Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Evaluation of denoising digital breast tomosynthesis data in both projection and image domains and a study of noise model on digital breast tomosynthesis image domain

Daniele Cristina Scarparo
Denis Henrique Pinheiro Salvadeo
Daniel Carlos Guimarães Pedronette
Bruno Barufaldi
Andrew Douglas Arnold Maidment

SPIE.

Daniele Cristina Scarparo, Denis Henrique Pinheiro Salvadeo, Daniel Carlos Guimarães Pedronette, Bruno Barufaldi, Andrew Douglas Arnold Maidment, "Evaluation of denoising digital breast tomosynthesis data in both projection and image domains and a study of noise model on digital breast tomosynthesis image domain," *J. Med. Imag.* **6**(3), 031410 (2019), doi: 10.1117/1.JMI.6.3.031410.

Evaluation of denoising digital breast tomosynthesis data in both projection and image domains and a study of noise model on digital breast tomosynthesis image domain

Daniele Cristina Scarparo,[a] Denis Henrique Pinheiro Salvadeo,[a,b,*] Daniel Carlos Guimarães Pedronette,[a] Bruno Barufaldi,[b] and Andrew Douglas Arnold Maidment[b]
[a]São Paulo State University (Unesp), Institute of Geosciences and Exact Sciences, Rio Claro, São Paulo, Brazil
[b]University of Pennsylvania, Hospital of the University of Pennsylvania, Department of Radiology, Philadelphia, Pennsylvania, United States

Abstract. Digital breast tomosynthesis (DBT) is an imaging technique created to visualize 3-D mammary structures for the purpose of diagnosing breast cancer. This imaging technique is based on the principle of computed tomography. Due to the use of a dangerous ionizing radiation, the "as low as reasonably achievable" (ALARA) principle should be respected, aiming at minimizing the radiation dose to obtain an adequate examination. Thus, a noise filtering method is a fundamental step to achieve the ALARA principle, as the noise level of the image increases as the radiation dose is reduced, making it difficult to analyze the image. In our work, a double denoising approach for DBT is proposed, filtering in both projection (prereconstruction) and image (postreconstruction) domains. First, in the prefiltering step, methods were used for filtering the Poisson noise. To reconstruct the DBT projections, we used the filtered backprojection algorithm. Then, in the postfiltering step, methods were used for filtering Gaussian noise. Experiments were performed on simulated data generated by open virtual clinical trials (OpenVCT) software and on a physical phantom, using several combinations of methods in each domain. Our results showed that double filtering (i.e., in both domains) is not superior to filtering in projection domain only. By investigating the possible reason to explain these results, it was found that the noise model in DBT image domain could be better modeled by a Burr distribution than a Gaussian distribution. Finally, this important contribution can open a research direction in the DBT denoising problem. © *2019 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JMI.6.3.031410]

Keywords: digital breast tomosynthesis; double denoising; Gaussian noise; Poisson noise; noise model; Burr distribution.

Paper 18222SSR received Oct. 1, 2018; accepted for publication Jan. 29, 2019; published online Feb. 20, 2019.

1 Introduction

Computed tomography (CT) was developed to obtain a more accurate diagnosis through internal images by the transmission of γ- or x-rays, reducing dramatically the use of surgery or incision (invasive method) in the patient's body for diagnosis. In tomography, the idea is to obtain visual "slices" of a body. For that, projections of the body being scanned are acquired by using a radiation source and analyzing the residual radiation detected after to cross the object. In CT, the source emits radiation with a certain intensity and at different angles around the body, acquiring noisy projections. To obtain the desired image, a set of projections is used in a tomographic reconstruction stage. In digital breast tomosynthesis (DBT),[1] the principle is practically the same as in CT, but the projections are acquired at limited angles and with low radiation dose, by definition.

However, the noise level becomes higher when the radiation dose is reduced. This noise in acquired projections is commonly modeled by Poisson (quantum noise)[2] or Poisson–Gaussian (quantum and electronic noise)[3] distributions.

Thus, methods to smooth noise in DBT are essential in the case of DBT, that is usually done implicitly in an iterative optimization process by a reconstruction algorithm, such as simultaneous algebraic reconstruction technique[2] and total variation (TV).[4] In addition, the traditional method for DBT reconstruction is the filtered backprojection (FBP) algorithm,[2] which performs a filtering step on the projections before backprojection step by using a high-pass filter (ramp filter) plus a windowing technique (e.g., Hann filter) to attenuate a general high frequency noise. It should be noted that using only a ramp filter will result in a higher noise level.

In turn, only a few denoising methods have been proposed specifically to deal with Poisson or Poisson–Gaussian noise in projections or for image domain in DBT, defining an explicit step for filtering before or after reconstruction. Basically, we can highlight the following methods in literature, mainly to denoise DBT projections: pointwise Wiener filter (PWF) in an Anscombe domain with FBP reconstruction,[5] and one of its extensions, by evaluating a window adaptive PWF;[6] a nonlocal version of the bilateral filter with reconstruction given by a maximum likelihood convex algorithm;[7] block matching and 3-D filtering (BM3D) method in Anscombe domain,[3] or directly in projections domain[8] (both with FBP reconstruction); RF3D (a version of BM3D) in Anscombe domain with FBP reconstruction;[9] and nonlocal means (NLM) with FBP reconstruction.[8] In turn, NLM with TV reconstruction was the only approach found that was applied on the DBT image domain.[10]

However, there is a lack of studies in DBT focusing on both filtering steps. Therefore, this paper proposes to use a double

*Address all correspondence to Denis Henrique Pinheiro Salvadeo, E-mail: denis.salvadeo@unesp.br

2329-4302/2019/$25.00 © 2019 SPIE

denoising in DBT by filtering in both pre- and postreconstruction steps. A double filtering basically consists of filtering the projection data (prefiltering), backprojecting them, and then applying a filtering step (postfiltering) on the reconstructed image. This process is illustrated in Fig. 1.

The main motivation to evaluate a double denoising approach in DBT was a previous study[11] performed for CT, which shows that double denoising achieved results superior to denoising CT data before or after reconstruction only. So, it is necessary to validate this approach for DBT as well.

In addition, this work includes a study of several combinations made with some existing filtering methods for Poisson and Gaussian noise found in literature, in order to achieve a better balance between detail preservation and noise reduction through double denoising in DBT. It is worth mentioning that several filters were individually evaluated (eight for prereconstruction and six for postreconstruction steps, respectively). To the best of our knowledge, most of such filtering methods are being evaluated for the first time for DBT denoising. Moreover, considering this number of individual filters, we tested a total of 48 combinations in a double filtering approach.

In summary, the main contributions of this paper are threefold: (i) the evaluation of double filtering in DBT; (ii) a comprehensive study and evaluation of denoising methods in DBT; and (iii) a study of the noise model on DBT image domain, especially regarding the reconstructed image after a prefiltering step. Items (i) and (iii) consist of the main novelties of this work.

Finally, this paper is organized as follows. In Sec. 2, the methodology and the definitions of the methods used are presented. In Sec. 3, the experimental evaluation is described. In Sec. 4, the results are shown and discussed. Finally, in Sec. 5, the conclusions and future proposals are presented.

2 Methodology

This work focuses on the stages of denoising prereconstruction data (projection domain), denoising postreconstruction data (image domain), and their combinations (double filtering) in DBT, as previously mentioned.

To denoise projections, whose noise can be modeled by a Poisson distribution, we evaluate two approaches: without or with the use of the Anscombe transform (AT).[12] In the first case, we have filters applied directly on projections since these filters are derived supposing a Poisson distribution. In turn, the use of the AT[12] has the objective to stabilize the noise variance. Because Poisson noise is a signal-dependent one, we use the AT to transform a signal-dependent noise to an additive noise, making it approximately Gaussian, with zero mean and unitary variance. The approaches using AT are illustrated in Fig. 2.

Traditionally, in tomography, AT is applied as a first step to filter projection data before its reconstruction. Next, the data are filtered in the Anscombe domain and then the inverse AT is used so that the image is prepared to go through the reconstruction process.

For a stage of noise filtering in 2-D projections, whose projection data have not yet undergone a reconstruction process and are corrupted by Poisson noise, the following filters were used:

- directly filters (without AT): pointwise maximum *a posteriori* (MAP)[13] and Poisson nonlocal means (P-NLM);[14]
- using AT: NLM,[15] BM3D,[16] PWF,[13] and some ones of their contextual versions as generalized Wiener filter (GWF),[17] Wiener filters with isotropic (IWF),[18] and separable (SWF)[18] Markov random fields (MRFs).

Getting the filtered 2-D projections, the next step is defined by a 3-D tomographic reconstruction computed by the FBP

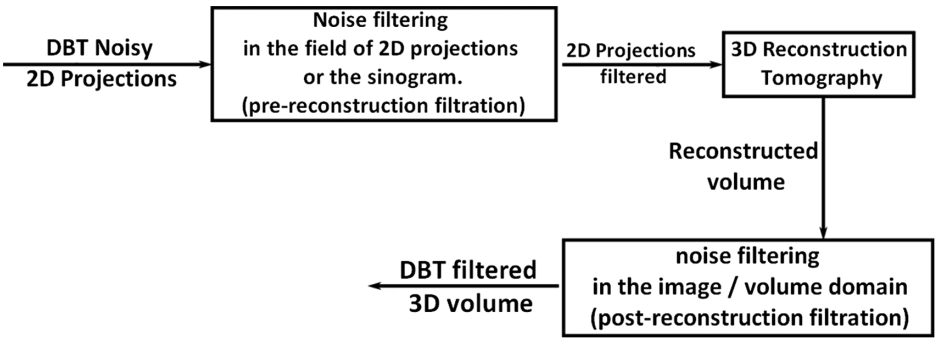

Fig. 1 Block diagram for double filtering in DBT.

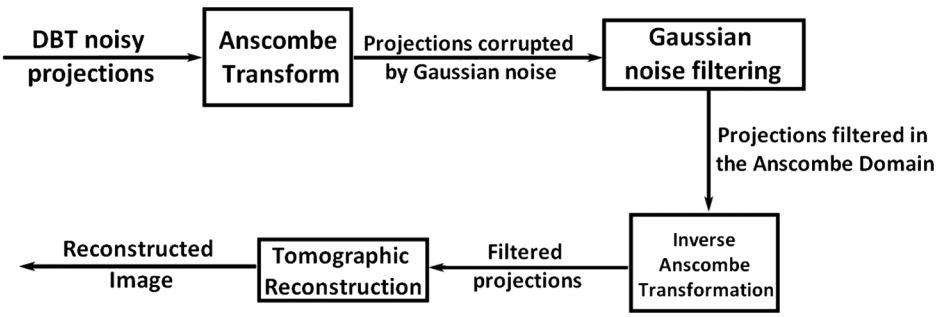

Fig. 2 Block diagram for denoising in Anscombe domain.

algorithm[19] to reach the DBT reconstructed volumes. The main reason to use this algorithm is due to its low computational cost in comparison with iterative reconstruction methods, aiming to identify a quick, suitable tool to a full DBT reconstruction (filtering + reconstruction + filtering).

The reconstructed images are still corrupted by noise, whose distribution can be approximated by a Gaussian (by invoking the central limit theorem[20]) and, therefore, a 3-D noise filtering step is required. For this step, we use the following filters suitable for Gaussian noise:

- regarding information of each reconstructed slice separately (as a 2-D image): BM3D,[16] NLM,[15,21] GWF,[17] IWF,[18] and SWF[18] MRFs;[22]
- regarding information in volumetric data by using two or more slices (as a 3-D image): 3-D version of PWF.[23]

So, we present in the next subsections, a brief description of the filters used in this work.

2.1 *Pointwise Wiener Filter*

Defined as the linear minimum mean square error[24] estimate of a desired signal, through a noisy signal, the Wiener filter is one of the more traditional methods for filtering images.

The definition in the 2-D case is based on the following equation:

$$g(x,y) = f(x,y) + v(x,y), \tag{1}$$

where $g(x,y)$ is a noisy observed signal, $f(x,y)$ is the noise-free signal to be estimated, and $v(x,y)$ describes the noise.

The Wiener filter equation of the proposal of Ref. 24 in the pointwise version is presented below:

$$\hat{f}(i,j) = \mu_f(i,j) + \frac{\sigma_f^2(i,j)}{\sigma_f^2(i,j) + \sigma_v^2(i,j)}(g(i,j) - \mu_g(i,j)), \tag{2}$$

where $\mu_f(i,j)$ and $\mu_g(i,j)$ are the local means of the original and noisy images, in this order, and $\sigma_f^2(i,j)$ and $\sigma_v^2(i,j)$ denote the local variances of the original and noisy images, respectively.

This filter was proposed to denoise an additive Gaussian noise. So, in the stage of filtering projections, it was applied in the Anscombe domain. The calculation of the local variance and mean statistics is performed in a 2-D window for prefiltering step and in a 3-D window for postfiltering.

2.2 *Generalized Wiener Filter*

The Wiener filter with Fisher information,[17] also called GWF, is a derivation of the PWF method previously mentioned, where Fisher information represents the amount of information that a certain observed random variable has on the parameter to be estimated, as shown in Ref. 17. The equation that describes this filter is as follows:[17]

$$\hat{f}(i,j) = \mu_f(i,j) + \frac{\sigma_f^2(i,j)}{\sigma_f^2(i,j) + \sigma_v^2(i,j)}\left[\alpha(g(i,j) - \mu_g(i,j)) + (1-\alpha)\sum_{g(k,l)\in\eta(i,j)}(g(k,l) - \mu_g(i,j))\right], \tag{3}$$

where $\alpha \in [0,1]$ represents a compromise between contextual ($\alpha = 0$) and pointwise ($\alpha = 1$) filtering. For projection data, it was applied in Anscombe domain.

2.3 *Wiener Filters with Isotropic and Separable Markov Random Fields*

Considering an image in lexicographic notation, with the estimate $\hat{f}$ of the original image of N pixels defined as a linear combination of g, being:[18]

$$\hat{f} = \sum_{n=0}^{N-1} \alpha_n g[n], \tag{4}$$

where $g[n]$ is a pixel of the noisy image and α_n are the weights that minimize $\epsilon = f - \hat{f}$. Thus, we expect to minimize the expected value of $\|\epsilon\|^2$ based on the orthogonality principle.[18]

Thus, we obtain a system with N equations and N unknowns, which admits a unique solution, described by[18]

$$R_{gg}\alpha = R_{fg}. \tag{5}$$

Moreover, it is important to mention that R_{fg} is a vector that indicates the autocorrelation between the central (current) pixel of the original image f with all other values observed in a window obtained from f and that each row of the matrix R_{gg} represents the autocorrelation between each noisy pixel g and another point of the noisy window.

According to this process, the difference between contextual Wiener with SWF and IWF MRF is in the definition of autocorrelation matrices.

For the Wiener version with SWF MRF,[18] the autocorrelation matrices of each central pixel (i,j) of a $W \times W$ window are defined by

$$R_{gg} = \begin{cases} \sigma_f^2(i,j) + \sigma_v^2(i,j), & \text{main diagonal} \\ \sigma_f^2(i,j)\rho_V^{|i'-i''|}\rho_H^{|j'-j''|}, & \text{remainder} \end{cases}, \tag{6}$$

$$R_{ff} = (\sigma_f^2(i,j)\rho_V^{|i'-i|}\rho_H^{|j'-j|}), \tag{7}$$

where ρ_V and ρ_H are the vertical and horizontal correlation coefficients, respectively, with values of 0.95 each and (i',j') and (i'',j'') are pixel positions in a window.

The Wiener version with IWF MRF[18] presents the following definition for the autocorrelation matrices:

$$R_{gg} = \begin{cases} \sigma_f^2(i,j) + \sigma_v^2(i,j), & \text{main diagonal} \\ \sigma_f^2(i,j)\rho^{\sqrt{(i'-i'')^2+(j'-j'')^2}}, & \text{remainder} \end{cases}, \tag{8}$$

$$R_{ff} = \left[\sigma_f^2(i,j)\rho^{\sqrt{(i'-i)^2+(j'-j)^2}}\right], \tag{9}$$

where ρ is the correlation coefficient with a value of 0.95.

Finally, IWF and SWF were applied in the Anscombe domain for prefiltering stage.

2.4 *Pointwise Maximum* a Posteriori

In order to obtain a point estimator of the signal and when we know their statistics, the MAP[13] estimation can be used. In this work, we use a version of this filter derived from a Poisson and

Gaussian densities modeling likelihood and an a priori knowledge, respectively. The MAP estimator is given by the following equation:[13]

$$\hat{g} = \frac{\mu - \sigma^2 \sqrt{(\sigma^2 - \mu)^2 + 4\sigma^2 y}}{2}, \tag{10}$$

where μ, σ^2, and $\hat{g}$ are the mean, variance, and estimated value, respectively.

Remember that this filter was used to filter data on projection domain, applied directly on the data corrupted by Poisson noise.

2.5 *Nonlocal Means*

Based on the redundancy of regions/patches (may have different forms) in an image, the NLM method was created by Ref. 21. Its filtering parameters are a patch size (P) and a search window size (W), whose values in our applications were 3×3 and 11×11 sizes, respectively, as well as a parameter h to handle the filter smoothing.

To obtain the estimated value of the noise-free pixel, NLM calculates a weighted average of the noisy pixels in a given region. The equation of this method is represented as[14] follows:

$$\hat{u}_s = \frac{\sum_{t \in W} \omega(s,t) v_t}{\sum_{t \in W} \omega(s,t)}, \tag{11}$$

where $\hat{u}_s$ is the noise-free estimate of one pixel of the image, W represents the search window, v_t is a noisy pixel in that window, and $\omega(s,t)$ represents the weights/similarity between

Table 1 PSNR results for virtual phantoms.

Methods	FS	Mean	SD	MBB	*I*
Noisy image	—	23.92	3.13	—	—
MAP	Pre	27.89	3.24	25.33	2.20
P-NLM	Pre	26.83	2.24	23.73	0.59
AT-NLM	Pre	26.94	2.24	23.82	0.68
AT-BM3D	Pre	28.68	2.14	25.15	2.01
AT-2D-PWF	Pre	**29.91**	2.51	26.31	**3.17**
AT-GWF	Pre	29.64	2.56	26.08	2.94
AT-SWF	Pre	29.80	2.71	26.29	3.15
AT-IWF	Pre	29.79	2.71	26.28	3.14
NLM	Post	26.48	3.25	29.02	0.16
3D-PWF	Post	23.92	3.13	28.86	0.00
GWF	Post	10.13	2.44	8.91	−19.95
IWF	Post	23.92	3.13	28.86	0.00
SWF	Post	23.92	3.13	28.86	0.00
BM3D	Post	**29.14**	2.32	27.74	−1.12
AT-2D-PWF + NLM	Both	**30.01**	2.31	—	—
AT-IWF + NLM	Both	29.98	2.47	—	—
AT-2D-PWF + 3D-PWF	Both	29.91	2.51	—	—
AT-GWF + NLM	Both	29.72	2.36	—	—
AT-GWF + 3D-PWF	Both	29.64	2.56	—	—
P-NLM + BM3D	Both	26.25	2.40	—	—

Note: Significance test by calculating *p*-values in Microsoft Excel. Only one method for each filtering stage (FS) was used. *p*-values should be <0.05 for a significant difference.
Based on PSNR:
1. AT-2D-PWF + NLM versus AT-2D-PWF: 0.918 (not significant).
2. AT-2D-PWF + NLM versus BM3D: 0.331 (not significant).
3. AT-2D-PWF versus BM3D: 0.406 (not significant).

Bold values represent the best results achieved for each filtering stage (FS). For column *I*, bold values represent the methods that contributed most in a double filtering scheme.

Table 2 SSIM results for virtual phantoms.

Methods	FS	Mean	SD	MBB	*I*
Noisy image	—	0.25	0.06	—	—
MAP	Pre	0.41	0.07	0.40	0.12
P-NLM	Pre	0.26	0.03	0.24	−0.04
AT-NLM	Pre	0.27	0.03	0.25	−0.03
AT-BM3D	Pre	0.40	0.06	0.36	0.08
AT-2D-PWF	Pre	**0.51**	0.06	0.46	**0.18**
AT-GWF	Pre	**0.51**	0.06	0.46	**0.18**
AT-SWF	Pre	0.50	0.06	0.46	**0.18**
AT-IWF	Pre	0.50	0.06	0.46	**0.18**
NLM	Post	0.35	0.06	0.43	0.01
3D-PWF	Post	0.25	0.06	0.42	0.00
GWF	Post	0.20	0.05	0.33	−0.09
IWF	Post	0.25	0.06	0.42	0.00
SWF	Post	0.25	0.06	0.42	0.00
BM3D	Post	**0.40**	0.05	0.31	−0.10
AT-2D-PWF + NLM	Both	**0.51**	0.06	—	—
AT-IWF + NLM	Both	**0.51**	0.06	—	—
AT-2D-PWF + 3D-PWF	Both	**0.51**	0.06	—	—
AT-GWF + NLM	Both	**0.51**	0.06	—	—
AT-GWF + 3D-PWF	Both	**0.51**	0.06	—	—
P-NLM + BM3D	Both	0.22	0.03	—	—

Note: Significance test by calculating *p*-values in Microsoft Excel. Only one method for each filtering stage (FS) was used. *p*-values should be <0.05 for a significant difference.
Based on SSIM:
1. AT-2D-PWF + NLM versus AT-2D-PWF: 0.773 (not significant).
2. AT-2D-PWF + NLM versus BM3D: 2.70×10^{-06} (significant).
3. AT-2D-PWF versus BM3D: 1.29×10^{-05} (significant).

Bold values represent the best results achieved for each filtering stage (FS). For column *I*, bold values represent the methods that contributed most in a double filtering scheme.

the patches centered in s and t. The weights of this average can be calculated based on the Euclidean distance to measure the similarity of the central patch with their neighbors. The following equation shows how to obtain these weights:[14]

$$\omega(s,t) = \exp\left(-\frac{1}{h}\sum_{k\in P}|v_{s,k} - v_{t,k}|^2\right), \quad (12)$$

where k symbolizes the k'th elements of patches v_s and v_t.

The original version of the NLM was proposed for additive white Gaussian noise, being used in the Anscombe domain to perform the filtration in the prereconstruction stage. In addition, an adapted version of the NLM for for Poisson noise (P-NLM) proposed by Ref. 14 was also used. In this method, the main difference from original NLM is that the similarity measure to obtain the weights was derived based on a Poisson distribution. This measure for P-NLM is defined by[14]

$$\omega(s,t) = \sum_{k\in P}\left[v_{s,k}\log v_{s,k} + v_{t,k}\log v_{t,k} - (v_{s,k}+v_{t,k})\log\left(\frac{v_{s,k}+v_{t,k}}{2}\right)\right]. \quad (13)$$

2.6 *Block-Matching and 3-D Filtering*

The BM3D was proposed by Ref. 16 and it is now considered the most important method for Gaussian noise filtering. The BM3D filtering approach consists of the union of wavelet thresholding, Wiener filtering, and nonlocal filtering approaches.

The method performs a collaborative filtering in a grouping model, in which mutually similar d-dimensional patches are stacked in a $(d+1)$-dimensional array and filtered together in a transformation domain.

Table 3 PSNR results for physical phantoms.

Methods	FS	50%	70%	85%	Mean	SD	MBB	*I*
Noisy image	—	25.07	26.41	27.12	26.20	1.04	—	—
MAP	Pre	28.83	30.07	30.83	29.91	1.01	28.80	2.01
P-NLM	Pre	28.25	28.82	29.26	28.78	0.51	27.19	0.41
AT-NLM	Pre	28.41	29.06	29.56	29.01	0.58	27.38	0.59
AT-BM3D	Pre	**30.58**	**31.77**	**32.48**	**31.61**	0.96	29.47	**2.69**
AT-2D-PWF	Pre	30.11	31.35	32.08	31.18	0.99	29.35	2.57
AT-GWF	Pre	29.54	30.86	31.73	30.71	1.10	28.89	2.10
AT-SWF	Pre	30.03	31.30	32.04	31.12	1.02	29.35	2.57
AT-IWF	Pre	30.08	31.33	32.06	31.16	1.00	29.39	2.61
NLM	Post	28.73	29.96	30.70	29.80	0.99	30.60	0.16
3D-PWF	Post	25.07	26.41	27.12	26.20	1.04	30.44	0.00
GWF	Post	19.98	20.77	21.24	20.66	0.64	20.30	−10.13
IWF	Post	25.07	26.41	27.12	26.20	1.04	30.44	0.00
SWF	Post	25.10	26.43	27.14	26.22	1.04	30.44	0.00
BM3D	Post	**30.77**	**31.76**	**32.30**	**31.61**	0.78	30.16	−0.28
AT-BM3D + 3D-PWF	Both	**30.58**	**31.77**	**32.48**	**31.61**	0.96	—	—
AT-IWF + NLM	Both	30.39	31.56	32.24	31.40	0.94	—	—
MAP + BM3D	Both	30.51	31.46	32.00	31.32	0.76	—	—
AT-2D-PWF + 3D-PWF	Both	30.11	31.35	32.08	31.18	0.99	—	—
AT-GWF + 3D-PWF	Both	29.54	30.86	31.73	30.71	1.10	—	—
P-NLM + BM3D	Both	27.73	28.24	28.64	28.20	0.46	—	—

Note: Significance test by calculating *p*-values in Microsoft Excel. Only one method for each filtering stage (FS) was used. *p*-values should be <0.05 for a significant difference.
Based on PSNR:
1. AT-BM3D + 3D-PWF versus AT-BM3D: 1.000 (not significant).
2. AT-BM3D + 3D-PWF versus BM3D: 0.999 (not significant).
3. AT-BM3D versus BM3D: 0.999 (not significant).

Bold values represent the best results achieved for each filtering stage (FS). For column *I*, bold values represent the methods that contributed most in a double filtering scheme.

With the transformation of the 3-D group, the local correlation between pixels in each block and the nonlocal one between the corresponding pixels of different blocks takes place at the same time.

In this way, the artifacts present in the group are very scarce, which results in a very efficient separation of noise and signal, using coefficient shrinkage. After the inverse transformation, we obtain estimates of each grouped block, which are then aggregated adaptively in their original locations.

The method is divided into two main parts. In the first one, named basic estimate, similar regions are determined using the BM algorithm. Then, wavelet thresholding is performed using hard-thresholding in the wavelet 3-D domain, thus creating several noise-free estimates for each block. Then, an aggregation (nonlocal weighted mean) of the filtered blocks is performed.

In the second one, named final estimate, the procedure is practically the same, except that instead of the wavelet thresholding, the Wiener filtering is used, considering the previous estimate as an initial estimate of the noise-free image.

Finally, it is important to remember that it is a suitable filter for Gaussian noise. Thus, it was necessary to use the AT in the prefiltering stage. Also, it is noteworthy that the BM3D was adapted to operate in the Anscombe domain by the definition of variance equal to 1.

3 Experimental Evaluation

To evaluate the use of a double denoising in DBT, we developed all filters used in this work (prereconstruction, postreconstruction, or double filtering) as well as the experiments by using the MATLAB. In addition, it is important to highlight that we used

Table 4 SSIM results for physical phantoms.

Methods	FS	50%	70%	85%	Mean	SD	MBB	*I*
Noisy Image	—	0.25	0.28	0.31	0.28	0.03	—	—
MAP	Pre	0.36	0.39	0.42	0.39	0.03	0.38	0.08
P-NLM	Pre	0.24	0.24	0.25	0.24	0.01	0.24	−0.06
AT-NLM	Pre	0.24	0.26	0.26	0.25	0.01	0.25	−0.06
AT-BM3D	Pre	0.36	0.38	0.39	0.37	0.02	0.35	0.05
AT-2D-PWF	Pre	**0.40**	**0.42**	**0.44**	**0.42**	0.02	0.39	**0.09**
AT-GWF	Pre	**0.40**	**0.42**	**0.44**	**0.42**	0.02	0.39	**0.09**
AT-SWF	Pre	0.39	**0.42**	**0.44**	**0.42**	0.02	0.39	**0.09**
AT-IWF	Pre	0.39	**0.42**	**0.44**	**0.42**	0.02	0.39	**0.09**
NLM	Post	**0.34**	**0.38**	**0.40**	**0.38**	0.03	0.36	0.00
3D-PWF	Post	0.25	0.28	0.31	0.28	0.03	0.37	0.00
GWF	Post	0.25	0.29	0.32	0.29	0.03	0.36	−0.01
IWF	Post	0.25	0.28	0.31	0.28	0.03	0.37	0.00
SWF	Post	0.25	0.28	0.31	0.28	0.03	0.37	0.00
BM3D	Post	0.29	0.31	0.31	0.31	0.01	0.26	−0.10
AT-BM3D + 3D-PWF	Both	0.36	0.38	0.39	0.37	0.02	—	—
AT-IWF + NLM	Both	**0.40**	**0.42**	0.43	0.41	0.02	—	—
MAP + BM3D	Both	0.29	0.30	0.31	0.30	0.01	—	—
AT-2D-PWF + 3D-PWF	Both	**0.40**	**0.42**	**0.44**	**0.42**	0.02	—	—
AT-GWF + 3D-PWF	Both	**0.40**	**0.42**	**0.44**	**0.42**	0.02	—	—
P-NLM + BM3D	Both	0.19	0.20	0.20	0.20	0.01	—	—

Note: Significance test by calculating *p*-values in Microsoft Excel. Only one method for each filtering stage (FS) was used. *p*-values should be <0.05 for a significant difference.
Based on SSIM:
1. AT-2D-PWF + 3D-PWF versus AT-2D-PWF: 1.000 (not significant).
2. AT-2D-PWF + 3D-PWF versus NLM: 0.104 (not significant).
3. AT-2D-PWF versus NLM: 0.104 (not significant).

Bold values represent the best results achieved for each filtering stage (FS). For column *I*, bold values represent the methods that contributed most in a double filtering scheme.

the version of FBP implemented available in an open source distribution.[25]

Our dataset has noisy and noise-free projection data from 15 phantoms generated by a virtual clinical trial software called OpenVCT,[26] developed by the University of Pennsylvania.

For each phantom, 2-D projections are acquired in 15 different angles with a total range of 15 deg. Each 2-D projection has 2048×1792 pixels and the volume has $2048 \times 1792 \times 64$ voxels, with $0.1\ \text{mm} \times 0.1\ \text{mm} \times 1.0\ \text{mm}$ per voxel. In addition, in this phantom, some lesions, such as masses or microcalcifications, were inserted by OpenVCT software.

Moreover, DBT projection data acquired from an anthropomorphic physical breast phantom using a clinical DBT machine (Selenia Dimensions, Hologic, Bedford, Massachusetts) generated by Ref. 8 were also tested. These DBT data comprise three different radiation doses: 50%, 70%, and 85% of the standard radiation dose. In addition, ground-truth is defined as an approximation given by averaging data from 10 acquisitions with standard radiation dose.

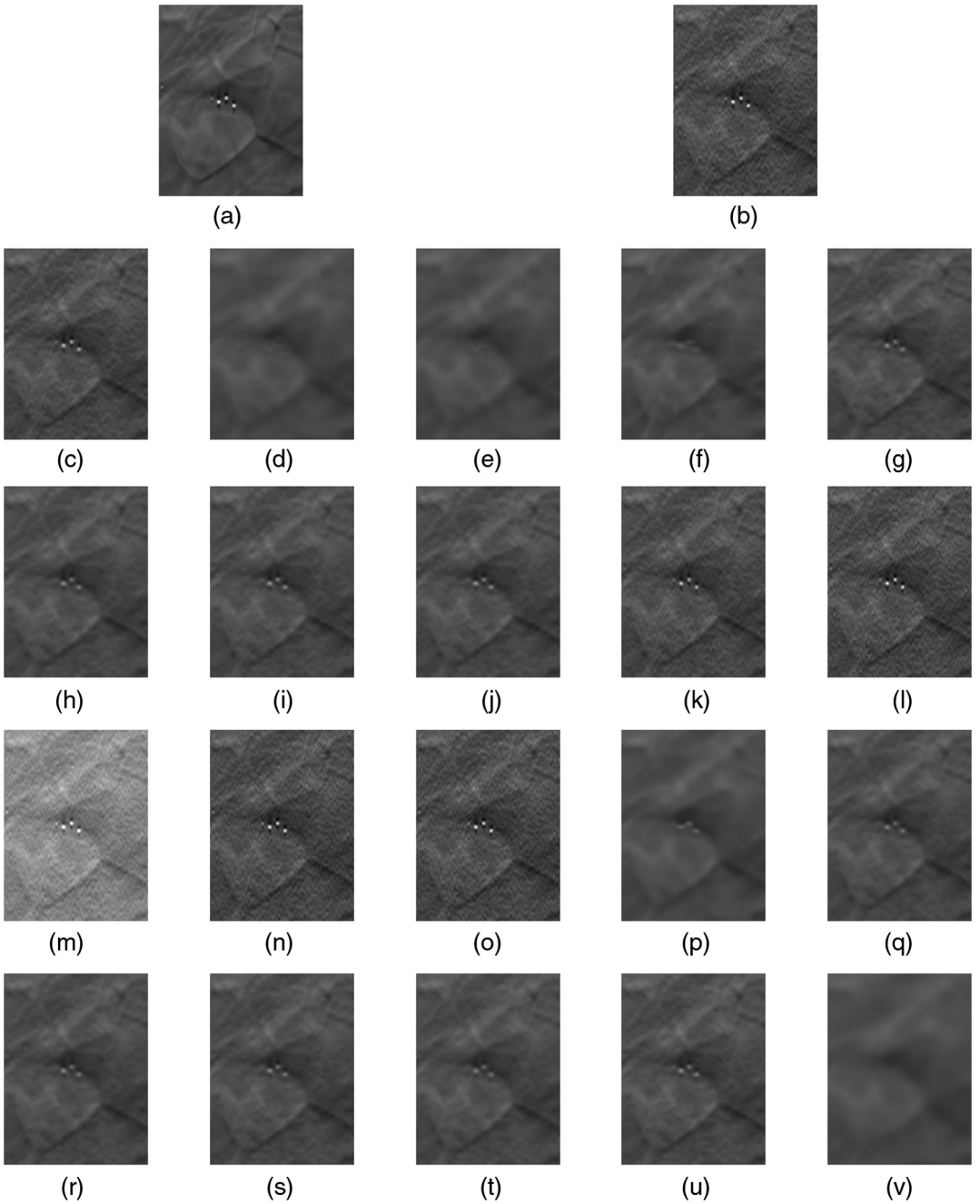

Fig. 3 Some filtering results for a RoI from a virtual phantom containing microcalcifications: (a) ideal, (b) noisy, (c) MAP, (d) P-NLM, (e) AT-NLM, (f) AT-BM3D, (g) AT-2D-PWF, (h) AT-GWF, (i) AT-IWF, (j) AT-SWF, (k) NLM, (l) 3D-PWF, (m) GWF, (n) IWF, (o) SWF, (p) BM3D, (q) AT-2D-PWF + NLM, (r) AT-IWF + NLM, (s) AT-2D-PWF + 3D-PWF, (t) AT-GWF + NLM, (u) AT-GWF + 3D-PWF, and (v) P-NLM + BM3D.

We performed quantitative and qualitative evaluations. For the quantitative evaluation, a mean of peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM)[27] was used, always calculated in regions of interest (RoIs) about final processing (after the reconstruction process or postfiltering). Comparisons were made in results achieved regarding only pre-reconstruction filters, only postreconstruction filters, and double denoising steps.

Finally, in order to obtain the noise variance that was used in the postfiltering stage for all the images used, the method proposed in Ref. 28 was employed after FBP algorithm. Remember that only NLM does not use noise variance.

4 Results

Tables 1–4 show the results of mean of PSNR and SSIM measurements calculated on RoIs from breast region, comparing ideal, noisy, and filtered reconstructed images. We recall that the measurements are calculated on the reconstructed 3-D volume and the RoIs were selected from a slice of interest, mainly

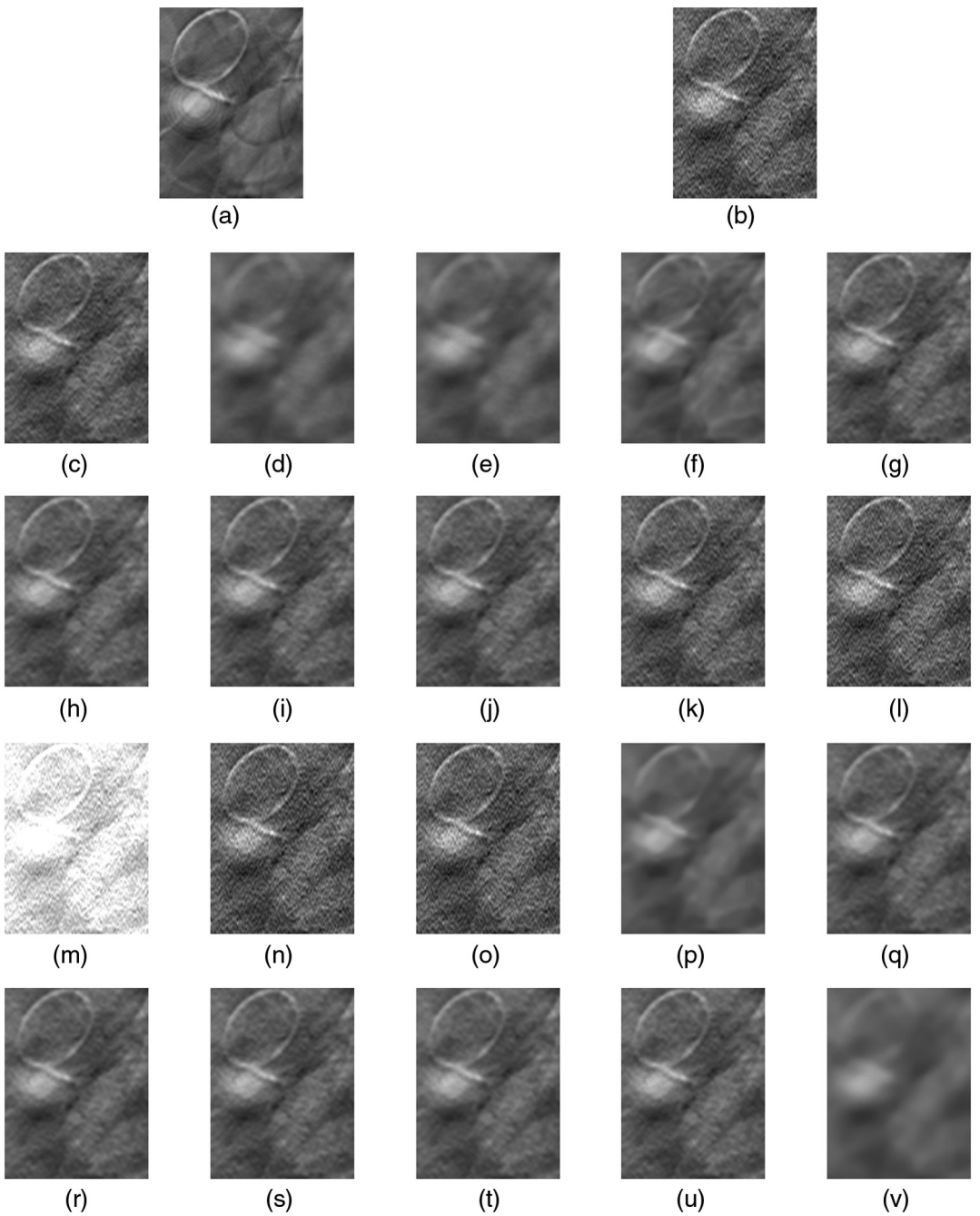

Fig. 4 Some filtering results for a RoI from a virtual phantom containing masses: (a) ideal, (b) noisy, (c) MAP, (d) P-NLM, (e) AT-NLM, (f) AT-BM3D, (g) AT-2D-PWF, (h) AT-GWF, (i) AT-IWF, (j) AT-SWF, (k) NLM, (l) 3D-PWF, (m) GWF, (n) IWF, (o) SWF, (p) BM3D, (q) AT-2D-PWF + NLM, (r) AT-IWF + NLM, (s) AT-2D-PWF + 3D-PWF, (t) AT-GWF + NLM, (u) AT-GWF + 3D-PWF, and (v) P-NLM + BM3D.

those containing some lesion. In the case of virtual phantoms, the slice of interest was the center slice.

In addition, some definitions are necessary to understand Tables 1–4, where column FS defines a filtering stage that a given method was applied, SD is a standard deviation of the PSNR and SSIM results for a given method. In turn, column MBB corresponds to a mean by base method, which are calculated by a mean of PSNR or SSIM results regarding the use of a certain method. For instance, MBB for MAP in Table 1 was averaged regarding all PSNR results, whose MAP was the filter used in projection domain in a double filtering scheme. In turn, I represents an improvement given by a current method in a double filtering scheme related to the expected value of pre- or postfiltering steps only, dependent on the stage that the respective method was applied. For instance, regarding the mean of PSNR results for all postfiltering methods (named by MPost), I for MAP in Table 1 was calculated as difference between the mean of PSNR (mean column) for MAP and MPost. For a postfiltering method, the mean column was compared with a mean of PSNR results for all prefiltering methods (MPre) to determine I value. It should be noted that negative values for I represent that a method affects negatively in a double filtering scheme by reducing the balance between noise reduction and detail preservation.

Moreover, Tables 1–4 include results for all pre- and postfiltering methods, but only the best five methods and the worse method in terms of PSNR and SSIM for a double denoising scheme. It is important to mention that 3D-PWF, IWF, and SWF show very similar results. So, only one of these methods was selected in results.

4.1 *Results for Virtual Phantoms*

As we can see in quantitative terms for virtual phantoms (see Tables 1 and 2), the AT-2D-PWF + NLM (double filtering) achieves the best overall results in terms of both SSIM and PSNR. However, several other methods achieve the same value in terms of SSIM, such as AT-IWF + NLM, AT-GWF + NLM, AT-GWF + 3D-PWF for double denoising approach and AT-2D-PWF and AT-GWF for prefiltering methods. In terms of PSNR, AT-2D-PWF + NLM was only 0.1 dB superior to AT-2D-PWF, which is insignificant. It should be noted that BM3D was the best postfiltering method, but it was 0.87 dB lower in terms of PSNR and 0.11 in terms of SSIM than

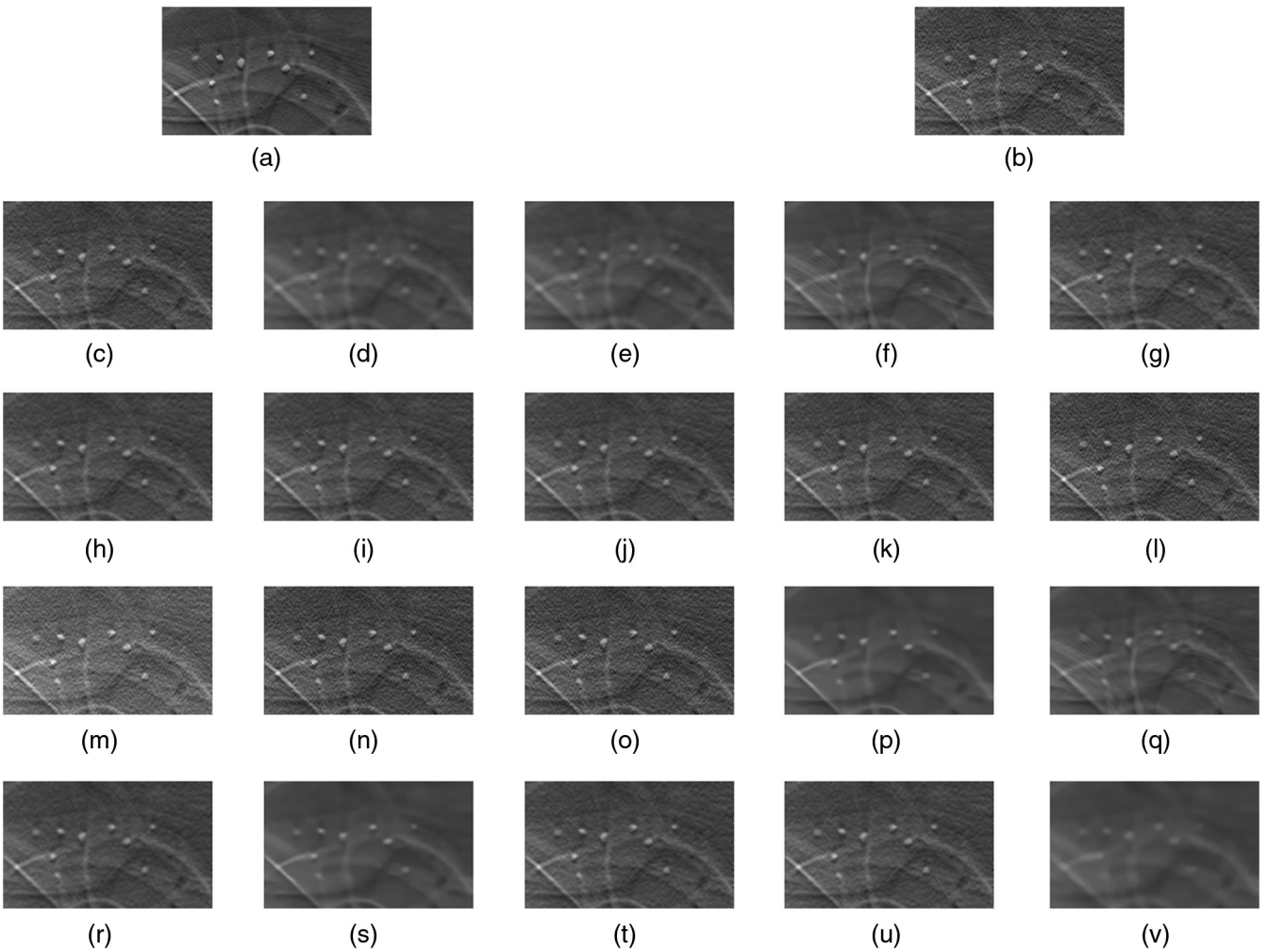

Fig. 5 Some filtering results for a RoI from a physical virtual phantom acquired with 50% of full dose: (a) ideal, (b) noisy, (c) MAP, (d) P-NLM, (e) AT-NLM, (f) AT-BM3D, (g) AT-2D-PWF, (h) AT-GWF, (i) AT-IWF, (j) AT-SWF, (k) NLM, (l) 3D-PWF, (m) GWF, (n) IWF, (o) SWF, (p) BM3D, (q) AT-BM3D + 3D-PWF, (r) AT-IWF + NLM, (s) MAP + BM3D, (t) AT-2D-PWF + 3D-PWF, (u) AT-GWF + 3D-PWF, and (v) P-NLM + BM3D.

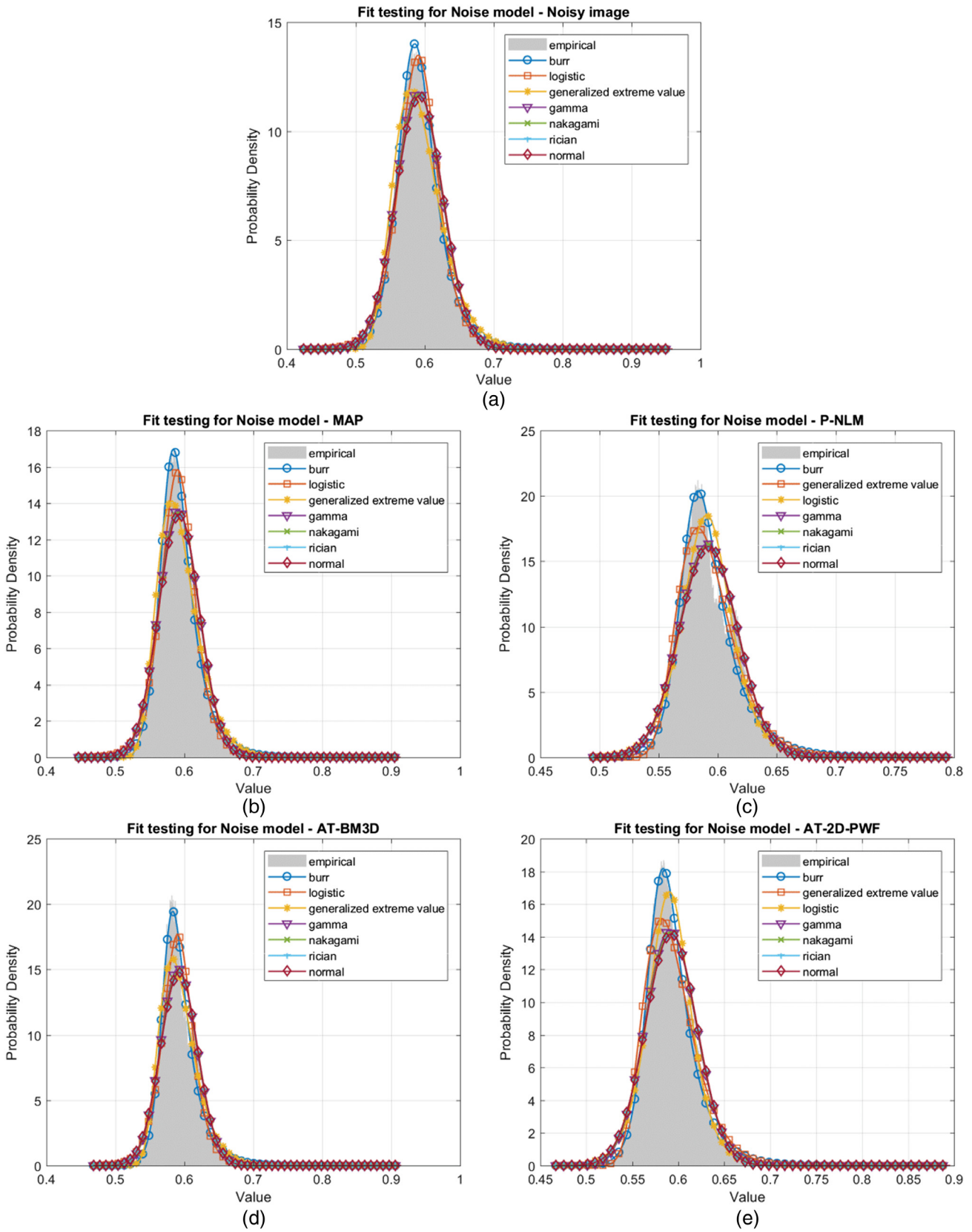

Fig. 6 Distribution fit testing for some noise models from a physical phantom data. (a) Noisy image, (b) MAP, (c) P-NLM, (d) AT-BM3D, and (e) AT-2D-PWF.

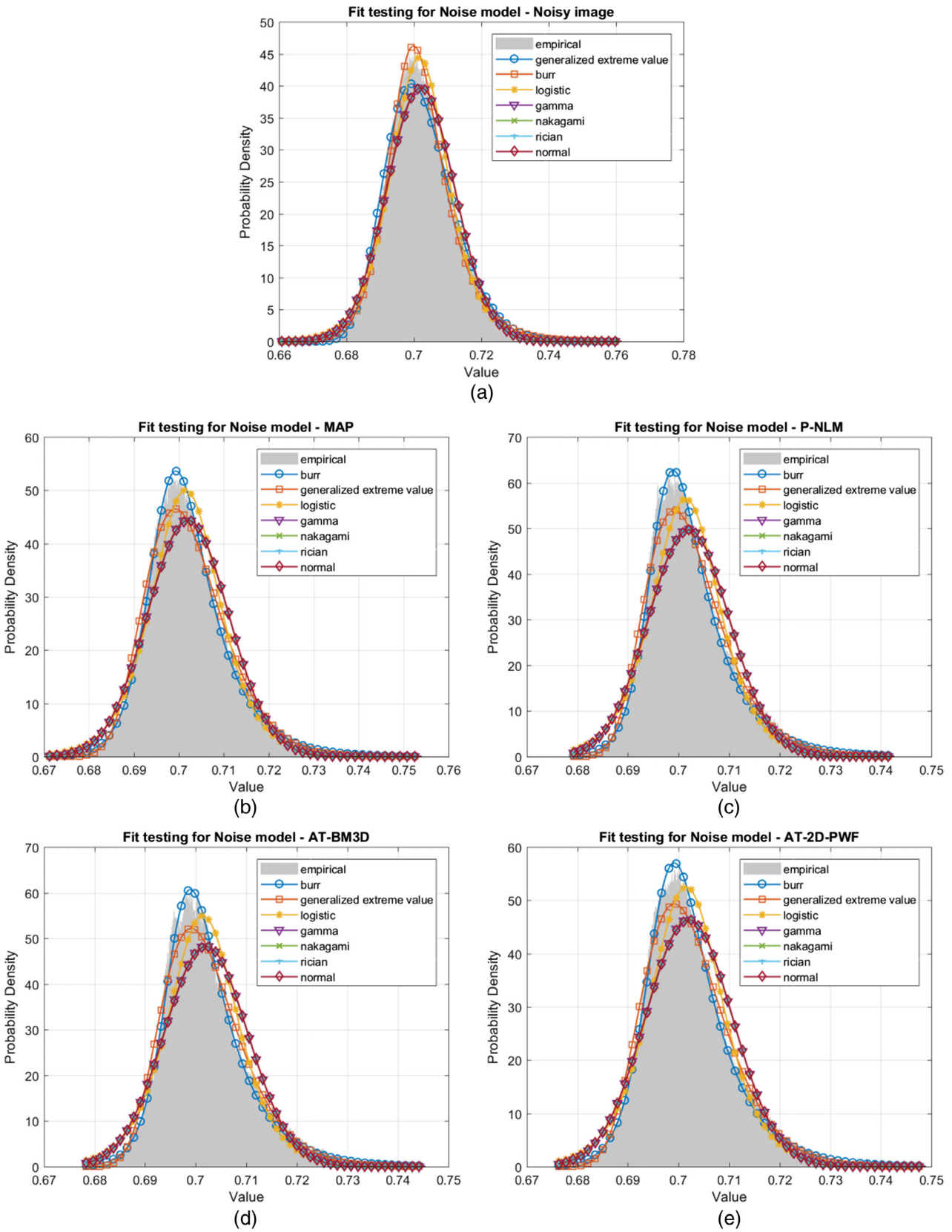

Fig. 7 Distribution fit testing for some noise models from a virtual phantom data. (a) Noisy image, (b) MAP, (c) P-NLM, (d) AT-BM3D, and (e) AT-2D-PWF.

AT-2D-PWF + NLM. In addition, it can be noted that, in general, prefiltering methods achieve better results than postfiltering methods.

In turn, in terms of improvement, AT-2D-PWF achieves the best results in both SSIM and PSNR terms (but similar to other Wiener filters applied on prefiltering step) while BM3D and mainly GWF negatively affect the double denoising results. Besides, P-NLM and AT-NLM were also not performed well.

Finally, looking at the images in Figs. 3 and 4 for a qualitative evaluation, it can be noted that P-NLM, AT-NLM, AT-BM3D, BM3D, and P-NLM + BM3D show an excessive blurring in images while GWF changes the image contrast. That can justify the performance achieved by these methods in quantitative terms. In addition, the best methods in terms of SSIM show similar visual results among them.

4.2 *Results for Physical Phantom*

Analyzing Tables 3 and 4 for the physical phantom, AT-BM3D (prefiltering method), BM3D (postfiltering method), and AT-2D-PWF + NLM (double filtering method) achieved the best results in terms of PSNR (31.61 dB), which is only 0.21 dB higher than the fourth method AT-IWF + NLM (postfiltering method). In terms of SSIM, the best results were achieved by AT-2D-PWF, AT-GWF, AT-SWF, AT-IWF for prefiltering methods and AT-GWF + NLM and AT-GWF + 3D-PWF for double filtering approach (0.42), which is 0.04 higher than the best result for a postfiltering method, which was achieved by NLM. In addition, it can be noted that, in general, prefiltering methods achieve better results than postfiltering methods.

In turn, in terms of improvement, AT-2D-PWF, AT-GWF, AT-SWF, and AT-IWF achieved the best results in terms of SSIM. In terms of PSNR, the best improvement was achieved by AT-BM3D (2.69 dB), but this improvement is about 0.1 dB higher than 2D-PWF, AT-SWF, AT-IWF, whose difference is negligible. However, mainly BM3D and GWF, in terms of SSIM and PSNR, affect negatively the double denoising results, respectively. In addition, P-NLM and AT-NLM were also not performed well.

Finally, looking at the images in Fig. 5 for a qualitative evaluation, it can be noted that P-NLM, AT-NLM, BM3D, and P-NLM + BM3D show an excessive blurring in images while GWF changes the image contrast. That can justify the performance achieved by these methods in quantitative terms. In addition, the best methods in terms of SSIM are shown similar visual results among them.

4.3 *Analysis of Noise Model in DBT Reconstructed Images*

As discussed in previous Secs. 4.1 and 4.2 about the achieved results in quantitative and qualitative terms from virtual and physical phantoms, a double denoising scheme did not show a significant improvement in comparison with prefiltering methods. In fact, prereconstruction filtering seems to define the behavior of the result. This situation does not correspond to the achieved results for CT, as shown in paper,[11] where a double denoising scheme seems to improve the results achieved by prefiltering methods only, which motivated its evaluation for DBT.

So, investigating a possible reason to justify this performance for double denoising in DBT, the following question arose: would Gaussian distribution be the most suitable noise model to describe the noise in DBT reconstructed images? It is important to note that a model for noise in DBT image domain is still an open problem in the literature.

Thus, to evaluate this question, a distribution fit testing was performed on data after FBP algorithm, regarding the following distributions: Gaussian (normal), beta, exponential, gamma, generalized extreme value, logistic, log normal, Nakagami, Rayleigh, Rician, Weibull, Burr, Birnbaum–Saunders. It evaluated reconstructed images from noisy or filtered projections (prefiltering step). Graphics represent these distribution fit tests, shown in Figs. 6 and 7, for physical and virtual phantoms, respectively. Basically, only the seven most fitted distributions are shown.

By analyzing the results, as displayed in Figs. 6 and 7, the Burr distribution[29,30] (also known as Burr type XII distribution or Singh–Maddala distribution[31]) seems to present a better fitting than Gaussian (normal) distribution, regarding reconstructed images from any prefiltering method or noisy projection, as well as in physical and virtual phantoms. This is a surprising result since Burr distribution is not usually considered as a model for image denoising problems. Finally, it should be noted that only FBP was tested as a DBT reconstruction method. That would need to be validated for other reconstruction methods.

5 Conclusion

In summary, this work presented a comprehensive study of denoising in DBT, evaluating a set of filters to denoise on projection or image domains, as well as in both domains (double filtering scheme), recalling that most of these noise filters had not yet been analyzed for DBT.

The qualitative and quantitative results achieved from virtual and physical phantoms show that the double filtering approach had no significant improvement in comparison with prefiltering methods, which is an unexpected result since a similar approach applied on CT data[11] shows superior results for the double filtering approach.

This unexpected result led us to investigate the noise model of DBT image domain after FBP algorithm from noisy or filtered projections (prefiltering). By using a distribution fit test, it was shown that a Burr distribution fit better than a Gaussian distribution on DBT image data, therefore suggesting that the noise in DBT image domain could be more suitably modeled by a Burr distribution. This is a primary contribution of this work. It is important to note that this contribution could enable the emergence of denoising methods involving a Burr noise model.

Finally, future works include the development of denoising methods for Burr noise model, validation of double denoising with postfiltering performed by these methods, evaluation of the noise in image domain regarding different DBT reconstruction algorithms, evaluation of denoising methods in terms of spatial resolution preserving, and evaluation of additional filters found in literature (e.g., based on deep neural networks).

Disclosures

We have no conflict of interests to declare.

Acknowledgments

The authors are grateful to São Paulo Research Foundation—FAPESP (Grant Nos. 2016/09714-4, 2017/17811-2, and 2017/25908-6) and Brazilian National Council for Scientific and Technological Development - CNPq (Grant No. 308194/

2017-9) for the financial support and to reviewers of this paper for the important contributions.

References

1. L. T. Niklason et al., "Digital tomosynthesis in breast imaging," *Radiology* **205**(2), 399–406 (1997).
2. A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania (1988).
3. L. R. Borges et al., "Pipeline for effective denoising of digital mammography and digital breast tomosynthesis," *Proc. SPIE* **10132**, 1013206 (2017).
4. E. Y. Sidky et al., "Image reconstruction in digital breast tomosynthesis by total variation minimization," *Proc. SPIE* **6510**, 651027 (2007).
5. M. A. C. Vieira et al., "Investigating poisson noise filtering in digital breast tomosynthesis," in *IX Workshop de Visão Computacional-WVC, IX*, Universidade Federal Fluminense (UFF) (2013).
6. M. A. C. Vieira, P. R. Bakic, and A. D. Maidment, "Effect of denoising on the quality of reconstructed images in digital breast tomosynthesis," *Proc. SPIE* **8668**, 86680C (2013).
7. G. Wu, J. G. Mainprize, and M. J. Yaffe, "Dose reduction for digital breast tomosynthesis by patch-based denoising in reconstruction," *Lect. Notes Comput. Sci.* **7361**, 721–728 (2012).
8. M. A. C. Vieira et al., "Feasibility study of dose reduction in digital breast tomosynthesis using non-local denoising algorithms," *Proc. SPIE* **9412**, 94122C (2015).
9. L. R. Borges et al., "Restoration of low-dose digital breast tomosynthesis," *Meas. Sci. Technol.* **29**(6), 064003 (2018).
10. M. Ertas et al., "An iterative tomosynthesis reconstruction using total variation combined with non-local means filtering," *BioMed. Eng. Online* **13**, 65 (2014).
11. V. C. Assis et al., "Double noise filtering in CT: pre- and post-reconstruction," in *28th SIBGRAPI Conf. Graphics, Patterns and Images*, pp. 313–320 (2015).
12. F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika* **35**, 246–254 (1948).
13. E. S. Ribeiro, "Novas propostas em filtragem de projeções tomográficas sob ruído Poisson," Master's Thesis, UFSCar, São Carlos (in portuguese) (2010).
14. C.-A. Deledalle, F. Tupin, and L. Denis, "Poisson NL means: unsupervised non local means for Poisson noise," in *17th IEEE Int. Conf. Image Process. (ICIP)*, IEEE, pp. 801–804 (2010).
15. C.-A. Deledalle, "Image denoising beyond additive Gaussian noise-patch-based estimators and their application to SAR imagery," PhD Thesis, Telecom ParisTech (2011).
16. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
17. A. L. M. Levada and N. D. A. Mascarenhas, "Filtragem adaptativa de ruído gaussiano em imagens através da minimização da informação de fisher observada," in *VI Workshop de Visão Computacional (WVC)*, pp. 7–12 (in portuguese) (2010).
18. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Inc., Upper Saddle River, New Jersey (1993).
19. L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A* **1**(6), 612–619 (1984).
20. C. Epstein, *Introduction to the Mathematics of Medical Imaging*, 2nd ed., Other Titles in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania (2008).
21. A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**, 490–530 (2005).
22. D. H. P. Salvadeo, "Filtragem de ruído em imagens tomográficas com baixa taxa de contagem utilizando uma abordagem Bayesiana contextual," PhD Thesis, UFSCar, São Carlos (in portuguese) (2013).
23. J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, Englewood Cliffs, New Jersey (1990).
24. D. T. Kuan et al., "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. Pattern Anal. Mach. Intell.* **7**(2), 165–177 (1985).
25. R. B. Vimieiro, L. R. Borges, and M. A. C. Vieira, "Open-source reconstruction toolbox for digital breast tomosynthesis," https://github.com/LAVI-USP/DBT-Reconstruction (2018).
26. B. Barufaldi et al., "OpenVCT: a GPU-accelerated virtual clinical trial pipeline for mammography and digital breast tomosynthesis," *Proc. SPIE* **10573**, 1057358 (2018).
27. Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.* **26**(1), 98–117 (2009).
28. X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising," *IEEE Trans. Image Process.* **22**, 5226–5237 (2013).
29. I. W. Burr, "Cumulative frequency functions," *Ann. Math. Stat.* **13**, 215–232 (1942).
30. P. R. Tadikamalla, "A look at the Burr and related distributions," *Int. Stat. Rev.* **48**, 337–344 (1980).
31. S. K. Singh and G. S. Maddala, "A function for size distribution of incomes," *Econometrica* **44**, 963–970 (1976).

Daniele Cristina Scarparo is currently an undergraduate student of computer science at São Paulo State University (UNESP), Rio Claro, Brazil. Her research interests include image processing, machine learning, and medical images.

Denis Henrique Pinheiro Salvadeo received his BSc, MSc, and PhD degrees in computer science from the Federal University of São Carlos (UFSCar), São Carlos, Brazil, in 2007, 2009, and 2013, respectively. He is a professor of computer science at São Paulo State University (UNESP), Rio Claro, Brazil, and nowadays, he is a visiting researcher at the University of Pennsylvania. His research interests include image processing, computer vision, medical images, random field models, and machine learning.

Daniel Carlos Guimarães Pedronette received his BSc degree in computer science (2005) from São Paulo State University (Brazil) and his MSc (2008) and PhD (2012) degrees in computer science (2008) from the University of Campinas (Brazil). He is currently an assistant professor at São Paulo State University, Rio Claro, Brazil. His research interests involve content-based image retrieval, unsupervised learning, reranking, rank aggregation, information retrieval, digital libraries, and image analysis.

Bruno Barufaldi received his PhD degree from the University of São Paulo in 2016 and he is currently in a postdoctoral researcher position at the University of Pennsylvania. He has expertise in software development for x-ray imaging systems. He works on radiation dose analysis on x-ray imaging systems, and quality assessment of digital mammography and digital breast tomosynthesis images. For the last two years, he has been active in the field of virtual clinical trials.

Andrew Douglas Arnold Maidment received his BASc degree in engineering science (1987) and PhD in medical biophysics (1993) from the University of Toronto, Canada. He is an associate professor of radiology at the Hospital of the University of Pennsylvania and chief of the physics section of the Department of Radiology. He is the author of more than 200 peer-reviewed papers. His research currently focuses on the development of advanced imaging modalities to improve breast cancer detection.