

An Unified Framework for Bayesian Denoising for Several Medical and Biological Imaging modalities

João M. Sanches Jacinto C. Nascimento Jorge S. Marques
 jmsr@isr.ist.utl.pt jan@isr.ist.utl.pt jsm@isr.ist.utl.pt
 Instituto de Sistemas e Robótica - Instituto Superior Técnico

Abstract—Multiplicative noise is often present in several medical and biological imaging modalities, such as MRI, Ultrasound, PET/SPECT and Fluorescence Microscopy. Noise removal and preserving the details is not a trivial task. Bayesian algorithms have been used to tackle this problem. They succeed to accomplish this task, however they lead to a computational burden as we increase the image dimensionality. Therefore, a significant effort has been made to accomplish this tradeoff, i.e., to develop fast and reliable algorithms to remove noise without distorting relevant clinical information. This paper provides a new unified framework for Bayesian denoising of images corrupted with additive and multiplicative noise. This allows to deal with additive white Gaussian and multiplicative noise described by Poisson and Rayleigh distributions respectively. The proposed algorithm is based on the *maximum a posteriori* (MAP) criterion, and an edge preserving priors are used to avoid the distortion of the relevant image details. The denoising task is performed by an iterative scheme based on Sylvester/Lyapunov equation. This approach allows to use fast and efficient algorithms described in the literature to solve the Sylvester/Lyapunov equation developed in the context of the Control theory. Experimental results with synthetic and real data testify the performance of the proposed technique, and competitive results are achieved when comparing to the of the state-of-the-art methods.¹

I. INTRODUCTION AND PRIOR WORK

The acquisition of medical Imaging introduces anatomical artifacts, which hamper the extraction of the relevant features for the medical diagnosis. Hence, denoising techniques must applied to significantly reduce the artifacts present in the image, and preserve the anatomical structure. A significant effort has been made to provide fast and reliable algorithms allowing noise removal and preserving relevant anatomical structures [1,2].

The white additive Gaussian noise is amongst the most popular models used to describe noisy signals [3]. This assumption is not acceptable in many imaging systems which produce images corrupted by multiplicative or impulsive noise [4,5].

Two types of multiplicative noise arise in a wide variety of imaging modalities such as medical and biological imaging: speckle noise and Poisson noise. The speckle noise usually appears in aquisition processes using coherent radiation, e.g., laser [6], ultrasound [7] and SAR [8] images. Poisson noise arises in systems involving counting procedures like PET/SPECT [9], functional MRI [10] and fluorescence confocal microscopy [11].

The additive Gaussian noise is commonly used in CT [12] and less frequently used in low intensity MRI [13]. Indeed, the noise corrupting the MRI images is usually approximated by a Rice and Gaussian distributions, for low and high intensities, respectively [13]. Common distributions describing speckle noise are: Rayleigh [13], Poisson [11], K-distribution [14], Nakagami [15], Fisher-Tippet [16] and *Generalized Gamma* (GG) [17].

In this paper a unifying denoising framework is presented to deal with additive Gaussian white noise and multiplicative noise described by Rayleigh and Poisson distributions. The paper shows that the MAP estimate of the image is the solution of a Sylvester-Lyapunov equation

$$\Phi X + X\Phi + Q = 0, \quad (1)$$

where Q is a known matrix in the case of the Gaussian noise and depends on X in the case of the Rayleigh and Poisson distributions. This paper formulates the denoising problem in terms of the Sylvester-Lyapunov equation which can be numerically solved by efficient algorithms. Different types of priors are also consider to preserve the discontinuities present in the noisy images. The proposed algorithm is compared with some of the most used despeckling filters, like the Wiener and the normal median filter, and a state-of-the-art adaptive median filter described in [18].

Prior work dealing with multiplicative noise can be framed into the following categories: *i*) Median filtering, *ii*) Wavelets and *iii*) Bayesian methods.

The median filter is the most popular choice since it is simple and provides high quality results. Variants addressing this approach are proposed in [18,19]. In the later, two versions are proposed: *i*) fixed - *Squeeze Box Filter* (SBF), and *ii*) variable - *Adaptive Weighted Median Filtering* (AWMF) size windows. In [18], it is illustrated that the proposed method outperforms other despeckling median filters, being a state-of-the-art method in this class.

Wavelet based despeckling algorithms have also been used, e.g., soft-thresholding denoising pioneered by Donoho [20]. These methods are based on multi-scale decomposition of the noisy image and in the processing of the image coefficients in coarser scales, e.g., [13,21,22].

Bayesian methods have also been proposed to deal with multiplicative noise reduction. This approach is closely related with the work presented herein. It formulates the denoising task as an estimation problem, where the likelihood function and a prior distribution are jointly maximized. The following three issues must be addressed: *i*) statistical observation model, *ii*) prior distribution and *iii*) optimization method [23]. Recently, denoising algorithms have been formulated in a Bayesian framework, and using a wavelet decomposition, achieving very efficient and fast algorithms [24,25].

This paper is organized as follows. Section II formulates the denoising problem. Section IV reports experimental results exhibiting the effectiveness of the proposed approach comparing its performance with a state-of-the-art method. Section V concludes the paper.

II. PROBLEM FORMULATION

Let X be a $N \times M$ unknown image to be estimated from the noisy image Y . The estimation of X is formulated in a Bayesian framework as the following optimization problem

$$\hat{X} = \arg \min_X E(X, Y) \quad (2)$$

where

$$E(X, Y) = \underbrace{E_Y(Y, X)}_{\text{Data fidelity term}} + \underbrace{E_X(X)}_{\text{Prior term}}. \quad (3)$$

$E_Y(X, Y)$, called *data fidelity term*, depends on the observation model and attracts the solution toward the data, and $E_X(X)$, called *prior term* or *internal energy*, regularizes the solution, removes the noise and incorporates a priori knowledge about the solution.

Usually it is assumed a priori that X is band-limited signal described by a *Markov Random Field* (MRF) defined using neighborhood interactions among set of pixels. By the Hammersley-Clifford theorem, this means that $P(X) = Ce^{-U(X)}$ is a Gibbs distribution where $C = 1/Z$ is the reciprocal of the partition function Z .

¹This work was partially supported by *FCT*, Portuguese Ministry of Science and Technology and Higher Education (which includes FEDER funds).

The interactions among neighbors, defined by the prior turn (2) into a well posed problem with a unique solution. However, the solution may appear smoothed and blurred at the transitions. To avoid this undesirable effect several authors have suggested different Gibbs distributions to model these interactions in order to minimize the distortions of the edges in the final solution. In general, this set of priors are called *edge preserving priors*, EPP. Two examples are the *total variation* and *Benford* priors [26].

Assuming statistical independence of the observations the energy function may be written as follows

$$E(X, Y) = - \sum_i p(y_i|x_i) + U(X) \quad (4)$$

where $p(y_i|x_i)$ is the distribution that models the acquisition process, $U(X) = \sum_k \rho(\delta_k)$ is the Gibbs energy, $\rho(\delta)$ is the so-called potential function, e.g., $\rho(x) = x^2$ is the simplest one, and δ_k are the intensity differences between pairs of neighboring pixels.

III. OPTIMIZATION

The minimization of (4) is an huge optimization task. It can be formulated in a matrix framework after vectorizing the images, $x = \text{vect}(X)$, in lexicographical order, i.e., by stacking the columns of X . Direct matrix processing of the data in this format is not possible, in practice, because the resulting matrices have huge dimensions, despite their sparseness structure. Therefore, the minimization of (4) is usually performed in a element wise basis.

In this paper an optimization method is proposed where the original matrix structure of the data is kept along the entire optimization process, making useless the vectorization of the data. Furthermore, the optimization problem is formulated by using the well known Sylvester/Lyapunov equation for which there are efficient and fast solvers described in the literature and the algorithm aims to be a generalized framework to cope with several image observation models and use several prior Gibbs distributions.

TABLE I

GAUSSIAN, POISSON AND RAYLEIGH MODELS FOR THE OBSERVATIONS.

Model	$p(y x)$	$\epsilon(x, y)$	$\partial\epsilon(x, y)/\partial x$	x^{ML}
Gauss	$K e^{-\frac{1}{2\sigma^2}(x-y)^2}$	$\frac{1}{2\sigma^2}(x-y)^2$	$\frac{x-x^{ML}}{\sigma^2}$	y
Poisson	$\frac{x^y e^{-x}}{y!}$	$x - y \log(x)$	$\frac{x-x^{ML}}{x}$	y
Rayleigh	$\frac{y}{x} e^{-y^2/(2x)}$	$\frac{y^2}{2x} - \log(\frac{y}{x})$	$\frac{x-x^{ML}}{x^2}$	$y^2/2$

The energy function to be minimized is the following,

$$E(Y, X) = \sum_{i,j}^{N,M} \epsilon(x(i, j), y(i, j)) + \alpha \sum_{i,j}^{N,M} [\omega_v(i, j)\delta_v^2(i, j) + \omega_h(i, j)\delta_h^2(i, j)] \quad (5)$$

where $\epsilon(x(i, j), y(i, j))$ are the terms associated with the data (see Table I), $y(i, j)$, and depends on the observation model and $\delta_\tau(i, j)$, with $\tau \in [h, v]$ are the horizontal and vertical differences between neighbors at each pixel location. $\omega_v(i, j)$ and $\omega_h(i, j)$ are weights updated in each iteration and obtained by using a specific MM algorithm proposed by [27]. These weights are related to the prior as follows,

$$\omega_\tau = \frac{1}{\delta p(\delta)} \frac{dp(\delta)}{d\delta} \quad (6)$$

where $p(\delta) = K e^{-\alpha\rho(\delta)}$ is the distribution of δ . Table II shows the weights computed with (6) for the three priors considered, L_2 , TV and Benford.

TABLE II
WEIGHTS FOR THE L_2 , TV AND BENFORD PRIORS, WHERE g^{t-1} DENOTES THE GRADIENT MAGNITUDE VALUE COMPUTED IN THE PREVIOUS ITERATION. SINCE WE ARE DEALING WITH ISOTROPIC FIELDS $\omega_h = \omega_v$.

Weights	L_2	TV	Benford
w	1	$1/g^{t-1}$	$[1/g^{t-1}]^2$

The partial derivative of $E(X, Y)$ w.r.t. $x(i, j)$ computed from (5) lead to

$$\begin{aligned} \frac{\partial E(X, Y)}{\partial x(k, l)} &= \frac{d\epsilon(x(k, l), y(k, l))}{dx(k, l)} \\ &+ 2\alpha [\omega_v(k, l)\delta_v(k, l) + \omega_h(k, l)\delta_h(k, l)] \\ &- 2\alpha [\omega_v(k+1, l)\delta_v(k+1, l) + \omega_h(k, l+1)\delta_h(k, l+1)] \end{aligned} \quad (7)$$

where $\omega_v(k, l) = \omega_h(k, l) = \omega(k, l)$ because we are using an isotropic MRF. Assuming the following simplifications

$$\begin{aligned} \omega_v(k+1, l) &\approx \omega_v(k, l) = \omega(k, l) \\ \omega_h(k, l+1) &\approx \omega_h(k, l) = \omega(k, l) \end{aligned}$$

and taking into account the equation (7) the stationary point may be computed by solving the following set of equations

$$\begin{aligned} \frac{1}{\omega(k, l)} \frac{d\epsilon(x(k, l), y(k, l))}{dx(k, l)} + 2\alpha [2x(i, j) - x(i-1, j) - x(i+1, j)] + \\ 2\alpha [2x(i, j) - x(i, j-1) - x(i, j+1)] = 0 \end{aligned} \quad (8)$$

Vertical differences
Horizontal differences

$0 \leq k, l \leq N, M$

TABLE III

MATRIX Σ FOR THE GAUSSIAN, POISSON AND RAYLEIGH OBSERVATION MODELS AND FOR THE L_2 , TV AND BENFORD PRIORS. THE OPERATOR \odot REPRESENTS A ELEMENT WISE OR HADAMMARD OPERATION.

Σ	Gaussian	Poisson	Rayleigh
L_2	$1/\sigma^2$	$1 \odot / X$	$1 \odot / X^{\odot 2}$
TV	G/σ^2	$G \odot / X$	$G \odot / X^{\odot 2}$
Bfd	$G^{\odot 2}/\sigma^2$	$G^{\odot 2} \odot / X$	$G^{\odot 2} \odot / X^{\odot 2}$

The set of equations (8) can be written using matrix notation

$$\Sigma \odot (X - X^{ML}) + 2\alpha(\Theta_N X + X\Theta_M) = 0 \quad (9)$$

where \odot is the Hadamard operator, the elements of X^{ML} are defined in Table I and Σ is a $N \times M$ matrix depending on the observation and prior models, as shown in table III. The matrix Θ_L , with $L \in [N, M]$, is as follows,

$$\Theta = \begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & \dots & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & \dots & \dots & 0 & 0 & -1 & 2 \end{pmatrix} \quad (10)$$

Equation (9) is non linear and must be iteratively solved. By using the fixed point method the following recursion is obtained

$$\Sigma^{t-1} \odot (X^{t-1} - X^{ML}) + 2\alpha(\Theta_N X + X\Theta_M) = 0 \quad (11)$$

where Σ^{t-1} is computed using the estimate X^{t-1} obtained in the previous iteration. Equation (11), solved in each iteration, is the well known Sylvester equation,

$$\Phi_N X + X\Phi_M + C^{t-1} = 0 \quad (12)$$

where

$$\begin{aligned} \Phi_N &= \frac{1}{2} I_N + 2\alpha\sigma^2\Theta_N \\ \Phi_M &= \frac{1}{2} I_M + 2\alpha\sigma^2\Theta_M \\ C^{t-1} &= \Sigma^{t-1} \odot (X^{t-1} - X^{ML}). \end{aligned}$$

IV. EXPERIMENTAL RESULTS

In this section we present two sets of experiments. In both sets we present a performance comparison of the following denoising methods: *i*) median filtering, *ii*) Wiener filtering, *iii*) *Squeeze Box filter* (SBF), *iv*) L_2 prior, *v*) *Total Variation* (TV) prior, and *vi*) *Benford* prior.

In the first set of experiments (Section IV-A), we present the performance of the mentioned methods in synthetic images analyzing the SNR for each technique. Then, we illustrate the results in several images modalities: *Ultrasound* (US), *Computed Tomography* (CT).

A. Synthetic Images

To assess the performance of the above mentioned methods, we performed Monte Carlo tests in two scenarios. We consider a synthetic image, which consists of a white centered square in a black background. This image corrupted with *i*) gaussian, and *ii*) Rayleigh noise distributions. The experiments comprise the following: in *i*) we change the σ parameter in the set $\sigma \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$, and we carried out 20 experiments for each σ value; in *ii*) we vary the $\sigma \in \{10, 100, 500, 1000, 5000\}$ keeping the same number of the experiments for each value of σ . Fig. 1 shows the mean of SNR for the different values of σ . The SNR is computed in a diagonal profile of the image from top-left to bottom-right.

It can be seen for gaussian noise (Fig. 1 (a)) that for small values of σ the Wiener method exhibits the best SNR, since is tailored to deal with this kind of noise. However, it strongly degrades as the σ increases. The median filtering has a similar poor performance for higher values of σ . The SBF has a good and almost constant SNR, nevertheless, the best SNR are achieved by the TV and Benford priors. Fig. 1 (b) reports the performance of the methods under a rayleigh noise corrupted image. The SBF and TV, Benford priors exhibit remarkable results. As in the previous experiment, the TV and Benford prior render the most attractive results.

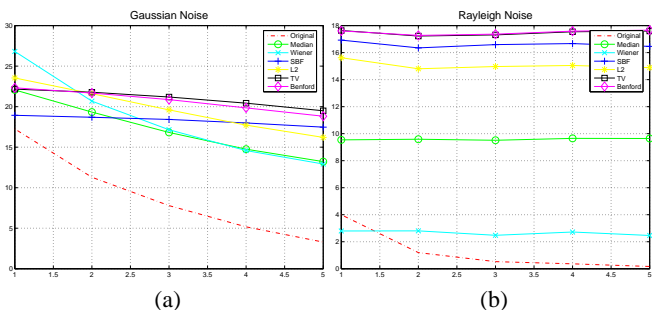


Fig. 1. SNR obtained for various methods: (a) gaussian, (b) rayleigh noise distributions.

A second experiment is shown in Fig. 2. This figure shows an input and a simulation of an ultrasound image to evaluate the performance to the various despeckling algorithms considered² This image consists of a background region class with a “one” pixel value, and other two classes with “zero” (dark region) and “ten” pixel values. We compute the SNR in a vertical line passing through the centers of the circles, in which the image profile is computed. In Fig. 2 it is illustrated the output image for each method, as well as the obtained intensity profiles: red line is the profile of the input image; the blue line depicts the resulting image profile for technique.

The best SNR results is achieved by the SBF method, however the TV and Benford priors can provide competitive results.

B. Real Images

To illustrate the performance of the denoising algorithms, we use real images from several medical modalities. The images output and the respective intensity profiles are shown. These profiles are obtained

²this test is similar to the one presented in [18].

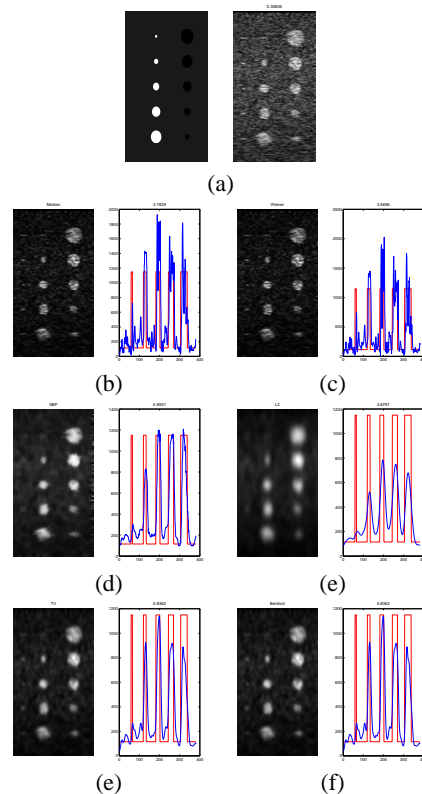


Fig. 2. SNR obtained for various methods: (a) original image; (b) median - $SNR = 3.78$; (c) Wiener - $SNR = 3.57$; (d) SBF - $SNR = 5.86$; (e) L_2 - $SNR = 4.88$; (f) TV - $SNR = 5.84$, (g) Benford - $SNR = 5.84$.

along a diagonal line (from upper left corner to bottom right corner). The original intensity profile is shown in blue, whereas the black line refers to the intensity of the filtered image. Due to the lack of space only a subset of images are provided.

Fig. 3, shows the resulting results in a US image. For the three Bayesian methods (L_2 , TV and Benford), the noise is modeled by a Rayleigh distribution. The comparison of the results obtained by the several methods shows that the best results are obtained by the TV and Benford priors as well as the SBF, which is tailored to deal with US images. This is illustrated in Fig. 4 where it is displayed the image profiles obtained in each method (blue line).

In Fig. 5 other modality (MRI) is presented. In this case the noise is modeled as being white, additive and Gaussian. It is desirable to compare the results with the Wiener method since it is a benchmark when dealing with Gaussian noise. From this point of view it is interesting to note that the EPP filters provides amongst the best results.

V. CONCLUSIONS

The application of a Bayesian methodology to image denoising processes is usually time and memory consuming. The usual procedure to deal with images as vectors leads to huge dimension matrices which are difficult to manipulate. In this paper we present an approach where the original 2D structure of the images is kept during the denoising process.

Herein, we show that the solution of the denoising problem with multiplicative and Gaussian noise can be obtained by using the Sylvester/Lyapunov matrix equation. This equation can be efficiently solved using numerical methods described in the literature and included in software packages (e.g., Matlab).

MAP filtering with non quadratic priors were also considered. We propose an iterative version of the gaussian MAP algorithm to be used with non-gaussian priors or with deblurring problems.

VI. ACKNOWLEDGEMENT

We are very grateful to Peter C. Tay, Scott T. Acton and John A. Hossack for providing the source code of the SBF method proposed in

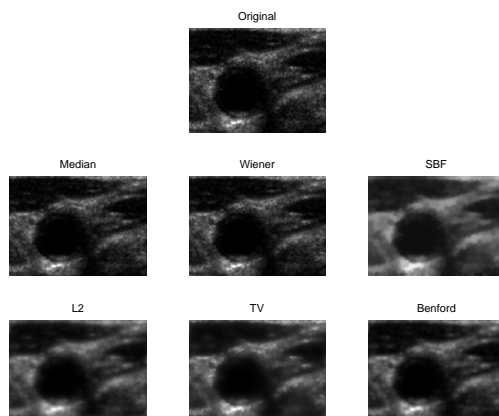


Fig. 3. Output of the methods in US images: (first row) heart, (second row) carotid, (third row) thyroid.

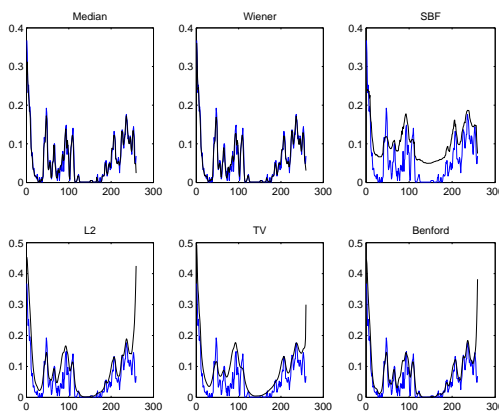


Fig. 4. Output of the methods in MRI images: (first row) head, (second and third rows) knee.

[18] as well as the image presented in Fig. 2, which assuredly was helpful to enrich the quality of the paper.

REFERENCES

- [1] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. on Medical Imaging*, vol. 25, no. 8, pp. 987–1010, 2006.
- [2] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.*, vol. 2, 2000.
- [3] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing*. pub-PH, 2000.
- [4] A. K. Jain, *Fundamentals of digital image processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.
- [5] A. Toprak and I. Güler, "Suppression of impulse noise in medical images with the use of fuzzy adaptive median filter," *J. Med. Syst.*, vol. 30, no. 6, pp. 465–471, 2006.
- [6] B. Zagar and H. Weiss, "Processing of laser speckle interferometer signals for high precision strain measurement," in *Proc. of the 8th International Symposium on Artificial Intelligence Based Measurement and Control*, vol. 1, Kyoto, Japan, 1991, pp. 99–105.
- [7] C. Burckhardt, "Speckle in ultrasound b-mode scans," *IEEE Trans. on Sonics and Ultrasonics*, vol. SU-25, no. 1, pp. 1–6, 1978.
- [8] P. Xu, "Despeckling sar-type multiplicative noise," *Int. Journal of Remote Sensing*, vol. 20, no. 13, pp. 2577–2596, 1999.
- [9] J. M. Ollinger and J. A. Fessler, "Positron emission tomography," *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 43–55, 1997.
- [10] G. E. Hagberg, G. Zito, F. Patria, and J. N. Sanes, "Improved detection of event-related functional mri signals using probability functions," *NeuroImage*, vol. 2, pp. 1193–1205, 2001.

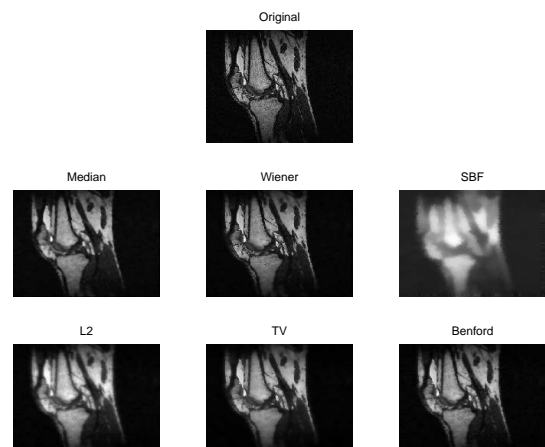


Fig. 5. Output of the methods in CT images: (first row) heart, (second row) lung, (third row) vascular.

- [11] G. van Kempen, L. van Vliet, and P. Verveer, "Application of image restoration methods for confocal fluorescence microscopy," in *Proceedings SPIE, 3-D Microscopy: Image Acquisition and Processing IV*, C. Cogswell, J. Conchello, and T. Wilson, Eds., vol. 2984. SPIE, 1997, pp. 114–124.
- [12] P. Gravel, illes Beaudoin, and J. A. D. Guise, "A method for modeling noise in medical images," *IEEE Trans. Medical Imaging*, vol. 23, no. 10, pp. 1221–1232, 2004.
- [13] P. Bao and L. Zhang, "Noise reduction for magnetic resonance images via adaptative multiscale products thresholding," *IEEE Trans. on Medical Imaging*, vol. 22, no. 9, 2003.
- [14] T. K. Keyes and W. T. Tucker, "The k-distribution for modeling the envelope amplitude of a backscattered signal," *IEEE Trans. on Ultrason., Ferroelect. and Freq. Control*, vol. 46, no. 4, pp. 883–887, 1999.
- [15] P. Shankar, "A general statistical model for ultrasonic backscattering from tissues," *IEEE Trans. Ultrason., Ferroelect. Freq. Control*, vol. 47, pp. 727–736, 2000.
- [16] J. M. Sanches and J. S. Marques, "Compensation of log-compressed images for 3d ultrasound," *Ultrasound in Medicine & Biology*, vol. 29, no. 2, pp. 239–253, 2003.
- [17] O. V. Michailovich and A. Tannenbaum, "Despeckling of medical ultrasound images," *IEEE Trans. Ultrason., Ferroelect. Freq. Control*, vol. 53, no. 1, pp. 64–78, 2006.
- [18] P. C. Tay, S. T. Acton, and J. A. Hossack, "Ultrasound despeckling using an adaptative window stochastic approach," in *Proc. of the Int. Conf. on Image Processing*, Atlanta, GA, USA, 2006, pp. 2549–2552.
- [19] T. Loupas, W. McDicken, and P. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Trans. on Circuits and Systems*, vol. 36, pp. 129–135, 1989.
- [20] D. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 613–627, 1994.
- [21] S. Gupta, L. Kaur, R. C. Chauhan, and S. C. Saxena, "A wavelet based statistical approach for speckle reduction in medical ultrasound images," *Medical and Biological Engineering and computing*, vol. 42, 2004.
- [22] A. Nicolaidis and et. al., "Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery," *IEEE Trans. Visualization and Computer Graphics*, vol. 52, no. 10, pp. 1653–1669, 2005.
- [23] C. R. Vogel, *Computational methods for inverse problems*. pub-SIAM, 2002.
- [24] D. Gleich and M. Datcu, "Gauss-markov model for wavelet-based sar image despeckling," *IEEE Signal Processing Letters*, vol. 13, no. 6, pp. 365 – 368, 2006.
- [25] J. Xu and S. Osher, "Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising," *IEEE Trans. on Image Processing*, vol. 16, no. 2, pp. 534–544, 2007.
- [26] J. Sanches and J. S. Marques, "Image reconstruction using the benford law," in *Proceedings ICIP 2006*. Atlanta GA, USA: IEEE International Conference on Image Processing, October 2006.
- [27] J. Dias, "Fast gem wavelet-based image deconvolution algorithm," 2003.