06% ± 0.09%
3	**Bias^2**	34.21%	2.41%	1.92%		
	Variance	8.73%	3.32%	4.38%		
	MNSE	42.94% ± 0.50%	5.73% ± 0.05%	6.30% ± 0.06%		

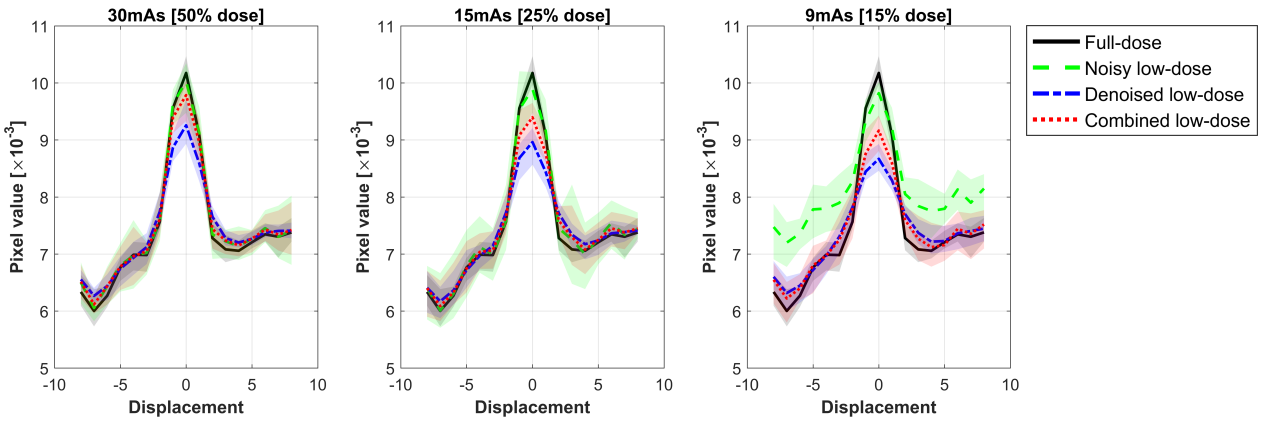


Figure 7: Cross-sections of the top calcification shown in Figure 6. The shaded areas show the min-max envelope over the ten realizations. The cross-sections were centered at the maximum of the microcalcification within the ground-truth in-plane slice; displacement is reported in pixels.

Table 2: Full width at half maximum (FWHM) calculated at the top calcification shown in Figure 6, at the in-plane slice.

Full width at half maximum (FWHM) (mm ± std)				
mAs	**Noisy**	**Denoised**	**Combined**	**Standard Full Dose (60 mAs)**
30	0.51 ± 0.07	0.55 ± 0.02	0.52 ± 0.06	0.54 ± 0.04
15	0.59 ± 0.07	0.60 ± 0.05	0.59 ± 0.05	
9	0.66 ± 0.16	0.60 ± 0.05	0.59 ± 0.06	
3	0.73 ± 0.26	0.59 ± 0.08	0.61 ± 0.13	

Next, the analysis of data obtained after denoising shows that it was able to remove virtually all noise. This is supported by visual analysis of the second column of ROIs shown in Figure 5 and Figure 6. The variance errors presented in Table 1 also indicate that the denoising was efficient at suppressing the noise, as the errors due to noise variance decrease drastically, reaching values close to zero.

Another important aspect related to the denoising is the larger bias when compared to the unfiltered data. The bias is mostly caused by the smoothing and smearing introduced by the denoising process. As the dose is decreased, the denoising becomes more aggressive and the bias increases, as observed throughout the experiments.

The combination between noisy and denoised projections resulted in images with noise statistics very similar to a standard full-dose acquisition. The similarity can be visually appreciated in Figure 5 and Figure 6. The FWHM for the combined images at 50% (30 mAs) of the dose reaches values close to that of the standard full-dose acquisition, indicating that the microcalcification was preserved during the restoration process. However, at doses lower than 50%, the restoration pipeline did not yield FWHM values comparable to that obtained at standard full dose. The objective measurements reported in Table 1 also supports that it is possible to approximate a standard full-dose DBT by denoising and processing a half-dose measurement. The experiments on lower-dose acquisitions show the limitations of the proposed processing pipeline. Combined projections at 25%, 15% and 5% dose yielded considerably higher signal and noise errors than expected for a standard full-dose acquisition. This can be seen from the differences in MNSE, bias, variance, and FWHM. At those dose ranges, even though the denoising manages to recover part of the signal, important features are not resolved, making the images unsuitable for clinical use.

8 Conclusions and future work

Based on a comprehensive characterization of the noise statistics of DBT projections, in this work we proposed a processing pipeline meant to approximate a standard full-dose (60 mAs) DBT starting from a lower-dose acquisition. The pipeline includes a multiframe denoising step for suppressing the noise in the noisy lower-dose projections, and a weighted averaging step where the noisy and denoised projections are combined to generate an image with the target expectation and noise variance. The combined signal is then fed into a reconstruction algorithm, here used as a black box, that generates slices of the reconstructed DBT.

We applied the proposed restoration pipeline to DBT projections acquired at 4 different reduced dose levels (50%, 25%, 15%, and 5%), aiming at approximating a standard full-dose measurement. Two different metrics have been used to objectively assess the fidelity of these approximations. All experiments have been performed using real data acquired from a clinical DBT system. While we are able to generate a good approximation of the standard full-dose images from the 50% dose set, it is clear that approximations from acquisitions at doses 25% or lower bear no clinical value due to the loss of details in the reconstructed slices.

While we do not claim that the dose for clinical DBT examinations could be reduced to 50% without affecting the detection and characterization of lesions (this would require a separate study), in this work the results show that no significant differences can be measured between standard full-dose and restored acquisitions at 50% dose.

Although the system used in our experiments features a detector with minimal pixel crosstalk, this issue may be more prominent with other systems. Thus, it is worth emphasizing that the proposed pipeline is based on pixelwise observation models, a pixelwise VST, and pixelwise weighted averaging, which are directly compatible with data subject to crosstalk [34, 14]. The pipeline's modular design (see Figure 1) allows to replace the AWGN denoising with filters for stationary non-white noise; notably, RF3D itself can cope with noise models more general than AWGN, including spatially (cross-talk) and temporally (fixed-pattern) correlated noise [27].

Future works include the addition of an image sharpening stage combined with the denoising. This could cope with the loss of high-frequency details which limited the performance of the restoration pipeline at 25% or lower doses. Furthermore, subjective reading tests should be performed with radiologists, to investigate the impact of the restoration pipeline on the detection and characterization of lesions.

MATLAB implementations of the RF3D denoising algorithm, and of the forward and inverse VSTs can be downloaded through the links provided by references [27] and [20], respectively.

Acknowledgements

This work was supported by the São Paulo Research Foundation (FAPESP grants 2013/18915-5 and 2016/25750-0), the Brazilian Foundation for the Coordination of Improvement of Higher Education Personnel (CAPES grant 88881.030443/ 2013-01), the Burroughs Wellcome Fund (IRSA 1016451), the Komen Foundation (grant IIRI326610), the National Institutes of Health and National Cancer Institute (grant 1R01CA154444), the Academy of Finland (project 310779), and the European Commission (FP7-PEOPLE-ITN-2013-607290).

The authors would like to thank Real Time Tomography for providing the reconstruction software, and Kristen Lau and Bruno Barufaldi for their support during the acquisition of images.

References

[1] International Atomic Energy Agency, "Appendix II: Medical exposure," in *International Basic Safety Standards for Protection against Ionizing Radiation and for the Safety of Radiation Sources*, 1996, pp. 45–56.

[2] A. B. de Gonzalez, C. D. Berg, K. Visvanathan, and M. Robson, "Estimated Risk of Radiation-Induced Breast Cancer From Mammographic Screening for Young BRCA Mutation Carriers," *Journal of the National Cancer Institute*, vol. 101, no. 3, pp. 205–209, 2009.

[3] M. J. Yaffe and J. G. Mainprize, "Risk of Radiation-induced Breast Cancer from Mammographic Screening," *Radiology*, vol. 258, no. 1, pp. 98–105, 2011.

[4] A. Haus and M. Yaffe, "Screen-film and digital mammography. Image quality and radiation dose considerations," *Radiologic Clinics of North America*, vol. 38, no. 4, pp. 871–898, July 2000.

[5] W. Huda, A. M. Sajewicz, K. M. Ogden, and D. R. Dance, "Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system," *Medical Physics*, vol. 30, no. 3, p. 442, 2003.

[6] R. S. Saunders, J. A. Baker, D. M. Delong, J. P. Johnson, and S. Ehsan, "Does image quality matter? Impact of resolution and noise on mammographic task performance," *Medical Physics*, vol. 34, no. 10, pp. 3971–3981, 2007.

[7] M. K. Kalra, M. M. Maher, D. V. Sahani, M. A. Blake, P. F. Hahn, G. B. Avinash, T. L. Toth, E. Halpern, and S. Saini, "Low-dose CT of the abdomen: evaluation of image improvement with use of noise reduction filters—pilot study," *Radiology*, vol. 228, no. 1, pp. 251–256, 2003.

[8] A. Manduca, L. Yu, J. D. Trzasko, N. Khaylova, J. M. Kofler, C. M. McCollough, and J. G. Fletcher, "Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT," *Medical Physics*, vol. 36, no. 11, pp. 4911–4919, 2009.

[9] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Transactions on Medical Imaging*, vol. 11, no. 2, pp. 221–232, Jun 1992.

[10] M. Lysaker, A. Lundervold, and X.-C. Tai, "Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1579–1590, Dec 2003.

[11] P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 425–441, April 2008.

[12] G. Wu, J. G. Mainprize, and M. J. Yaffe, "Dose reduction for digital breast tomosynthesis by patch-based denoising in reconstruction," in *International Workshop on Digital Mammography*. Springer, 2012, pp. 721–728.

[13] M. A. d. C. Vieira, P. R. Bakic, and A. D. A. Maidment, "Effect of denoising on the quality of reconstructed images in digital breast tomosynthesis," *Proceedings SPIE Medical Imaging*, vol. 8668, p. 86680C, 2013.

[14] L. R. Borges, I. Guerrero, P. R. Bakic, A. Foi, A. D. A. Maidment, and M. A. C. Vieira, "Method for simulating dose reduction in digital breast tomosynthesis," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2331–2342, Nov 2017.

[15] L. R. Borges, P. R. Bakic, A. Foi, A. D. A. Maidment, and M. A. C. Vieira, "Pipeline for effective denoising of digital mammography and digital breast tomosynthesis," *Proceedings SPIE Medical Imaging*, vol. 10132, p. 1013206, 2017.

[16] S. Young, P. Bakic, K. J. Myers, and S. Park, "Performance tradeoffs in a model breast tomosynthesis system," in *Imaging and Applied Optics Congress*. Optical Society of America, 2010, p. DTuA3. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=DIPA-2010-DTuA3

[17] J. M. Boone, "X-ray production, interaction, and detection in diagnostic imaging," in *Handbook of Medical Imaging: Physics and Psychophysics*, R. Van Metter, J. Beutel, and H. Kundel, Eds. SPIE, 2000, pp. 1–78.

[18] I. A. Cunningham, "Applied linear-system theory," in *Handbook of Medical Imaging: Physics and Psychophysics*, R. Van Metter, J. Beutel, and H. Kundel, Eds. SPIE, 2000, pp. 79–159.

[19] M. Mäkitälo and A. Foi, "Optimal Inversion of the Anscombe Transformation in Low-Count Poisson Image Denoising," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 99–109, Jan 2011.

[20] M. Mäkitalo and A. Foi, "Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 91–103, 2013. [Online]. Available: http://www.cs.tut.fi/~foi/invansc

[21] Z. Xu, U. Bagci, J. Seidel, D. Thomasson, J. Solomon, and D. J. Mollura, "Segmentation based denoising of PET images: An iterative approach via regional means and affinity propagation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2014, pp. 698–705.

[22] T. Bölke, L. Krapf, R. Orzekowsky-Schroeder, T. Vossmeyer, J. Dimitrijevic, H. Weller, A. Schüth, A. Klinger, G. Hüttmann, and A. Gebert, "Data-adaptive image-denoising for detecting and quantifying nanoparticle entry in mucosal tissues through intravital 2-photon microscopy," *Beilstein Journal of Nanotechnology*, vol. 5, pp. 2016–2025, 2014.

[23] S. Yang and B.-U. Lee, "Poisson-Gaussian noise reduction using the hidden Markov model in contourlet domain for fluorescence microscopy images," *PLOS ONE*, vol. 10, no. 9, pp. 1–19, 2015.

[24] S. Harizanov, J. de Dios Pont, S. Ståhl, and D. Wenzel, "Noise removal and feature extraction of 2D CT radiographic images," in *Advanced Computing in Industrial Mathematics*. Springer, 2018, pp. 57–70.

[25] L. Azzari and A. Foi, "Variance Stabilization for Noisy+Estimate Combination in Iterative Poisson Denoising," *IEEE Signal Processing Letters*, vol. 23, no. 8, pp. 1086–1090, 2016.

[26] J. L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press, 1998, pp. 263–265, Appendix A1. [Online]. Available: http://www.multiresolution.com/cupbook.pdf

[27] M. Maggioni, E. Sánchez-Monge, and A. Foi, "Joint Removal of Random and Fixed-Pattern Noise Through Spatiotemporal Video Filtering," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4282–4296, Oct 2014. [Online]. Available: http://www.cs.tut.fi/~foi/GCF-BM3D

[28] J. Kuo, P. A. Ringer, S. G. Fallows, P. R. Bakic, A. D. Maidment, and S. Ng, "Dynamic reconstruction and rendering of 3D tomosynthesis images," in *Proceedings of SPIE*, vol. 7961, no. 796116-1, 2011, pp. 796 116–11.

[29] A.-K. Carton, P. Bakic, C. Ullberg, H. Derand, and A. D. Maidment, "Development of a physical 3D anthropomorphic breast phantom," *Medical Physics*, vol. 38, no. 2, pp. 891–896, 2011.

[30] L. Cockmartin, P. R. Bakic, H. Bosmans, A. D. Maidment, H. Gall, M. Zerhouni, and N. W. Marshall, "Power spectrum analysis of an anthropomorphic breast phantom compared to patient data in 2D digital mammography and breast tomosynthesis," in *International Workshop on Digital Mammography*. Springer, 2014, pp. 423–429.

[31] G. Boracchi and A. Foi, "Uniform motion blur in Poissonian noise: Blur/noise tradeoff," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 592–598, 2011.

[32] Y. Lu, H.-P. Chan, J. Wei, M. Goodsitt, P. L. Carson, L. Hadjiiski, A. Schmitz, J. W. Eberhard, and B. E. Claus, "Image quality of microcalcifications in digital breast tomosynthesis: Effects of projection-view distributions," *Medical Physics*, vol. 38, no. 10, pp. 5703–5712, 2011.

[33] N. Marshall, "Calculation of quantitative image quality parameters: Notes describing the use of OBJ_IQ_reduced," *NHSBSP Equipment Report 0902*, 2009.

[34] L. Azzari and A. Foi, "Variance stabilization in poisson image deblurring," in *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 728–731.