

Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Image-guided filtering for improving photoacoustic tomographic image reconstruction

Navchetan Awasthi
Sandeep Kumar Kalva
Manojit Pramanik
Phaneendra K. Yalavarthy

SPIE.

Navchetan Awasthi, Sandeep Kumar Kalva, Manojit Pramanik, Phaneendra K. Yalavarthy, "Image-guided filtering for improving photoacoustic tomographic image reconstruction," *J. Biomed. Opt.* **23**(9), 091413 (2018), doi: 10.1117/1.JBO.23.9.091413.

Image-guided filtering for improving photoacoustic tomographic image reconstruction

Navchetan Awasthi,^a Sandeep Kumar Kalva,^b Manojit Pramanik,^b and Phaneendra K. Yalavarthy^{a,*}

^aIndian Institute of Science, Department of Computational and Data Sciences, Bangalore, India

^bNanyang Technological University, School of Chemical and Biomedical Engineering, Singapore

Abstract. Several algorithms exist to solve the photoacoustic image reconstruction problem depending on the expected reconstructed image features. These reconstruction algorithms promote typically one feature, such as being smooth or sharp, in the output image. Combining these features using a guided filtering approach was attempted in this work, which requires an input and guiding image. This approach act as a postprocessing step to improve commonly used Tikhonov or total variational regularization method. The result obtained from linear backprojection was used as a guiding image to improve these results. Using both numerical and experimental phantom cases, it was shown that the proposed guided filtering approach was able to improve (as high as 11.23 dB) the signal-to-noise ratio of the reconstructed images with the added advantage being computationally efficient. This approach was compared with state-of-the-art basis pursuit deconvolution as well as standard denoising methods and shown to outperform them. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.23.9.091413]

Keywords: photoacoustic imaging; guided image filtering; Lanczos Tikhonov; basis pursuit deconvolution; total variation; model-based reconstruction.

Paper 180031SSRRR received Jan. 16, 2018; accepted for publication Jun. 1, 2018; published online Jun. 25, 2018.

1 Introduction

Photoacoustic tomography (PAT) is a noninvasive imaging modality combining benefits of optical contrast and ultrasonic resolution.^{1–6} This modality uses a short pulsed laser as the radiation source to irradiate the biological tissue. The tissue absorbs the optical energy resulting in small temperature rise, in turn generates pressure waves by photon absorption and thermoelastic expansion. These pressure (acoustic) waves are detected by the ultrasound transducers placed at the boundary of the imaging domain. The major strength of the photoacoustic (PA) imaging is that the detected acoustic waves have less scattering and attenuation in biological tissues, making them propagate through thick biological tissues easily. The detected PA waves on the boundary are typically utilized to generate an initial pressure distribution inside the tissue. The initial pressure distribution is proportional to the product of optical fluence and absorption coefficient. The optical absorption provides information about tissue components, such as melanin, bilirubin, lipids, water, hemoglobin, and oxyhemoglobin.^{7,8} The biochromophores are helpful in assessing the pathophysiological condition of the tissue. The applications of PA imaging include monitoring tissue health condition in the fields of cardiology,⁹ ophthalmology,¹⁰ oncology,^{11,12} dermatology,^{13,14} and neurosciences.¹⁵

The important step in the PA tomographic imaging is the image reconstruction, which is an initial value problem. In this, the aim is to estimate the initial pressure at time $t = 0$, given the boundary pressure data at time “ t .” Several reconstruction methods are available for estimating the initial pressure. The standard reconstruction algorithms include analytical algorithms

[backprojection (BP) and delay-and-sum] and time-reversal (TR) algorithms.^{16–19} These are computationally efficient as compared to the model-based reconstruction algorithms but require a large amount of data to obtain meaningful results. The requirement of a large amount of data results in either increased scan time or expensive instrumentation setups. Moreover, typical setups used for the PA tomographic measurements record the ultrasound waves over an aperture, which may not cover the object,^{20–22} resulting in only limited data.

Reconstruction of PA images in these limited data cases using either analytical or TR methods results in poor quality images, often having limited ability in terms of much required contrast. The model-based image reconstruction algorithms were proposed in the literature for these limited data cases, which improve the quantitative accuracy of reconstructed images.^{23–26} These algorithms especially iterative in nature also provide robustness to noise in the boundary data,^{16,27} making them attractive in real-time. These model-based iterative reconstruction algorithms tend to be computationally complex compared to analytical algorithms like linear backprojection (LBP) and require utilization of regularization to constrain the solution space. Often, the reconstructed image characteristics obtained by analytical algorithms are complementary to the model-based reconstruction algorithms.^{28–30}

To improve the PA imaging performance, the earlier attempts included applying signal enhancement methods on the PA data collected.^{31–35} These methods typically apply deconvolution on the raw data collected to improve recorded acoustic signal.^{31–35} These approaches have limited utility (will also be shown with an example using the method attempted in Ref. 33) compared to postprocessing of reconstructed PA images. In addition,

*Address all correspondence to: Phaneendra K. Yalavarthy, E-mail: yalavarthy@iisc.ac.in

applying an enhancement method for the recorded PA data may be computationally expensive, especially in cases where large amounts of data are collected.³⁶ It will be ideal to develop a method that work independent of the amount of the boundary data collected.

Postprocessing of reconstructed medical images is part of standard clinical diagnosis, which typically involves enhancing the contrast of images.³⁷ In this step, the reconstructed images are typically either normalized in terms of the histogram or applied a filter to enhance the diagnostic accuracy. Specific to PA imaging, there were attempts earlier to perform blind deconvolution to improve the resolution in optical-resolution PA microscopy (OR-PAM).³⁸ This blind deconvolution process involves estimation of the blur kernel or point spread function using a Wiener filter type estimator, which needs prior information about the signal-to-noise ratio (SNR) of the input image.³⁸ In continuation of these efforts, assuming that regularization in the model-based image reconstruction can be used to provide the blurring matrix, the image reconstruction was considered as a two step process.^{28,39} The first step being model-based image reconstruction and the second one being the deblurring using the estimated model-resolution matrix via sparse recovery-based methods that utilized the ℓ_1 -based regularization.²⁸ Often, these postprocessing methods, especially like deblurring/deconvolution steps, are computationally demanding requiring computational times on par with the initial step of model-based reconstruction.^{28,39}

The recent advances in image fusion in computer vision enabled researchers to effectively combine the best features from the input and guiding images to provide an output image that is superior to both of them.⁴⁰ Especially, in areas such as multicamera as well as multifocus fusion, the results have been encouraging with utility of guided filtering, which can be seen as a state-of-the-art image fusion method with little to no computational complexity.⁴¹ In this work, the same guided filtering⁴² based approach has been used to improve the reconstruction results obtained from various reconstruction schemes that are typically used in PA image reconstruction. The proposed approach also can be seen as a two step process, initially, one performs reconstruction using two reconstruction methods [examples could be LBP^{28,30,43} and total variational (TV) based regularization method]^{44,45} and the reconstruction results are combined using guided filtering to provide superior reconstruction result compared to both of them. It is also demonstrated that the standard denoising methods, such as wavelet thresholding and nonlocal means (NLM), which are typically applied in standard postprocessing of reconstructed images, provides very limited utility in cases discussed here. It is shown using both numerical and experimental phantom along with *in-vivo* results that the guided filtering approach provides superior results compared to the established basis pursuit deconvolution (BPD) method.^{28,39} More importantly, these results also establish the utility of the image fusion methods to improve PA image reconstruction. All discussion in this work is limited to two-dimensional imaging as the aim is to show the utility of the proposed guided filtering approach in PA imaging.

2 Photoacoustic Image Reconstruction

The forward problem in the PAT involves collecting the pressure data by ultrasound detectors on the boundary of the imaging domain, given an initial pressure distribution. The PA wave equation can be written as follows:⁶

$$\nabla^2 P(x, t) - \frac{1}{c^2} \frac{\partial^2 P(x, t)}{\partial t^2} = \frac{-\beta}{C_p} \frac{\partial H(x, t)}{\partial t}, \quad (1)$$

where $P(x, t)$ is the pressure at position x and time t , c is the speed of sound in the medium, β is the thermal expansion coefficient, C_p is the specific heat capacity, and $H(x, t)$ represents the energy deposited per unit time per unit volume. The reconstruction (inverse) problem involves an estimate of the initial pressure [$P(x, t)$ at $t = 0$] inside the imaging domain, given the boundary measurements at time t , making the reconstruction (inverse) problem equivalent to an initial value problem.

2.1 System Matrix Building

The numerical experiments setup carried out in this work is the same as in Refs. 29, 30, and 46. For completeness, it is briefly reviewed here. The imaging domain has a dimension of $n \times n$ (herein, $n = 201$). It is vectorized by stacking the columns into a $n^2 \times 1$ vector and represented as x . The system matrix (\mathbf{A}) has a dimension of $m \times n^2$. Here, m is the number of detectors multiplied with number of time samples. Each column of the system matrix represents the impulse response corresponding to every entry in the vectorized image. The columns of the data are also stacked into a long vector (b) (dimension $m \times 1$).^{29,30}

The forward model for PA imaging can be written as follows:

$$\mathbf{A}x = b, \quad (2)$$

where x is a long column vector having a dimension of $n^2 \times 1$ ($n = 201$) and b is the measurement vector with a dimension of $m \times 1$ ($m = 100 \times 500$).

LBP image (x_{LBP}) can be obtained as follows:^{28,43}

$$x_{\text{LBP}} = \mathbf{A}^T b, \quad (3)$$

where \mathbf{T} indicates the transpose of the matrix. BP methods are noniterative in nature making them computationally efficient. The drawback of BP methods is that they provide only qualitative results especially in limited data cases, and hence avoided for quantitative imaging.²⁷ This BP image (x_{LBP}) is typically used as an initial guess for iterative model-based image reconstruction methods.^{26,30} In this work, x_{LBP} was used as guiding image to improve the reconstruction results obtained from model-based reconstruction methods.

2.2 k-Wave Time Reversal Method

The TR is a standard one-step image reconstruction method that can be performed using the open-source k-wave toolbox.⁴⁷ In this method, the solution of the wave equation $P(x, t)$ vanishes outside the time point N (the longest time for which the PA wave is traveling inside the domain).⁴⁸ The zero initial conditions are imposed and the model is solved backward in time, thus giving the solution [$P(x, 0)$] at $t = 0$. Even though TR provides a model-based solution, often the reconstruction results are dependent on the amount of available data as well as bandwidth of acoustic detectors.^{18,48} As the boundary data considered here are limited, the interpolated data were used to estimate the initial pressure inside the imaging domain.⁴⁷ This method was utilized to reconstruct the initial pressure distribution using both limited bandwidth as well as full bandwidth data. Note that as all

experimental acoustic detectors are always band-limited, the full bandwidth results are limited to numerical phantom cases considered here.

2.3 Lanczos Tikhonov Regularization Method

In cases of limited data, the PA tomographic image reconstruction problem becomes ill-conditioned. The Tikhonov regularization method is the most commonly used regularization method for solving these ill-conditioned problems. Tikhonov regularization works on the principle of minimizing the residual of the linear equations along with smooth regularization term. The objective function for Tikhonov regularization can be written as follows:

$$\Omega = \|\mathbf{A}x - b\|_2^2 + \alpha\|x\|_2^2, \quad (4)$$

where α is the regularization parameter, which provides a balance between the residual and the expected initial pressure distribution. The function to be minimized is Ω and the l_2 -norm is represented by $\|\cdot\|_2$. Equation (4) minimized with respect to x results in

$$(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})x_{\text{Tikh}} = \mathbf{A}^Tb. \quad (5)$$

The regularization parameter (α) controls the reconstructed initial pressure distribution characteristics. Higher regularization tends to oversmooth the image, while a lower value of regularization parameter amplifies the noise in the image. Solving these linear systems of equations [Eq. (5)] requires $O(n^6)$ operations with $n = 201$ for the problem at hand, making it computationally demanding.

The Lanczos bidiagonalization⁴⁹ provides dimensionality reduction of the problem, making the reconstruction problem more tractable in terms of required number of operations as well as storage. This method was previously described in Ref. 29 for estimation of the optimal regularization parameter (α) in the Tikhonov-framework that utilizes the QR decomposition. The dimensionality reduction is performed on system matrix, \mathbf{A} , via Lanczos bidiagonalization, as given in Ref. 49. The left and right Lanczos matrices and the bidiagonal matrix are related to the system matrix \mathbf{A} as follows:^{29,46}

$$\mathbf{M}_{k+1}(\beta_0 e_1) = b, \quad (6)$$

$$\mathbf{A}\mathbf{N}_k = \mathbf{M}_{k+1}\mathbf{B}_k, \quad (7)$$

$$\mathbf{A}^T\mathbf{M}_{k+1} = \mathbf{N}_k\mathbf{B}_k^T + \alpha_{k+1}n_{k+1}e_{k+1}^T, \quad (8)$$

where \mathbf{B}_k represents the lower bidiagonal matrix, \mathbf{M}_k and \mathbf{N}_k represent the left and right orthogonal Lanczos matrices, and β_0 is the l_2 -norm of b . The dimensions of left and right orthogonal Lanczos matrices are $(m \times k)$ and $(n^2 \times k)$, respectively, with k representing the number of iterations for which bidiagonalization was performed. The unit vector of dimension k is represented as e_k ($=1$ at the k 'th and 0 otherwise).

Equations (6)–(8) can be utilized to rewrite as follows:

$$b - \mathbf{A}x = \mathbf{M}_{k+1}(\beta_0 e_1 - \mathbf{B}_k x^{(k)}); \quad x = \mathbf{N}_k x^{(k)}, \quad (9)$$

where $x^{(k)}$ represents the dimensionality reduced version of x with $k \ll n^2$. Substituting Eq. (9) in Eq. (4) and using the

property $\mathbf{M}_{k+1}^T \mathbf{M}_{k+1} = \mathbf{I}$, the cost function can be rewritten as follows:^{29,46}

$$\hat{\Omega} = \|\beta_0 e_1 - \mathbf{B}_k x^{(k)}\|_2^2 + \alpha\|x^{(k)}\|_2^2. \quad (10)$$

Considering the first-order condition, the solution becomes

$$x_{\text{est}}^{(k)} = (\mathbf{B}_k^T \mathbf{B}_k + \alpha \mathbf{I})^{-1} \beta_0 \mathbf{B}_k^T e_1; \quad x_{\text{est}} = \mathbf{N}_k x_{\text{est}}^{(k)}, \quad (11)$$

where $x_{\text{est}}^{(k)}$ is the estimate of $x^{(k)}$ for a fixed k and a given regularization parameter α .⁵⁰ Note that $x_{\text{est}} = x_{\text{Tikh}}$ only if $k = n^2$ for a given value of α . For other cases, x_{est} becomes an approximation of x_{Tikh} . There exists an optimal k , which reduces the error between x_{est} and x_{Tikh} within the precision limits.^{46,49} Since dimensionality reduction is obtained using the Lanczos algorithm, this method offers computationally efficient solutions as compared to the other methods, including the solution given in Eq. (5).

The performance of model-based reconstruction scheme depends on the choice of reconstruction parameter, such as α and k .^{51,52} Error estimate methods have been utilized for finding the number of iterations (k) for the Lanczos method^{51–53} along with α (readers are encouraged to refer to Bhatt et al.⁵³ for more details, it is briefly reviewed here). The residual of the system r is defined as follows:

$$r = b - \mathbf{A}x. \quad (12)$$

The norm of the error estimate can be written as follows:⁵³

$$\|e\|_2^2 \approx \eta_v^2 := e_0^{v-1} e_1^{5-2v} e_2^{v-3}, \quad (13)$$

where $v \in \mathbb{R}$, $e_0 := \|r\|_2^2$, $e_1 := \|\mathbf{A}^T r\|_2^2$, and $e_2 := \|\mathbf{A}\mathbf{A}^T r\|_2^2$. The error estimate for $v = 2$ becomes⁵⁵

$$\eta_2 = \frac{\|r\|_2 \|\mathbf{A}^T r\|_2}{\|\mathbf{A}\mathbf{A}^T r\|_2}. \quad (14)$$

The regularization parameter α is found by minimizing Eq. (14):

$$\eta_2(\alpha) = \frac{\|r\|_2^2 \|\mathbf{A}^T r\|_2^2}{\|\mathbf{A}\mathbf{A}^T r\|_2^2}. \quad (15)$$

For the minimum value of α , the number of steps k is calculated. This error estimate approach⁵³ was utilized to get the optimal values for k and α in this work and the reconstructed image is represented by x_{LTO} . For the case of k and α being chosen heuristically, it is represented by x_{LTH} .

2.4 Basis Pursuit Deconvolution Method

In Ref. 28, BPD was utilized to deblur the solution obtained using the Lanczos-Tikhonov regularization method (discussed in Sec. 2.2). As regularization blurs the solution, the effect of regularization can be overcome by the BPD method, so it is typically used for deblurring the x_{LTH} . It utilizes the split augmented Lagrangian shrinkage algorithm (SALSA)⁵⁴ to minimize the following objective function, which uses ℓ_1 -type regularization to promote sharp features:

$$\overline{\Omega} = \|\mathbf{H}x^{(k)} - x_{\text{est}}^{(k)}\|_2^2 + \sigma\|x^{(k)}\|_1, \quad (16)$$

Algorithm 1 Algorithm of SALSA.

Input: \mathbf{A} , b , λ , $k = 0$, v_0 and d_0

Output: Solution vector: x_{k+1}

- 1 Repeat 2-5 till convergence is satisfied
- 2 $x_{k+1} = \arg \min_x \|\mathbf{A}x - b\|_2^2 + \lambda \|x - v_k - d_k\|_2^2$
- 3 $v_{k+1} = \arg \min_v \tau\phi(v) + (\lambda/2) \|x_{k+1} - v - d_k\|_2^2$
- 4 $d_{k+1} = d_k - (x_{k+1} - v_{k+1})$
- 5 $k = k + 1$

where \mathbf{H} is the model resolution matrix, which represents the blur matrix, defined as

$$\mathbf{H} = (\mathbf{B}_k^T \mathbf{B}_k + \alpha \mathbf{I})^{-1} \mathbf{B}_k^T \mathbf{B}_k. \quad (17)$$

The solution obtained by minimizing Eq. (16) with respect to $x^{(k)}$ is denoted as $x_d^{(k)}$ (solved by utilizing Algorithm 1) and the final deblurred solution becomes

$$x_{\text{BPD}} = \mathbf{N}_k x_d^{(k)}. \quad (18)$$

2.5 Total Variational Regularization Method

The Tikhonov regularization [Eq. (4) or Eq. (10)] promotes a smooth solution due to the quadratic nature of regularization. The TV regularization method is used for solving the ill-posed problem by introducing a regularization term minimizing the variation in the solution (x). The cost function for the same can be written as follows:^{44,45}

$$\Gamma = \|\mathbf{A}x - b\|_2^2 + \lambda \|x\|_{\text{TV}}, \quad (19)$$

where λ is the regularization parameter. Here, $\|\cdot\|_{\text{TV}}$ denotes the isotropic TV utilizing the form in Ref. 55. For TV regularization, the alternating direction method of the multipliers framework given in Ref. 56 was utilized, and the algorithm known as SALSA⁵⁴ (given in Algorithm 1) was deployed in this work.

2.6 Denoising Using Wavelet Thresholding (Hard and Soft)

One of the standard techniques used for denoising of images is wavelet thresholding. Its application ranges from noise reduction, signal and image compression up to signal recognition.⁵⁷ Thresholding works by setting the high frequency sub-band coefficients that are less than a threshold to zero.⁵⁷ Both hard thresholding (HT) and soft thresholding (ST) have been used in this work.⁵⁸ The advantage of this method is that the denoising approach is model-free and can be applied as a postprocessing step. As the aim of this work was to show that the guided filtering performs better than the standard denoising method, here, the input image was chosen to be x_{LTO} (which has best image quality among the reconstructed results). The input image x_{LTO} is transformed into the wavelet domain:

$$\hat{a} = \text{DWT}(x_{\text{LTO}}), \quad (20)$$

where \hat{a} represents the image in the wavelet domain and DWT stands for discrete wavelet transform (here, Haar wavelets were utilized).^{58,59} The coefficients in the wavelet domain are represented by a . After thresholding is performed on all wavelet coefficients, the coefficients are represented as y . The image is again transformed from wavelet domain to x domain, which is represented as \hat{x} :

$$\hat{x} = \text{IDWT}(y), \quad (21)$$

with IDWT representing inverse DWT. As it is expected that noise is present in the higher order wavelet coefficients, this thresholding will eliminate the noise.

2.6.1 Hard thresholding

It works on the rule that if a wavelet coefficient value is less than a threshold, it is set to zero:⁵⁷

$$y = \begin{cases} a & \text{if } |a| \geq \sigma \\ 0 & \text{if } |a| < \sigma \end{cases}, \quad (22)$$

where σ is the threshold value, a is the input wavelet coefficient, and y is the output wavelet coefficient. A large value of threshold leads to a large bias error while a smaller value increases the variance.

2.6.2 Soft thresholding

It is defined as follows:⁵⁷

$$y = \begin{cases} \text{sgn}(a)f(|a| - \sigma) & \text{if } |a| \geq \sigma \\ 0 & \text{if } |a| < \sigma \end{cases}, \quad (23)$$

where $f(a)$ is a linear function, which is a straight line with slope being dependent on σ ⁵⁷ and $\text{sgn}(a)$ is defined as follows:

$$\text{sgn}(a) = \begin{cases} -1 & \text{if } a < 0 \\ 0 & \text{if } a = 0 \\ 1 & \text{if } a > 0 \end{cases}. \quad (24)$$

The ST provides a better alternative to HT. In simple terms, the ST can be considered as linear transformed version of HT results with an aim to enhance the signal content in the reconstructed image along with suppression of noise. Note that σ was chosen in this work based on the SNR of the raw/boundary data.

2.7 Denoising Using Nonlocal Mean Filter

The other standard method for denoising of images is NLM filtering.⁶⁰ Similar to the wavelet denoising method, this also does not rely on any imaging model and can be applied as a postprocessing method to denoise images that are corrupted with Gaussian noise. It is defined as follows:

$$\text{NL}[x_{\text{LTO}}](i) = \sum_{j \in I} w(i, j) x_{\text{LTO}}(j), \quad (25)$$

where the weights $w(i, j)$ depend on the similarity between pixels and satisfy conditions $0 \leq w(i, j) \leq 1$ and $\sum_j w(i, j) = 1$. The similarity between two pixels depend on the closeness of the intensity gray level vectors $x_{\text{LTO}}(N_i)$ and $x_{\text{LTO}}(N_j)$. Here, N_k denotes a grid of fixed size having center at pixel k . The

similarity is a decreasing function of the weighted Euclidean distance, $\|x_{\text{LTO}}(N_i) - x_{\text{LTO}}(N_j)\|_{2,a}^2$, where $a > 0$ is the standard deviation of the Gaussian kernel.⁶⁰ The weights in Eq. (25) are defined as follows:

$$w(i, j) = \frac{1}{Z(i)} e^{-\frac{\|x_{\text{LTO}}(N_i) - x_{\text{LTO}}(N_j)\|_{2,a}^2}{h^2}}, \quad (26)$$

with $Z(i)$ being the normalizing constant, given as

$$Z(i) = \sum_j e^{-\frac{\|x_{\text{LTO}}(N_i) - x_{\text{LTO}}(N_j)\|_{2,a}^2}{h^2}}, \quad (27)$$

where h is the degree of filtering, which controls the decay of the exponential function. The filter parameters are chosen as described in Ref. 60. The NLM-based filtering is considered as the state-of-the-art denoising method that is purely image driven and has been widely applied for medical images.

2.8 Image-Guided Filtering (Proposed Method)

In many applications of computer vision, filtering is used to enhance the signal content and suppress the noise in images. Various linear translation-invariant (LTI) filters, such as Gaussian, Laplacian, and NLM (discussed above), have been used in image restoration, deblurring/sharpening of images, etc.⁶¹ These filters (LTI) are spatially invariant and independent of the imaging model. They are deployed depending on the application and the performance of these filters varies depending on the extra information that could be used in the filtering process, which can be provided by a different image. For example, in anisotropic diffusion,⁶² gradient information has been utilized as extra information, which helps in avoiding smooth edges. The weighted least squares filter⁶³ utilizes the filtering input as the guidance. The guiding/guidance image can be another image instead of the filtering image as in image matting,⁶⁴ haze removal,⁶⁵ and colorization.⁶⁶ This process optimizes a quadratic function involving large sparse matrices and is computationally demanding,^{63–66} making their utility limited in real-time. Another way to make use of the guiding image is by explicitly introducing it in the filter kernels as in bilateral filter.^{67–69} Here, the output becomes a weighted average of the nearby pixels and the weights are calculated using the intensity in the guiding image. The bilateral filter can smooth small fluctuations in the image and preserve edges, but it is often prone to gradient reversal artifacts.⁶³ Thus, there is a need for a filter that can perform image fusion, which is computationally efficient and simultaneously enhance the information of the input image based on guiding image.

A recently proposed guided filter produces the output by combining the content of an input (to be filtered) image and guiding image, by performing a linear transform of the guiding image.⁴² It has better behavior compared to the bilateral filter (no gradient reversal) and it is not a smoothing filter.⁴² It can transfer structures of the guiding image to the input (to be filtered) image, which found applications in filtering techniques, such as dehazing and guided feathering. Also, it uses a fast linear time algorithm, which is invariant to the kernel size or the intensity range of the image.

The guided filter has become a state-of-the-art method in computer vision and computer graphics fields with specific applications, including image fusion,⁴¹ edge aware smoothing,

structure transferring, high dynamic range compression, image feathering/matting, dehazing,⁷⁰ detail enhancement,⁷¹ flash/no-flash denoising,⁷² and joint upsampling. These applications have been encouraging to apply the guided filtering in the PA image reconstruction framework.

In PA imaging, the reconstructed image characteristics are dependent on reconstruction technique and there are both analytical and model-based reconstruction methods that are being widely used. Moreover, each reconstruction scheme has variable computational complexity, for example, the analytical-based ones being fast and model-based ones being relatively computationally complex. It will be ideal to develop a fusion method that can combine the features obtained by two different reconstruction methods to result in a superior fused/filtered image. In this work, with the help of guided filtering, the reconstructed image details were enhanced by combining results from two different reconstruction methods that provide distinct image characteristics. For example, the LBP image typically tends to be smooth but is prone to streak artifacts and also lacks sharp features, especially in limited data cases. The TV or any other sparse reconstruction method (such as Lanczos Tikhonov) will be able to provide these sharp features, but often they also enhance the noise present in the boundary data. It will be ideal to develop a image fusion method that can effectively fuse BP (analytical type result) and model-based (here either TV or Lanczos Tikhonov) reconstructed images, which is proposed here with utility of image-guided filter.

The joint image filtering (involving guiding and to-be-filtered image) frameworks, including the current one, is derived from minimizing an objective function involving a fidelity term and a regularization term. The regularization in the guided filtering case is implicitly imposed on the objective function with the fidelity term only. With addition of the smoothness term, the regularization scheme becomes explicit. There are methods that utilize the explicit regularization⁷³ in the joint image filtering, but are often computationally more expensive compared to implicit methods. In the implicit regularization case (applicable to the proposed method), the guiding image is utilized within a local window to implicitly result in a weight/filter function to be applied on the input image. This allows the structure of the guiding image to be transferred to the input image (similar to bilateral filter⁶⁷) with important features like edges as well as corners to be preserved while allowing noise suppression. This results in a spatially variant filtering kernel, depending on the local window, that is utilized in the guided filtering approach. Simply, the weighted averaging regularizes the filtering process, resulting in a regularized output image. As this process is applied locally, all implicit methods are easy to implement involving little to no computational burden. The detailed discussion about implicit and explicit image filtering approaches is present in Ref. 73 for the interested readers.

Specifically, the input image for image-guided filter can be the Lanczos Tikhonov image obtained either using the heuristic (x_{LTH}) or optimal (x_{LTO}), or the TV image (x_{TV}), which needs to be filtered (also known as to be guided image, represented by x_{TG}). The guiding image represented by $x_G (= x_{\text{LBP}})$ obtained via Eq. (3) can be used to make x_{TG} more structured and less noisy to result in the filtered image (x_F). In this approach, the output image (x_F) is a transformation of x_G in a window ω_p centred at pixel p and is given as follows:^{41,42}

$$x_F^i = a_p x_G^i + b_p, \quad \forall i \in \omega_p, \quad (28)$$

where a_p and b_p are linear coefficients that are constant in ω_p that needs to be determined. The model ensures that x_F has an edge only if x_G has an edge, as $\nabla x_F = a \nabla x_G$. Suppose n^i is the amount of noise present in the x_{TG} image, then the reconstructed image x_F can be represented as follows:

$$x_F^i = x_{TG}^i - n^i. \quad (29)$$

The determination of coefficients (a_p and b_p) were achieved by minimizing the cost function obtained after combining Eqs. (28) and (29) and is written as follows:⁴¹

$$E(a_p, b_p) = \sum_{i \in \omega_p} [(a_p x_G^i + b_p - x_{TG}^i)^2 + \epsilon a_p^2], \quad (30)$$

where ϵ is the regularization parameter penalizing larger values of coefficient a_p . The coefficients obtained after minimizing Eq. (30) are given as follows:⁴¹

$$a_p = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_p} x_G^i x_{TG}^i - \mu_p \bar{x}_{TG}^i}{\sigma_p^2 + \epsilon}, \quad (31)$$

$$b_p = \bar{x}_{TG}^p - a_p \mu_p, \quad (32)$$

with μ_p and σ_p^2 being the mean and variance of x_G , respectively, in ω_p , $|\omega|$ is the number of pixels in ω_p , and $\bar{x}_{TG}^i = \frac{1}{|\omega|} \sum_{i \in \omega_p} x_{TG}^i$ is the mean of x_{TG} in ω_p . For the modified filter, Eqs. (31) and (32) are changed as in Ref. 74. Since a pixel is involved in all the overlapping windows, x_F is computed using an average over all the windows and the output thus becomes⁴¹

$$x_F^i = \frac{1}{|\omega|} \sum_{p|i \in \omega_p} (a_p x_G^i + b_p). \quad (33)$$

Due to the symmetry, $\sum_{p|i \in \omega_p} a_p = \sum_{p \in \omega_i} a_p$, Eq. (33) can be written as follows:

$$x_F^i = \bar{a}_i x_G^i + \bar{b}_i. \quad (34)$$

Algorithm 2 lists all important steps including variations for the original guided filter. Note that in Algorithm 2, β acts as the control parameter and α is the parameter helping in making a sharper transition from the low-pass to high-pass filtering mode.

The guided filter has distinct advantages of improving image quality as it is an edge preserving, gradient preserving, and structure transferring filter.^{41,42} This method is applied as a post-processing step after reconstruction and its performance is compared against wavelet-based ST as well as HT techniques and NLM filtering-based denoising methods.

The edge preserving can be explained, when one assumes that the guiding image is the same as the input (to be guided) image. In this case, the guiding filter coefficients can be written as $a_p = \frac{\sigma_p^2}{\sigma_p^2 + \epsilon}$ and $b_p = (1 - a_p) \mu_p$. When $\epsilon > 0$, two cases arise:

- If the image x_G has high variance in ω_p , then $\sigma_p \gg \epsilon$, so $a_p \approx 1$ and $b_p \approx 0$. Thus, if a pixel is in a high variance area, then its value remains same, i.e., $x_F = x_G$.

Algorithm 2 Modified guided filter ($x_{TG}, x_G, r, \epsilon, \alpha, \beta$) with $f_a(\cdot)$ representing performing operation "a" on the arguments.

Input: x_{TG} -To be Filtered input image (x_{LTO}, x_{LTH}, x_{TV})

x_G – Guidance Image(x_{LBP})

r -Window radius (patch size)

ϵ, α, β – Regularization Parameters

Output: x_F – Filtered Image

1: $\text{mean}_{x_G} = f_{\text{mean}}(x_G)$;

$\text{mean}_{x_{TG}} = f_{\text{mean}}(x_{TG})$

2: $\text{corr}_{x_G} = f_{\text{mean}}(x_G \cdot x_G)$;

$\text{corr}_{x_G-TG} = f_{\text{mean}}(x_G \cdot x_{TG})$

3: $\text{var}_{x_G} = \text{corr}_{x_G} - \text{mean}_{x_G} \cdot \text{mean}_{x_G}$;

$\text{cov}_{x_G-TG} = \text{corr}_{x_G-TG} - \text{mean}_{x_G} \cdot \text{mean}_{x_{TG}}$

4: $a = [\text{cov}_{x_G-TG} / (\text{var}_{x_G} + \epsilon)]^\alpha$;

$b = \text{mean}_{x_{TG}} - \beta \cdot a \cdot \text{mean}_{x_G}$

5: $\text{mean}_a = f_{\text{mean}}(a)$;

$\text{mean}_b = f_{\text{mean}}(b)$

6: $x_F = \text{mean}_a \cdot x_G + \text{mean}_b$

- If the image x_G is almost constant in ω_p , then $\sigma_p \ll \epsilon$, so $a_p \approx 0$ and $b_p \approx \mu_p$. Thus, if the pixel is in a flat patch area, then its value becomes the average of neighboring pixels.

The guided filter performs gradient preserving since it uses a patchwise model. For the self-guiding image, $a_p < 1$ and b_p are constant. Suppose the detail layer is given as $d = x_{TG} - x_F$ and utilizing $\partial_x x_F = a_p \partial_x x_{TG}$, one can write

$$\partial_x d = \partial_x x_{TG} - \partial_x x_F = (1 - a_p) \partial_x x_{TG}, \quad (35)$$

which implies that $\partial_x x_d$ and $\partial_x x_{TG}$ are always in the same direction, thus, the gradient is always preserved.

2.9 Automated Wavelet Denoising of Recorded Photoacoustic Data

One of the standard denoising method in the data domain is the wavelet denoising,³³ using maximum overlap DWT (MODWT),⁷⁵ which is a non-orthogonal transform. The benefit of using the MODWT is that it is applicable for any sample size (not restricted to powers of 2) and hence zero padding is not necessary along with eliminating arbitrary truncation. The MODWT contrast to DWT forms a zero-phase filter, resulting in lining up features with original signal. For complete discussion of MODWT, please see Ref. 75, and with respect to application to the PA image, the readers are encouraged to see Ref. 33. In this work, we have utilized the MODWT to implement wavelet smoothing on the noisy PA signals, which automatically sets a denoising threshold using universal threshold

criteria.⁷⁶ Earlier attempts of using this method have been shown to improve the image quality by 22%.³³

3 Figures of Merit

To quantitatively assess reconstructed image quality using methods discussed till now, image metrics, such as root mean square error (RMSE), contrast-to-noise ratio (CNR), and SNR, were utilized in this study. The first two were utilized in cases of numerical phantoms, where the target/expected initial pressure distribution is available. The latter one is applied in experimental phantom cases, where the ground truth is not available.

3.1 Root Mean Square Error

The RMSE is a common metric for assessing the performance of reconstructed images in PAT^{23,77–79} and it is defined as follows:

$$\text{RMSE}(x^{\text{target}}, \hat{x}) = \sqrt{\frac{\sum (x^{\text{target}} - \hat{x})^2}{N}}, \quad (36)$$

where N is the total number of pixels of the imaging domain, x^{target} is the expected initial pressure distribution, and \hat{x} is the reconstructed (or filtered) initial pressure distribution. The lesser value of RMSE indicates the closeness of the reconstructed/denoised image with the expected image.

3.2 Contrast-to-Noise Ratio

CNR is defined as follows:^{36,80,81}

$$\text{CNR} = \frac{\mu_{\text{roi}} - \mu_{\text{back}}}{(\delta_{\text{roi}}^2 a_{\text{roi}} + \delta_{\text{back}}^2 a_{\text{back}})^{1/2}}, \quad (37)$$

where μ and δ represent the mean and the variance corresponding to the region of interest (roi) and the background (back) in the reconstructed/filtered initial pressure. The roi refers to the region of the phantom, where initial pressure values are greater than zero and the background refers to the region of the phantom, where it is 0. The $a_{\text{roi}} = \frac{A_{\text{roi}}}{A_{\text{tot}}}$ and $a_{\text{back}} = \frac{A_{\text{back}}}{A_{\text{tot}}}$ represent the area ratio.⁸⁰ A_{roi} represents area of the region of interest and A_{back} represents area of the background. A_{tot} is the sum of areas of the region of interest and the background. a_{roi} and a_{back} are the noise weights for the region of interest and the background, respectively. Higher value of CNR represents the better contrast recovery of the reconstructed/filtering method.⁸¹

3.3 Signal-to-Noise Ratio

The SNR is expressed as follows:⁸²

$$\text{SNR}_r \text{ (in dB)} = 20 \times \log_{10} \left(\frac{S}{n} \right), \quad (38)$$

where S denotes the peak initial pressure value and n corresponds to standard deviation of the whole image. This is standard metric for defining the image quality.⁶¹ Higher SNR represents the lesser noise in the reconstructed image and thus better performance of the reconstruction/filtering method. To differentiate it from the SNR of the boundary data, this figure of merit for the reconstructed image was denoted by SNR_r .

4 Numerical and Experimental Studies

The measurement of actual initial pressure distribution in real experiments is challenging. Comparing of performance of different reconstruction algorithms in these cases is challenging, so, typically, numerical phantoms are deployed for effective comparison. In this work, four such numerical phantoms were considered.

The computational imaging grid has a dimension of 501×501 (0.1 mm/pixel) and the detectors were placed on a circle of radius 22 mm. The schematic diagram showing the data acquisition setup along with acoustic detectors placement is given in Fig. 1. The experimental data were generated on a high dimensional grid having a dimension of 401×401 and the reconstruction was performed on a lower dimensional grid having a dimension of 201×201 to mimic real experimental scenario. The numerical phantoms spans the imaging domain of size 20.1 mm \times 20.1 mm. The data generated using the high dimensional grid were added with white Gaussian noise to result in SNR levels ranging from 20 to 60 dB. An open source toolbox *k-wave*⁴⁷ in MATLAB was used for generating the data. The sampling frequency used for the data collection is 20 MHz and the number of time samples for each detector was 500. A hundred acoustic detectors placed on the circle at equidistance (as depicted in Fig. 1) with center frequency of 2.25 MHz and 70% bandwidth collected the boundary data. The system matrix (\mathbf{A}) thus formed has a dimension of $50,000 \times 40,401$. The speed of sound was assumed to be 1500 m/s and the medium was assumed to be homogeneous with no dispersion or absorption of sound.

Initially, three numerical phantoms with maximum initial pressure distribution of 1 kPa [blood vessel (BV), modified Derenzo, and PAT, as given in Figs. 3(a)–3(c), respectively] were used to show the effectiveness of the proposed method. The numerical BV phantom [Fig. 3(a)] consists of thick and thin BV mimicking the typical BV structures. The modified Derenzo phantom [Fig. 3(b)], which consists of circular objects with varying diameter grouped according to size, was also used to assess the performance of the reconstruction method in terms

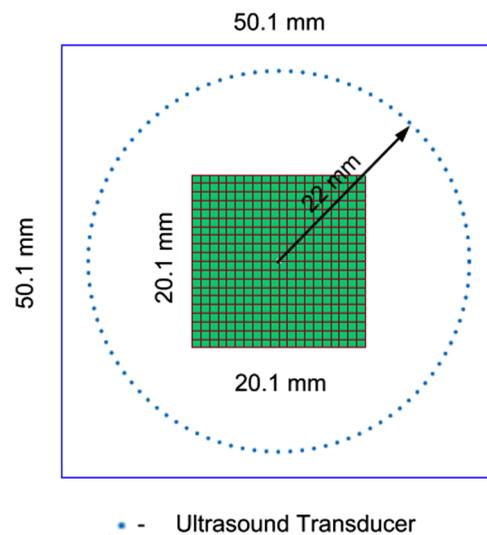


Fig. 1 Schematic diagram of PA data acquisition geometry along with depiction of position of 100 acoustic detectors (shown by dots) around the imaging domain. The computational imaging grid size is 50.1 mm \times 50.1 mm.

of size. Another phantom consisting of the letters “PAT” [Fig. 3(c)] was also used to determine the ability of the reconstruction method in terms of recovering sharp edges. The fourth numerical phantom mimics the single slice of real breast, which was created using contrast-enhanced imaging data.^{83,84} In this phantom, the initial pressure was varied from 0 to 5 and the expected initial pressure distribution is given in Fig. 3(d). The experiments were also performed using 50 detectors with the BV phantom to assess the ability to recover the structures with sparse data (here, the dimension of b and A is $25,000 \times 1$ and $25,000 \times 40,401$, respectively).

The schematic of the experimental setup for the collection of the PA data is shown in Fig. 2(a). It is similar to Fig. 1(e) of Ref. 85 (as well as Fig. 2 in Ref. 30). A Q-switched Nd:YAG laser (Continuum, Surelite Ex) was used to deliver a laser pulse of 5-ns duration at 10-Hz repetition rate having wavelength of 532 nm. Four right-angle uncoated prisms (PS911, Thorlabs) and one uncoated planoconcave lens L1 (LC1715, Thorlabs) were utilized to deliver the laser pulses to the sample. The laser energy density on the phantom was 9 mJ/cm^2 , which is within the safety limit ($<20 \text{ mJ/cm}^2$: ANSI safety limit⁸⁶). A triangular shaped horse hair phantom was utilized [photograph is shown in Fig. 2(c)] having side-length and diameter of 10 and 0.15 mm, respectively. The horse hair phantom was attached to the pipette tips adhered on acrylic slab.⁸⁷ The PA data were collected continuously around the horse hair phantom for full 360 deg using a 2.25-MHz flat ultrasound transducer (Olympus NDT, V306-SU) with a 13-mm diameter active area and 70% nominal bandwidth. Another experimental phantom was used

consisting of circular tubes. It was made using low density polyethylene tubes (5-mm inner diameter) and were filled with Indian black ink [photograph is shown in Fig. 2(b)]. The tubes were placed at 0 and 15 mm from the scanning center and affixed at the bottom of the acrylic slab. The reconstructed slice in these cases is the middle slice of the imaging phantom.

Using the same set-up, as in Fig. 2, *in vivo* rat brain imaging (healthy one) was carried out at 1064-nm wavelength as the light penetration is deeper in the near-infrared region. The laser energy density on the animal head (skin) was 9.7 mJ/cm^2 , which was well below the ANSI safety limit of 100 mJ/cm^2 at 1064 nm. Healthy female rats weighing $90 \pm 5 \text{ g}$ were procured from InVivos Pte Ltd. to conduct this *in vivo* animal experiment. All animal experiments were carried out according to the guidelines and regulations approved by the institutional Animal Care and Use committee of Nanyang Technological University, Singapore (Animal Protocol Number ARF-SBS/NIE-A0263). Before performing the experiments, rats were anesthetized using a mixture containing ketamine (85 mg/Kg) and xylazine (15 mg/Kg) by injecting a dosage of 0.2 mL/100 g intraperitoneally. The animals were placed in the PA scanner after trimming the hair on the head and epilating using hair removal cream. The animal was maintained under anesthesia during the experiments using 0.75% isoflurane gas (Medical Plus Pte Ltd., Singapore) along with oxygen (1.2 L/min). The anesthesia mixture was delivered through a nose cone with a breathing mask covering the nose and mouth of the animal. A custom made animal holder was used to place the animal in a sitting position on its abdomen

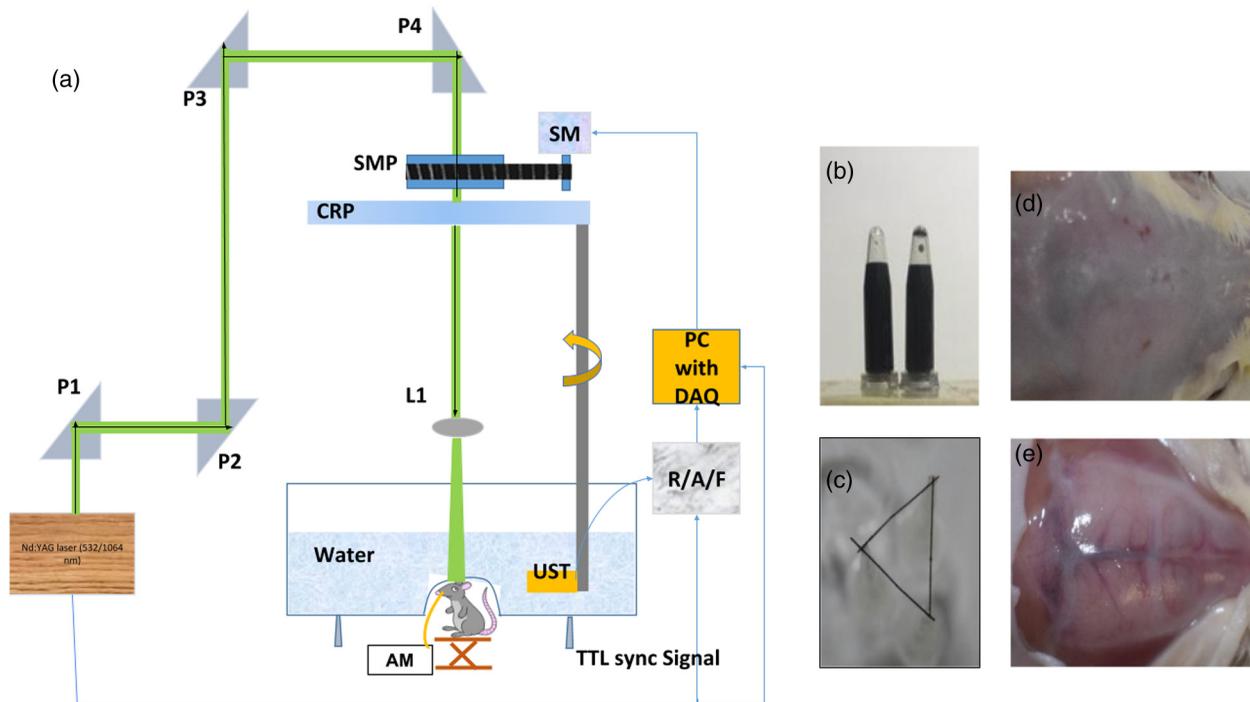


Fig. 2 (a) Schematic diagram showing the experimental PA data acquisition. The major components of the experimental set-up are: CRP, circular rotating plate; SM, stepper motor; SMP, stepper motor pulley unit; P1,P2,P3,P4, uncoated right-angled prisms; L1, planoconcave lens; R/A/F, receiver, amplifier, and filter for PA signal; DAQ, data acquisition card; UST, ultrasound transducer. (b) Photograph of circular-shaped tubes filled with India black ink. (c) Photograph of triangular-shaped horse hair phantom. (d) Photograph of rat brain after trimming and removal of hair on the head before PAT imaging. (e) Photograph of rat brain after cutting and open the top skin layer in the head region after the PAT imaging is done.

and surgical tapes were used to hold the animal in this position. Then, the animal along with the animal holder was mounted on a translational stage during the experiment. Height of the animal was adjusted using this translational stage in order to adjust the scanning plane of the rat brain. The photograph of the lateral plane with and without skin layer is shown in Figs. 2(d) and 2(e), respectively. Animals were euthanized after the experiment by injecting an overdose of Valbarb (sodium pentobarbitone) intraperitoneally.

The acquired PA signals were first amplified and filtered using a pulse amplifier (Olympus-NDT, 5072PR) and then recorded using a data acquisition (DAQ) card (GaGe, compscope 4227) using a single channel with a sampling frequency of 25 MHz inside a desktop (Intel Xeon 3.7 GHz 64-bit processor, 16 GB RAM, running windows 10 operating system). The synchronization of data acquisition with laser illumination was achieved using a sync signal from the laser. The reconstructed PA imaging region contains 200×200 pixels and has a size of $40 \text{ mm} \times 40 \text{ mm}$. The data collected have 2400 A-lines averaged over six times resulting in 400 detected signals collected by the ultrasound transducers acquiring the data continuously around the hair phantom and tube phantom in full 360 deg for an acquisition time of 240 s with a rotational speed of 1.5 deg/s. The data collected for *in vivo* animal brain had 4800 A-lines averaged over 48 times resulting in 100 detected signals acquired by the ultrasound transducer continuously in full 360 deg for an acquisition time of 480 s with a rotational speed of 0.75 deg/s. This averaging of A-lines reduced the energy fluctuations during multiple firings of the laser. The ultrasound transducer and the phantom were immersed in water to enable ultrasound coupling. The distance from the center of the system to the face of the ultrasonic transducer is 37.02 mm for hair phantom, 38.22 mm for tube phantom, and 50.76 mm for *in vivo* imaging. The speed of sound was considered to be 1500 m/s in all cases. The PA data were acquired with sampling frequency of 25 MHz (1024 samples), and subsampled at half the rate at 12.5 MHz to result in 512 time samples. The system matrix (\mathbf{A}) had a dimension of $51,200 \times 40,000$ (51,200: number of detectors are 100 with

each collecting 512 time samples, 40,000: dimensions of the imaging domain being 200×200). Determining the actual initial pressure rise (target values) in these experiments is not plausible, so the performance of each reconstruction method was quantitatively compared with SNR_r [Eq. (38)] being the figure of merit. Note that a Linux workstation with 32 cores of Intel Xeon processor having a speed of 3.1 GHz with 128 GB RAM was used for all computations performed in this work.

5 Results and Discussion

For completeness, in numerical phantom cases, initially, the TR method (based on k-wave) was utilized on full bandwidth data collected from 100 detectors with SNR of 40 dB and these results are compiled in Figs. 3(e)–3(h). As the number of detectors is still 100 (limited), the corners of these images have a blurring effect due to partial volume effect. The figures of merit—RMSE and CNR—for these results along with data SNR being 20 and 60 dB, are compiled in Table 2. Note that as real acoustic detectors are always band-limited, it is not possible to obtain these results in case of experimental and *in-vivo* phantom cases.

The reconstructed initial pressure distributions for the numerical BV phantom [Fig. 3(a)] limited bandwidth data collected from 100 detectors with SNR of 40 dB using k-wave TR method is shown in Fig. 4(a) and for LBP in Fig. 4(b). The result obtained using the Lanczos Tikhonov method using the heuristic parameters ($\alpha = 0.3$ and $k = 40$), x_{LTH} , is shown in Fig. 4(c). Note that the same k and α values were used for other experiments in obtaining x_{LTH} . The same effort in terms of performing reconstruction using the Lanczos Tikhonov method with the reconstruction parameters chosen by the error estimate method, known as x_{LTO} , is given in Fig. 4(d). One of the advantage of the BPD method lies with its ability to improve the reconstruction results that are obtained by heuristic choice of reconstruction parameters,²⁸ so the reconstruction result presented in Fig. 4(c) was deconvolved using BPD and the output is shown in Fig. 4(e). The reconstruction result using TV regularization, x_{TV} , is provided in Fig. 4(f). The proposed image-guided filtering results with input images being x_{LTH} , x_{LTO} , and x_{TV} are given correspondingly in Figs. 4(j)–4(l) with guiding image

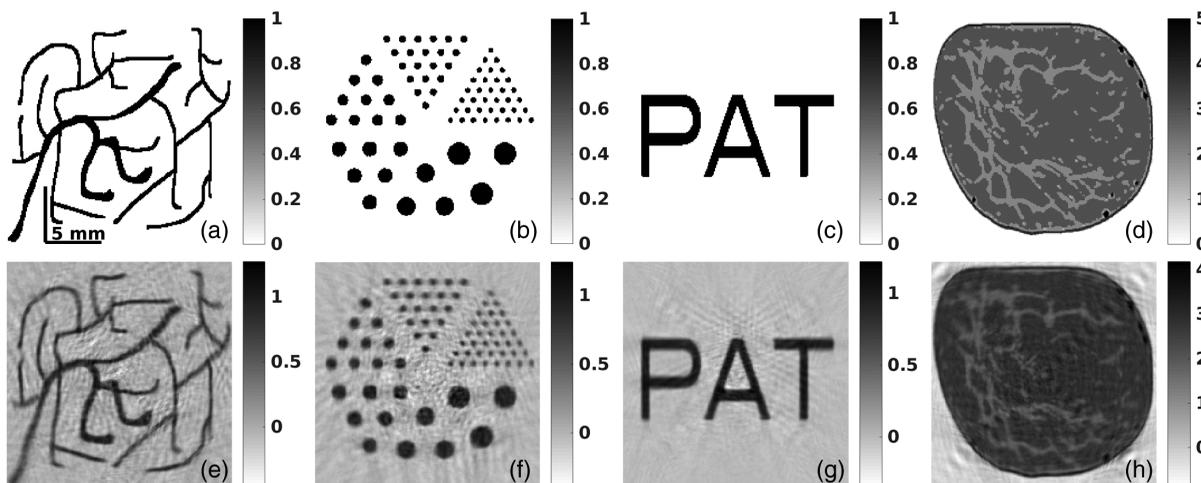


Fig. 3 The target numerical phantoms that were considered in this work: (a) BV phantom, (b) modified Derenzo phantom, (c) “PAT” phantom, and (d) realistic breast phantom. The reconstructed initial pressure rise using full bandwidth data with SNR of 40 dB corresponding to 100 detectors using k-wave TR corresponding to each target phantom (a, b, c, d) are given as (e, f, g, h). The corresponding figures of merit—RMSE and CNR—are given in Table 2 against the row “TR (full).”

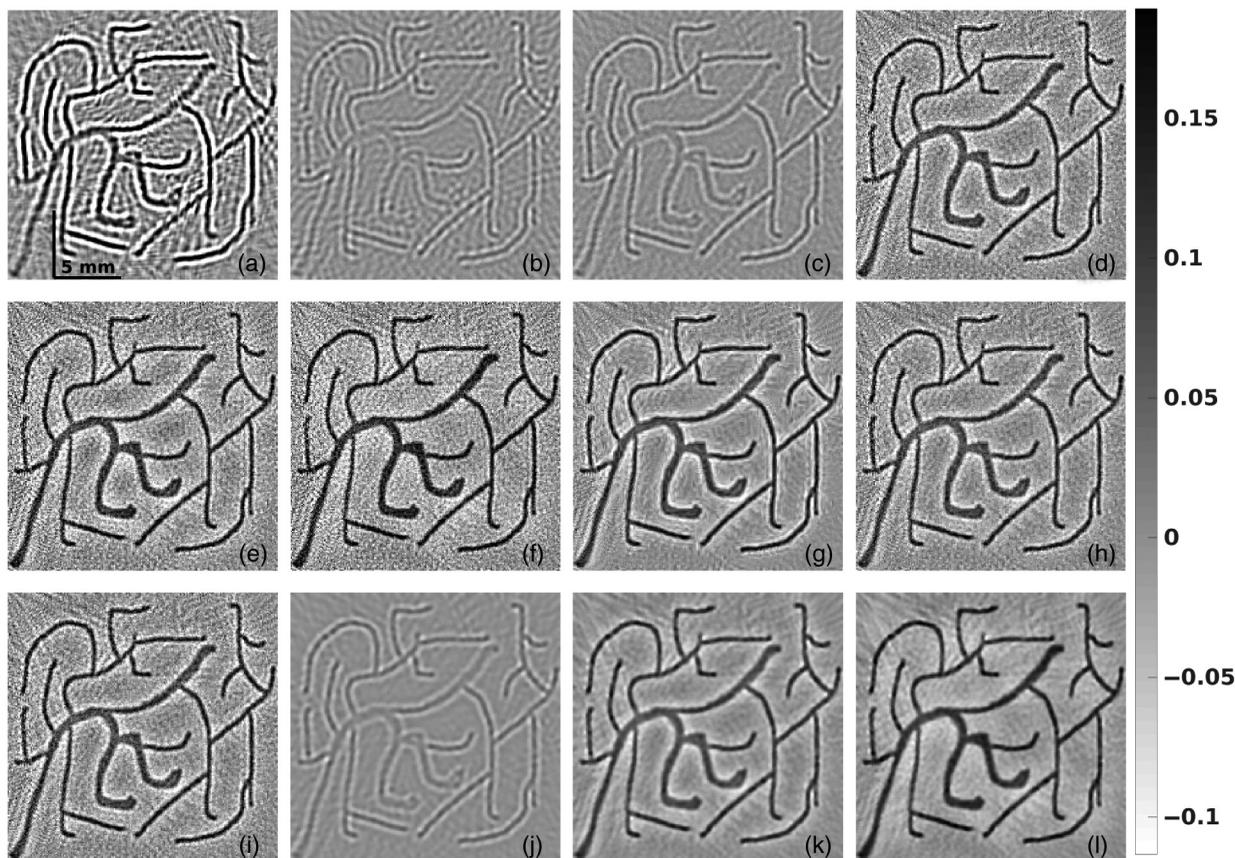


Fig. 4 The reconstructed initial pressure distribution using 100 detectors data with SNR of 40 dB corresponding to numerical BV phantom using (a) k-wave TR, (b) LBP, (c) LTH, (d) LTO, and (e) BPD with input image being (c), (f) TV, (g) NLM, wavelet, (h) ST, and (i) HT denoising with input image being (d). The proposed guided filter results with input images being (c), (d), and (f) are given in (j), (k), and (l), respectively, with (b) being the guiding image. The target/expected distribution is given in Fig. 3(a). The corresponding figures of merit—RMSE and CNR—are given in Table 2. The computational times corresponding to these methods are presented in Table 1.

being x_{LBP} [Fig. 4(b)]. For obtaining these results, the value of ϵ was chosen to be 1×10^{-3} . The patch size was taken as one (corresponding to 3×3 neighborhood), while the value of α and β being 1.05. The results using standard denoising methods applied to x_{LTO} [Fig. 4(d)] using NLM filtering and wavelet thresholding-based methods are shown in Figs. 4(g)–4(i), respectively. The computational time recorded for reconstructing these images is listed in Table 1. The third column of Table 2 provides recorded figures of merit, RMSE and CNR, for these results.

From these results, it is obvious that the guided filtering approach improves the reconstructed images using either Tikhonov or TV regularization methods without adding any significant computational burden. Even though the standard denoising methods (NLM and wavelet thresholding) were able to improve marginally these results, the computational complexity for methods like NLM is far superior to be utilized in real-

time. Also, the bipolar effect is more pronounced for the TR case mainly due to band-limited data utilized here [compare with the result shown in Fig. 3(e)]. From these results (given in Table 2), the guided filter results with x_{LTO} method, the improvement obtained in the RMSE is 45% while for the CNR, the improvement obtained is 79% for 40 dB SNR case compared to the TV method, which gives the best RMSE and CNR without filtering. The performance compared to methods like BPD is superior, improving the figures of merit by 100% [comparing guided filter (TV) result with BPD result in Table 2].

The automated wavelet denoising of recorded PA data using MODWT (described in Sec. 2.9) was utilized to denoise the signals recorded on the numerical BV phantom [Fig. 3(a)]. The SNR of the recorded data was 40 dB. An example of denoised and recorded noisy PA signals are given in Fig. 5. The RMSE between the noise-free PA signal and noisy signal (dashed line) is 0.744. The RMSE in the case of denoised PA signal (solid

Table 1 Recorded computational time (in seconds) for the results presented in Fig. 4.

Method	TR	LBP	LTH	LTO	BPD	TV	ST	HT	NLM	Guided filtering
Figure	4(a)	4(b)	4(c)	4(d)	4(e)	4(f)	4(g)	4(h)	4(i)	4(j)–4(l)
Time	78.12	0.5	38	166.89	0.001	5.92	0.05	0.05	375.15	0.01

Table 2 The figures of merit—RMSE and CNR—for the methods discussed in this work with SNR of data being 20, 40, and 60 dB, utilizing four numerical phantoms (BV, Derenzo, PAT, and realistic breast). The proposed method, guided filter, results are given as last three rows with input image (to be filtered) being given as argument. The entries in bold corresponds to highest RMSE and CNR observed for each column excluding the TR method that utilizes full bandwidth data (indicated in italics). Note that the entries in parenthesis corresponding to TR indicate the bandwidth of detected signals.

Phantom	BV (100 detectors)			Derenzo			PAT			BV (50 Detectors)			Breast																	
	20 dB	40 dB (Fig. 4)	60 dB	20 dB	40 dB (Fig. 9)	60 dB	20 dB	40 dB (Fig. 10)	60 dB	20 dB	40 dB (Fig. 11)	60 dB	20 dB	40 dB (Fig. 12)	60 dB															
Metric	RMSE	CNR	RMSE	CNR	RMSE	CNR	RMSE	CNR	RMSE	CNR	RMSE	CNR	RMSE	CNR	RMSE	CNR														
LBP	0.2487	1.31	0.2575	1.32	0.2561	1.38	0.2151	1.51	0.2118	1.48	0.2119	1.54	0.2319	0.95	0.2191	1.00	0.2189	1.00	0.2325	1.18	0.2636	1.31	0.2620	1.32	0.0596	0.31	0.0597	0.31		
LTH	0.2260	1.63	0.2354	1.83	0.2080	1.83	0.1387	2.05	0.1430	2.15	0.1447	2.15	0.2567	1.23	0.2596	1.36	0.2627	1.36	0.2235	1.39	0.2564	1.63	0.2488	1.63	0.0563	0.42	0.0553	0.44	0.0551	0.44
LTO	0.2308	1.49	0.1546	2.62	0.1738	2.16	0.1299	2.11	0.1226	3.13	0.1569	2.84	0.2390	1.06	0.2260	2.87	0.1680	3.33	0.2234	1.39	0.1842	2.08	0.1727	2.11	0.0579	0.37	0.0567	0.53	0.0585	0.54
BPD	0.2820	0.96	0.1509	2.47	0.1562	2.28	0.1567	1.76	0.1254	2.88	0.1226	2.97	0.2567	1.43	0.2126	2.80	0.2284	3.04	0.2522	0.60	0.1965	1.95	0.1977	2.05	0.0719	0.30	0.0618	0.59	0.0606	0.60
TV	0.2878	0.75	0.1561	2.66	0.1581	2.59	0.1959	1.52	0.1125	3.56	0.1125	3.24	0.2830	1.47	0.1988	3.90	0.1988	3.31	0.2407	0.63	0.1768	2.24	0.1912	2.11	0.0689	0.33	0.0516	0.97	0.0529	0.85
NLM	0.2181	0.85	0.1338	3.23	0.1738	2.16	0.1628	1.16	0.0905	4.38	0.1569	2.84	0.1944	1.15	0.2226	3.68	0.1681	3.36	0.2521	1.07	0.1614	2.50	0.1727	2.11	0.1016	0.40	0.0526	0.59	0.0583	0.56
HT	0.2250	1.37	0.1462	3.02	0.1737	2.18	0.1444	1.90	0.1165	3.48	0.1566	2.87	0.2568	1.02	0.2236	3.23	0.1673	3.37	0.2085	1.31	0.1768	2.38	0.1810	2.64	0.0640	0.28	0.0543	0.50	0.0547	0.50
ST	0.2678	1.07	0.1466	3.35	0.1738	2.16	0.1594	1.87	0.1226	3.14	0.1569	2.84	0.3386	1.06	0.2256	2.90	0.1680	3.33	0.2472	1.17	0.1727	2.59	0.1841	2.60	0.0644	0.32	0.0544	0.48	0.0547	0.49
TR (limited)	0.2148	1.66	0.1886	2.15	0.1902	2.16	0.1476	2.16	0.1483	2.44	0.1493	2.45	0.2261	1.72	0.2464	2.06	0.2418	2.06	0.2212	1.22	0.2119	1.45	0.2073	1.45	0.0545	0.61	0.0522	0.65	0.0521	0.65
TR (full)	0.1453	3.62	0.0895	4.98	0.0926	5.00	0.1135	6.13	0.0750	7.06	0.0721	7.08	0.0738	7.59	0.0478	12.87	0.0477	12.99	0.1896	2.80	0.1247	3.46	0.1267	3.46	0.0291	6.09	0.0502	7.14	0.0491	7.15
Guided filter (LTH)	0.2307	1.83	0.2423	2.00	0.2134	2.00	0.1440	2.36	0.1445	2.44	0.1464	2.44	0.2559	1.47	0.2469	1.58	0.2487	1.58	0.2223	1.56	0.2603	1.78	0.2509	1.78	0.0565	0.48	0.0574	0.48	0.0567	0.48
Guided filter (LTO)	0.2334	1.67	0.1329	3.84	0.0756	3.81	0.1352	2.43	0.0833	5.71	0.0591	7.51	0.2342	1.24	0.1976	3.96	0.0709	8.53	0.2222	1.56	0.1600	2.93	0.1119	4.13	0.0582	0.41	0.0558	0.63	0.0553	0.66
Guided filter (TV)	0.1950	1.57	0.0859	4.77	0.1030	4.67	0.1493	3.27	0.0668	6.54	0.0823	5.86	0.1539	3.27	0.1189	6.06	0.1596	5.07	0.1930	1.14	0.1350	3.21	0.1629	2.95	0.0483	1.33	0.0383	2.92	0.0364	2.15

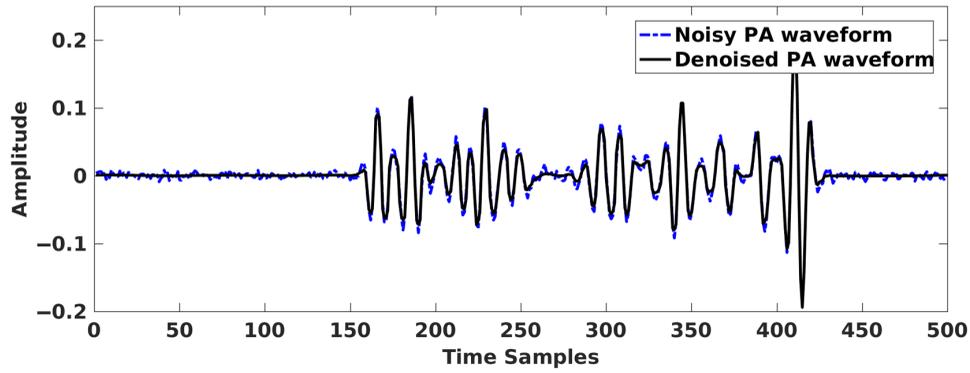


Fig. 5 The reconstructed initial pressure distribution for the BV phantom [Fig. 3(a)] with SNR of data being 40 dB for a single detector located at 12 'o' clock position. The dashed line represents the noisy PA signal and the solid black line represents the denoised PA signal using automated wavelet denoising method (Sec. 2.9). The reconstructed images are shown in Fig. 6.

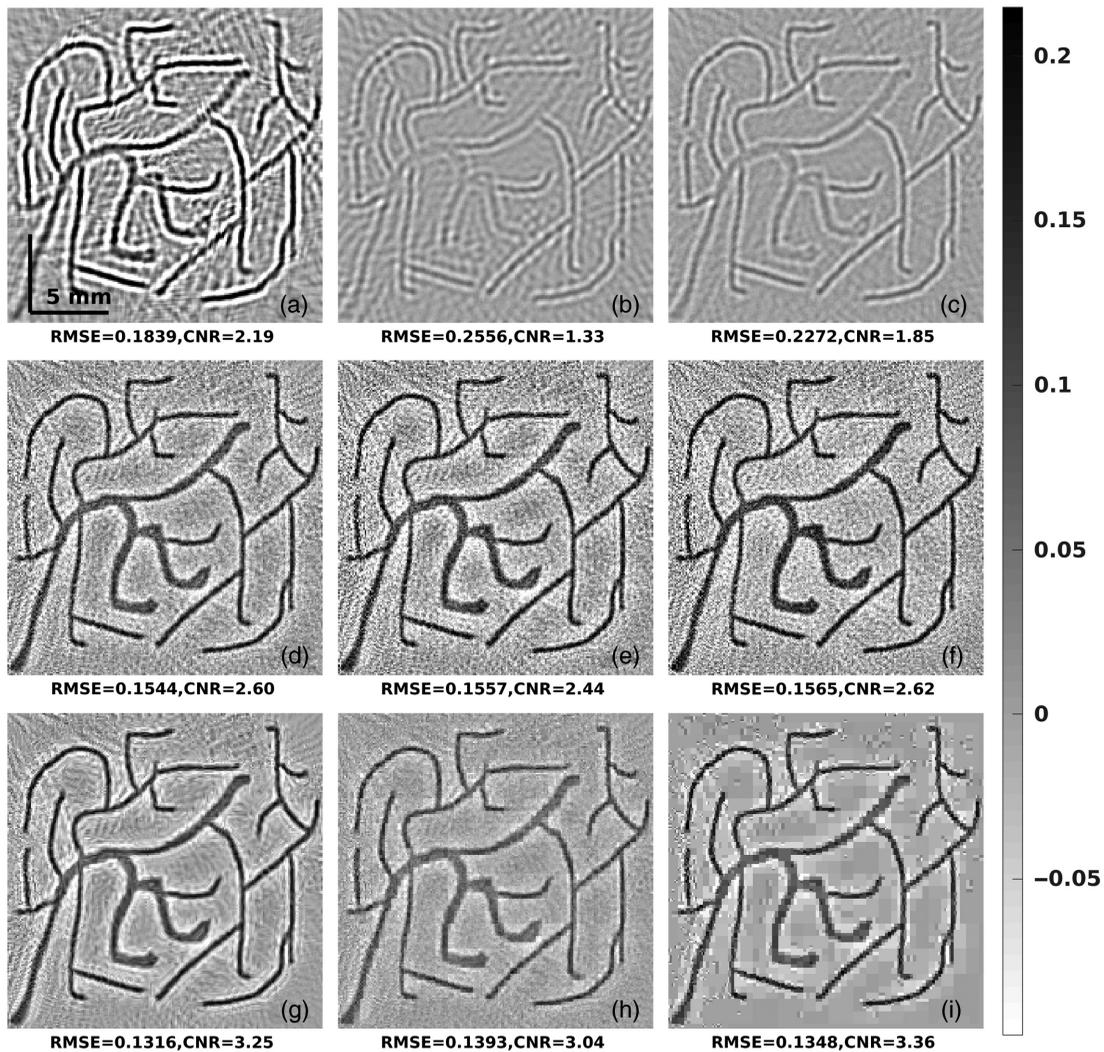


Fig. 6 The reconstructed initial pressure distribution using denoised (Sec. 2.9) PA data corresponding to numerical BV phantom using (a) k-wave TR, (b) LBP, (c) LTH, (d) LTO, and (e) BPD with input image being (c), (f) TV, (g) NLM, wavelet, (h) ST, and (i) HT denoising with input image being (d). The target/expected distribution is given in Fig. 3(a). The corresponding figures of merit—RMSE and CNR—are given below each figure.

line) is 0.114. The computational time required for denoising the recorded signals from 100 detectors is 0.3 s. These denoised PA signals were utilized as input to the reconstruction methods discussed in this work and results of the same are given in Fig. 6. The figures of merit—RMSE and CNR—are also reported in the same figure below each image. Note that the corresponding results with utilization of recorded raw data are given in Fig. 4 and figures of merit in Table 2. The improvement of utilizing the wavelet denoising on the raw data is less than 2% in all cases in terms of figures of merit and in comparison to the proposed guided filter approach. Note that for the data having SNR of 20 dB, the improvement is as high as 25% (results are not shown). It is also important to note that there are sophisticated methods proposed in the literature,^{34,35} which could improve the reconstruction performance significantly (latest one being based on deep learning³⁵), but the computational complexity of these methods is at least one order higher than the wavelet denoising method utilized here, which has the computational complexity similar to the proposed guided filter approach.

To understand changes in the reconstructed image frequency spectrum with the utilization of the guided filtering approach proposed here, the frequency spectrum plots pertaining to these along with input and guided images are shown in Fig. 7. The center of these plots represents the low spatial frequencies and the edges of these will provide high-frequency components. One observation from these plots is that the guided filtering approach significantly alters the frequency spectrum of both input and guiding image to give an output that is an improved version. More importantly, as the noise is known to be existing in the high frequency components, it is able to eliminate noise effectively. In these cases, the Pearson correlation^{88,89} (PC) between (a) and others was utilized to correlate these spectrums, which is known to be invariant to scale in the frequency spectrums to provide required normalization. The PC values given in Fig. 7 indicate that the guided filter applied to

x_{TV} was having the highest correlation, thus in turn, giving the closest approximation to the target/expected initial pressure distribution.

To study the effect of parameters utilized in the guided filter, patch size, and epsilon (ϵ) in Algorithm 2, a study was performed to record CNR by varying them. These results are reported in Fig. 8 for the case of the BV phantom. The patch size defines the neighborhood around the pixel used in the guided filtering. The ϵ defines the degree of smoothing (a threshold value on the variance). If a small value is specified, only neighborhoods with small variance will be smoothed and neighborhoods with large variance will not be smoothed. If a large value is specified, the neighborhoods with large variance will also get smoothed. The values of α and β in Algorithm 2 were chosen to be 1.05. From the results presented in Fig. 8, it is obvious that a patch size of 1 and ϵ of 1×10^{-3} provide the optimal results, these values were utilized for obtaining the results presented in this work.

The figures of merit—RMSE and CNR—for the cases of data with SNR of 20 and 60 dB (reconstructed images are not shown here) were compiled and listed in Table 2. Even for these results, it is obvious that the guided filtering approach provides the best performance with input image being either x_{LTH} or x_{LTO} . The computational time recorded for obtaining these results was in the similar orders as given in Table 1. It is also important to note that the computational time required for obtaining x_{LTO} is at least five times more compared to x_{LTH} with x_{TV} taking the least amount of computational time among the input images for the guided filter (refer to Table 1). The performance of x_{TV} being the input image for the guided filter compared to other input images (x_{LTH} or x_{LTO}) is comparable with the added advantage of computation of x_{TV} being efficient.

The reconstruction results from the similar effort utilizing data from 100 detectors with SNR being 40 dB for the modified

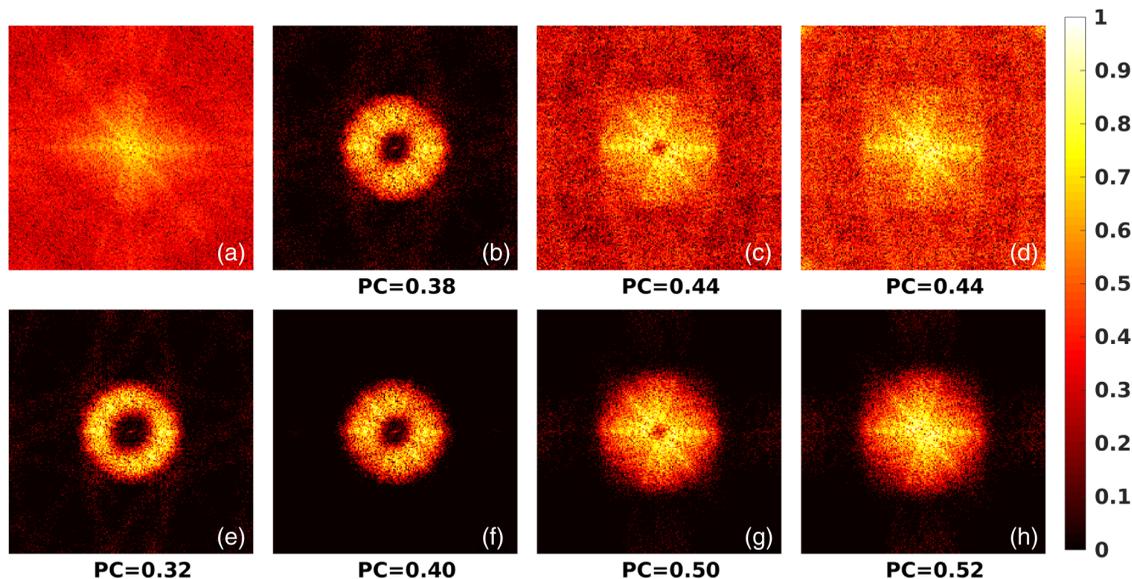


Fig. 7 The logarithm of amplitude of Fourier spectrum of results pertaining to (a) target BV phantom [spatial distribution is given in Fig. 3(a)], (b) LTH [spatial distribution is given in Fig. 4(c)], (c) LTO [spatial distribution is given in Fig. 4(d)], (d) TV [spatial distribution is given in Fig. 4(f)], and (e) LBP [spatial distribution is given in Fig. 4(b)]. The guided filter results are given in (f), (g), and (h) corresponding to spatial distribution given in Figs. 4(h), 4(j), and 4(k). The figure of merit, PC computed with (a) being the target, is given correspondingly below each image.

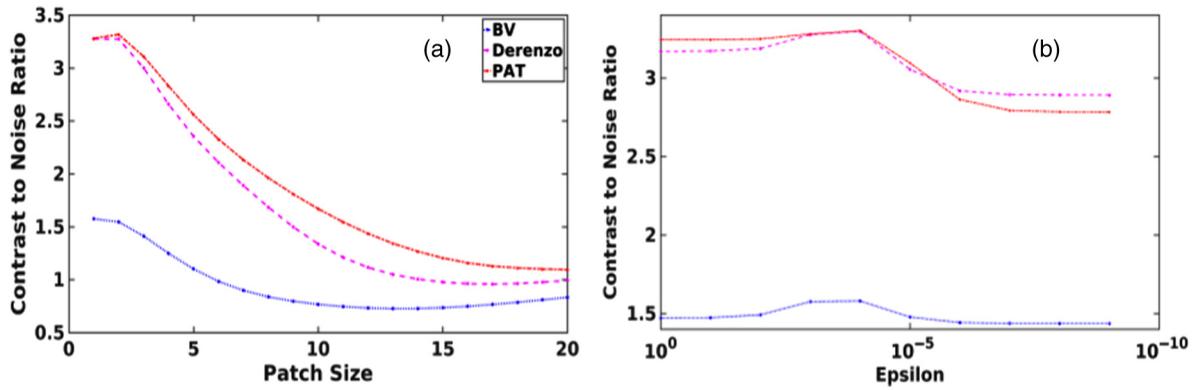


Fig. 8 The plot of figure of merit, CNR as a function of (a) patch size and (b) epsilon (ϵ) for three numerical phantoms considered in this work [given in the legend of (a)]. These results correspond to guided filter results with TV reconstructed image being input and LBP being guiding image. The patch size was varied for obtaining (a) with ϵ being constant (1×10^{-3}) and the ϵ was varied for obtaining (b) with patch size being constant (one).

Derenzo phantom are provided in Fig. 9. The reconstructed initial pressure distribution using TR and LBP (x_{LBP}) are shown in Figs. 9(a) and 9(b), respectively. The reconstruction using Lanczos-Tikhonov-Heuristic (LTH), Lanczos Tikhonov optimal (LTO), BPD, and TV is shown in Figs. 9(c)–9(f). The guided filtering approach results for the input image being LTH,

LTO, and TV are shown in Figs. 9(j)–9(l), respectively. The denoised results of LTO using NLM and wavelet (ST and HT) thresholding methods are given in Figs. 9(g)–9(i), respectively. The required computational time for obtaining these results is similar to the values given in Table 1. Similar to the BV phantom, the figures of merit corresponding to these

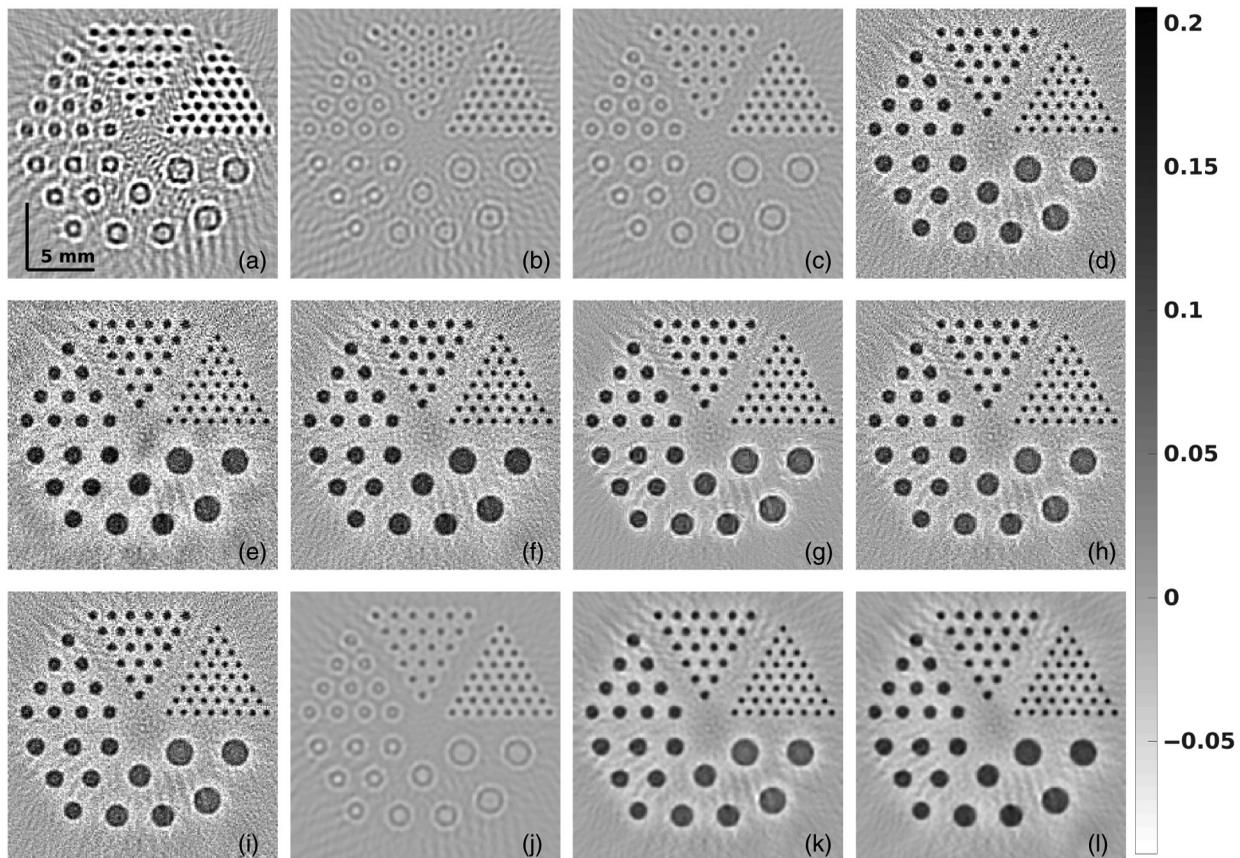


Fig. 9 The reconstructed initial pressure distribution using 100 detectors data with SNR of 40 dB corresponding to numerical Derenzo phantom using (a) k-wave TR, (b) LBP, (c) LTH, (d) LTO, and (e) BPD with input image being (c), (f) TV, (g) NLM, wavelet, (h) ST, and (i) HT denoising with input image being (d). The proposed guided filter results with input images being (c), (d), and (f) are given in (j), (k), and (l), respectively, with (b) being the guiding image. The target/expected distribution is given in Fig. 3(b). The corresponding figures of merit—RMSE and CNR—are given in Table 2.

results are given in Table 2 along with results pertaining to SNR of data being 20 and 60 dB. A similar trend as observed in the BV phantom case was observed here with the guided filtering approach performing superior to all other approaches including BPD. The improvement obtained for the guided filtering approach for the SNR of 40 dB with x_{TV} as the input images were 40% for RMSE and 83% for the CNR compared to the TV case. Similarly, the improvement obtained for SNR of 60 dB was 47% for RMSE and 131% compared to the TV method, which provides best results without filtering. As observed in the earlier case, the guided filtering approach with x_{TV} as the input provides the optimal performance in terms of improvement in figures of merit as well as required computational time. It should also be observed that bipolar (donut shaped objects in place of filled circles) objects in TR [Fig. 9(a)], LBP [Fig. 9(b)], and LTH [Fig. 9(c)] are mainly due to the limited bandwidth of the boundary data and in case of utilizing full bandwidth data, these bipolar objects become unipolar [see Fig. 3(f)].

Another phantom consisting of letters “PAT” was also considered to compare the performance of the discussed methods in terms of recovering sharp features. These results corresponding to SNR of 40 dB (100 detectors) were given in Fig. 10. The TR and LBP reconstruction results are given in Figs. 10(a) and 10(b) correspondingly. The results from LTH, LTO, BPD, and TV are shown in Figs. 10(c)–10(f) correspondingly. The proposed

guided filter approach results with input images being (c), (d), and (f) are, respectively, presented in (j)–(l). The denoised results of LTO using NLM and wavelet (ST and HT) thresholding methods are given in Figs. 10(g)–10(i), respectively. The figures of merit, similar to earlier cases, including results from data with SNR of 20 and 60 dB, are compiled in Table 2. The computational times observed were similar to those reported in Table 1. The improvement obtained in RMSE and CNR using the guided filter with TV results as the input image as compared to the TV method was 40% and 55%, respectively for SNR of 40 dB. For SNR of 60 dB, the improvement was 64% and 157% for RMSE and CNR, respectively, compared to the TV method, which gives the best result for the SNR of 60 dB. As observed earlier, the guided filter with x_{TV} as the input image provides the optimal performance in terms of improvement in figures of merit as well as required computational time.

To assess the observed trends with sparse data, having only 50 detectors, the experiments were repeated on the BV phantom. The results pertaining to this experiment with SNR of data being 40 dB are shown in Fig. 11. These results are more prone to streak artifacts as the data available is half compared to the ones presented in Fig. 4. The TR method in this case has significant aliasing artifacts as results presented in Fig. 4. Even the denoised results using standard methods (NLM and wavelet thresholding) show significant artifacts compared to the guided

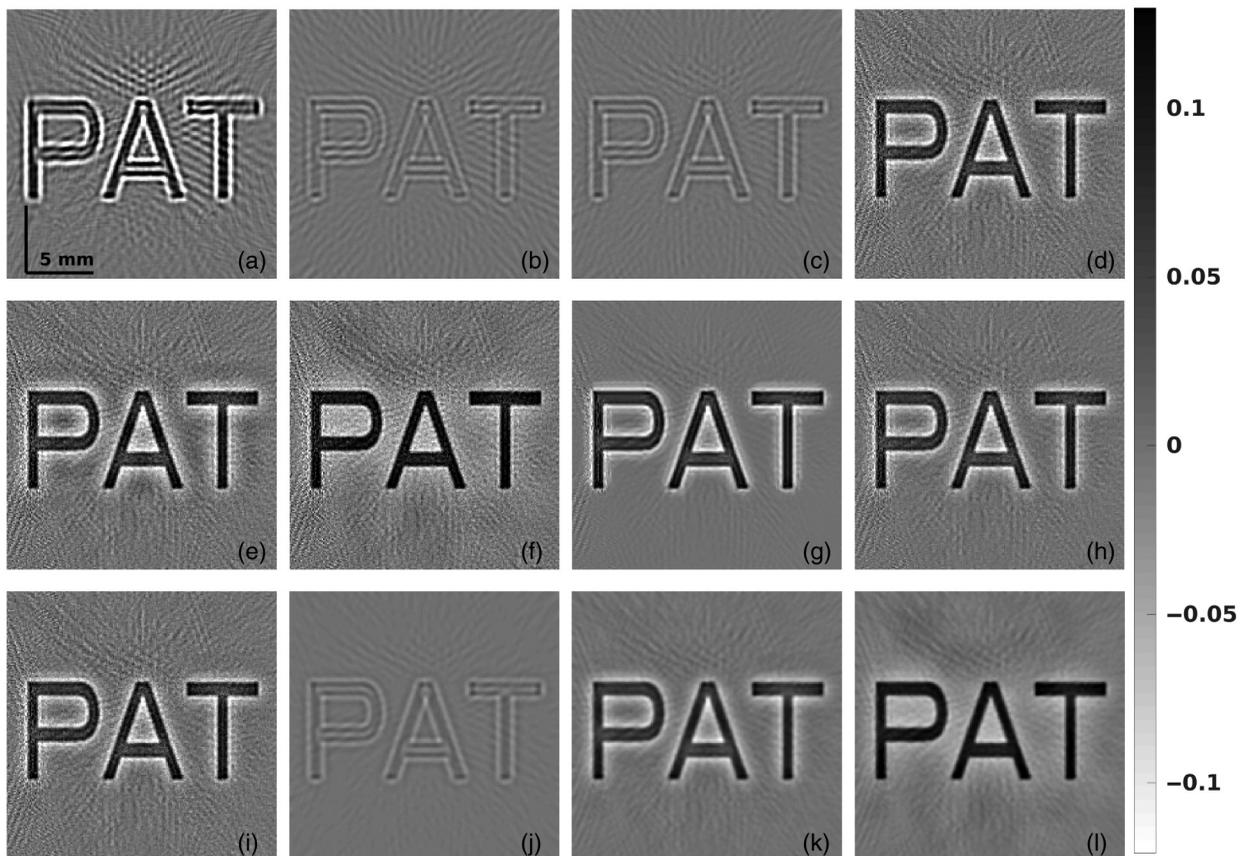


Fig. 10 The reconstructed initial pressure distribution using hundred detectors data with SNR of 40 dB corresponding to numerical “PAT” phantom using (a) k-wave TR, (b) LBP, (c) LTH, (d) LTO, and (e) BPD with input image being (c), (f) TV, (g) NLM, wavelet, (h) ST, and (i) HT denoising with input image being (d). The proposed guided filter results with input images being (c), (d), and (f) are given in (j), (k), and (l), respectively, with (b) being the guiding image. The target/expected distribution is given in Fig. 3(c). The corresponding figures of merit—RMSE and CNR—are given in Table 2.

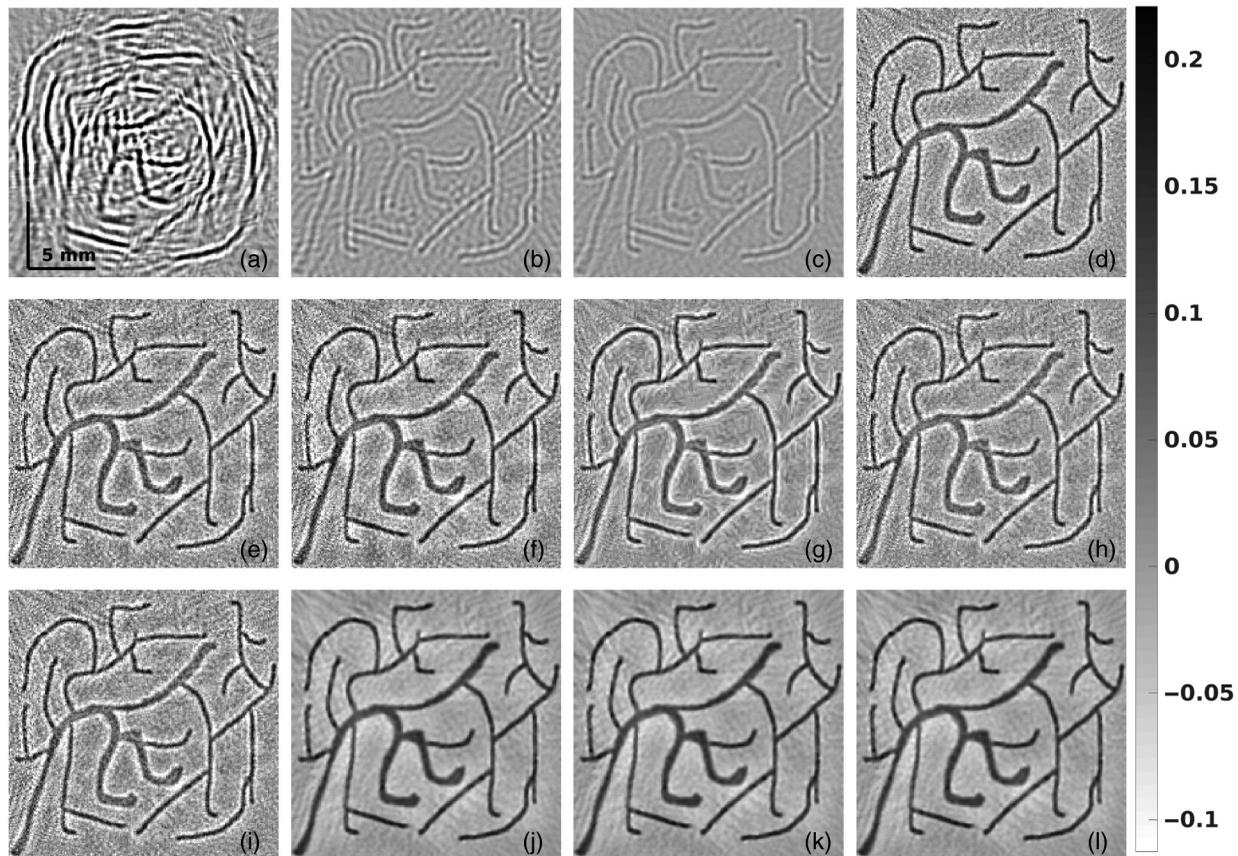


Fig. 11 The reconstructed initial pressure distribution using fifty detectors data with SNR of 40 dB corresponding to numerical BV phantom using (a) k-wave TR, (b) LBP, (c) LTH, (d) LTO, and (e) BPD with input image being (c), (f) TV, (g) NLM, wavelet, (h) ST, and (i) HT denoising with input image being (d). The proposed guided filter results with input images being (c), (d), and (f) are given in (j), (k), and (l), respectively, with (b) being the guiding image. The target/expected distribution is given in Fig. 3(a). The corresponding figures of merit—RMSE and CNR—are given in Table 2.

filter results. Similar to earlier cases, the RMSE and CNR were compiled for these results along with SNR of data being 20 and 60 dB in Table 2. The results show similar trends as observed earlier, with the guided filtering approach providing the best results with input the image being x_{TV} providing optimal results. This experiment also showed that the guided filtering approach can work well in the sparse data cases as well without adding any significant computational burden.

The reconstruction results pertaining to the realistic breast phantom [Fig. 3(d)] corresponding to the PA data with SNR of 40 dB (100 detectors) are given in Fig. 12. The TR and LBP reconstruction results are given in Figs. 12(a) and 12(b) correspondingly. The results from LTH, LTO, BPD, and TV are shown in Figs. 12(c)–12(f), respectively. The proposed guided filter approach results with input images being (c), (d), and (f) are, respectively, presented in (j)–(l). The denoised results of LTO using NLM and wavelet (ST and HT) thresholding methods are given in Figs. 12(g)–12(i), respectively. The figures of merit, similar to earlier cases, including results from data with SNR of 20 and 60 dB, were compiled in Table 2. The computational times observed were similar to those reported in Table 1. The improvement obtained in RMSE and CNR using the guided filter with TV results as the input image as compared to the TV method was 26% and 201%, respectively, for SNR of 40 dB. For SNR of

60 dB, the improvement was 31% and 152% for RMSE and CNR, respectively, compared to the TV method, which gives the best result for the SNR of 60 dB. These results, which were performed on a more realistic phantom, with more than two grayscale values of initial pressure show the improved performance with the guided filtering approach, showing its utility in real time. It should also be noted that, overall, the figures of merit are lower for this case compared to the remaining numerical phantom cases as the structures in these case are heterogeneous and complex.

The reconstructed results for the experimental horse-hair phantom are shown in Fig. 13. The LBP result is shown as Fig. 13(a) with LTO and TV results being Figs. 13(b) and 13(c), respectively. The guided filtering approach results using x_{LTO} (b) and x_{TV} (c) as the input images are given in Figs. 13(d) and 13(e) correspondingly. Note that the results pertaining to LTH and BPD were not considered here, as they were proven to be not as effective as LTO and TV in numerical phantom studies. Similarly, the standard denoising methods were also not utilized as they were proven to be inferior compared to the guided filter results in the numerical phantom cases. The figure of merit, SNR, for these reconstruction results is given at the bottom of each result. The improvement in terms of SNR, with the guided filter is at least 5.5 dB with the guided filter with LTO [Fig. 13(d)], as the input image showing the best

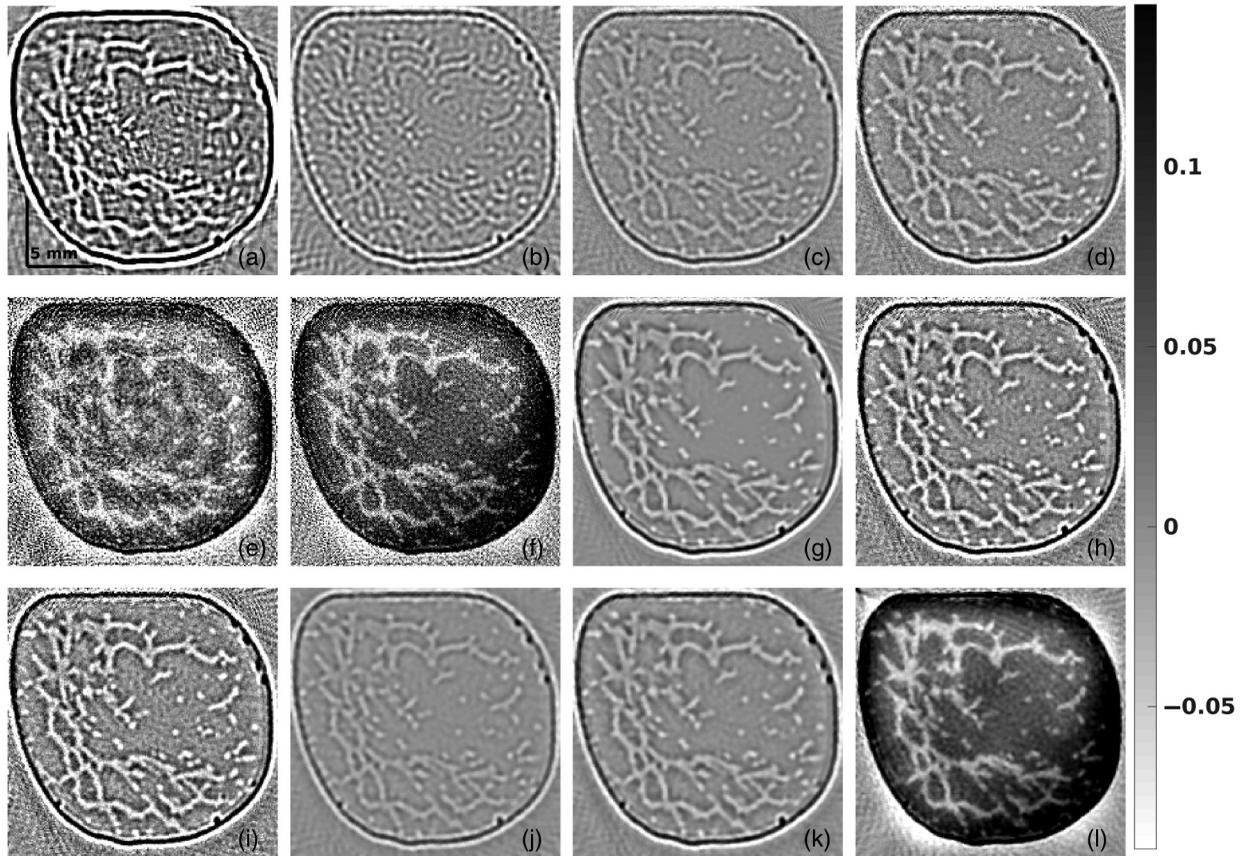


Fig. 12 The reconstructed initial pressure distribution using hundred detectors data with SNR of 40 dB corresponding to realistic breast phantom using (a) k-wave TR, (b) LBP, (c) LTH, (d) LTO, and (e) BPD with input image being (c), (f) TV, (g) NLM, wavelet, (h) ST, and (i) HT denoising with input image being (d). The proposed guided filter results with input images being (c), (d), and (f) are given in (j), (k), and (l), respectively, with (b) being the guiding image. The target/expected distribution is given in Fig. 3(d). The corresponding figures of merit—RMSE and CNR—are given in Table 1.

performance with improvement of 10.6 dB. As TV is computationally efficient, the guided filter result with TV as the input image provides optimal performance.

Results from similar efforts for the case of the experimental tube phantom are shown in Fig. 14. As observed earlier, the guided filter approach provides the best performance in terms of SNR_r of the reconstructed image. In the guided filter approach, the result pertaining to the input image being x_{LTO} is superior compared to the other result (SNR_r improvement of 9.38 dB). Even though the original result without the guided filter shows suboptimal performance compared to others [comparing Fig. 14(b) with 14(a) and 14(c)], the guided filter improves the Fig. 14(b) by at least 9.38 dB [Fig. 14(d)] in terms of observed SNR_r .

The results pertaining to the *in-vivo* rat brain imaging to show the effectiveness of the proposed method in preclinical imaging were presented in Fig. 15. The LBP reconstruction is shown in Fig. 15(a) and the reconstructions using LTO and TV are shown in Figs. 15(b) and 15(c), respectively. Similar to the previous observations, the guided filter outperforms others in terms of SNR_r of the reconstructed image. Among the guided filter reconstructions, the result obtained using the input image being x_{LTO} is superior compared to the other result (SNR_r improvement of 11.23 dB).

Overall, the guided filter approach has shown in all numerical, experimental, and *in-vivo* cases that it provides superior

performance. Specifically, the observed figures of merit for the output (filtered) image are exceeding the input (to be filtered) and guiding image figures of merit (Table 2, Figs. 13–15). This process adds little to no computational time as reported in Table 1, making it highly attractive. Also note that the guiding image in all cases discussed here was chosen to be the LBP image (x_{LBP}), which is used as an input image to the TV-regularized method. Thus, the guiding image is readily available after the reconstruction process using TV-regularized solution.

It is also important to note that it is possible to use other images (such as using x_{LTO}) as the guiding image rather than x_{LBP} , our experience shows that the improvement in these cases is minimal. The guided filtering approach provides optimal performance, when different characteristics in the input and guiding image have to be combined to provide a filtered image. The BP image being more smooth and x_{LTO} or x_{TV} promoting sharp edges due to their sparse recovery nature, this combination leads to superior results. The frequency spectrum presented in Fig. 7 gives an insight into this choice of x_{LBP} being the guiding image as its frequency spectrum is complimentary to others, such as x_{LTO} and x_{TV} . This method becomes ineffective if the frequency spectrum of both input and guiding images is similar (like x_{LTO} and x_{TV}). Moreover, if the choice is not proper, like other explicit filters, the guided filter can exhibit artifacts of unwanted smoothing near edges.⁴¹

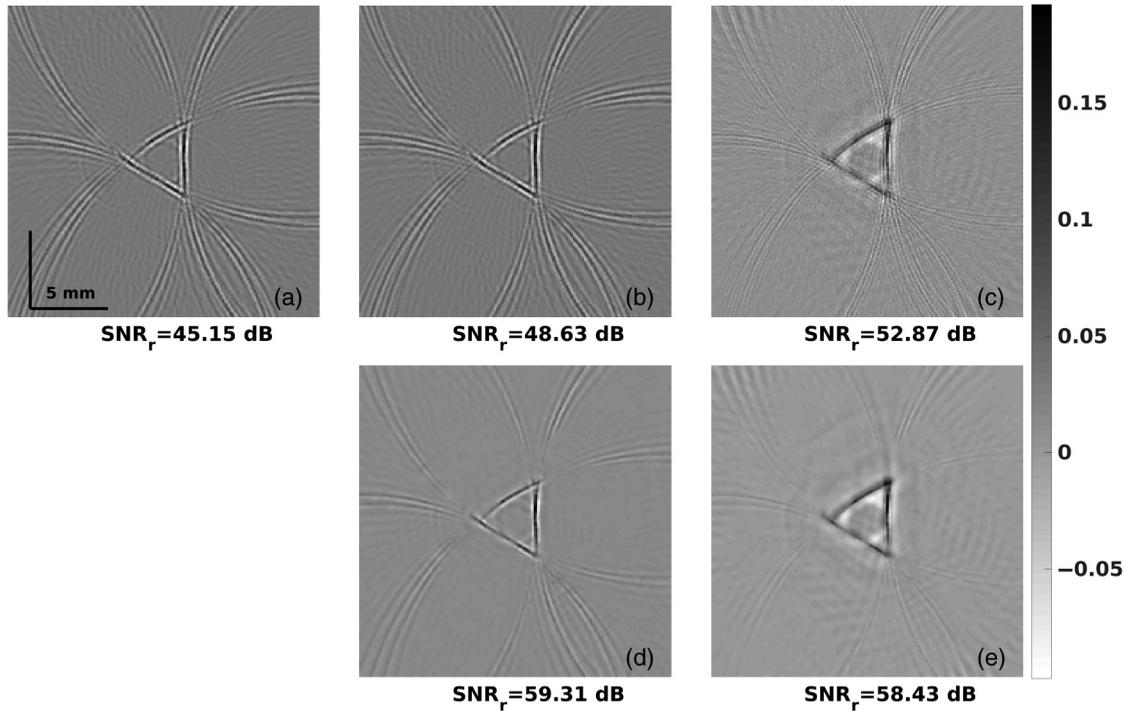


Fig. 13 The reconstructed initial pressure distribution with experimental horse hair phantom data using (a) LBP, (b) LTO, and (c) TV regularization. The guided filter results with input image being (b) and (c) are correspondingly given in (d) and (e). The photograph of the horse hair phantom is shown in Fig. 2(c). The figure of merit, SNR_r (in dB), of each reconstructed image is correspondingly given below. The computational time required for generating (a), (b), (c), (d), and (e) is 0.5, 11.74, 4.10, 0.01, and 0.01 s, respectively.

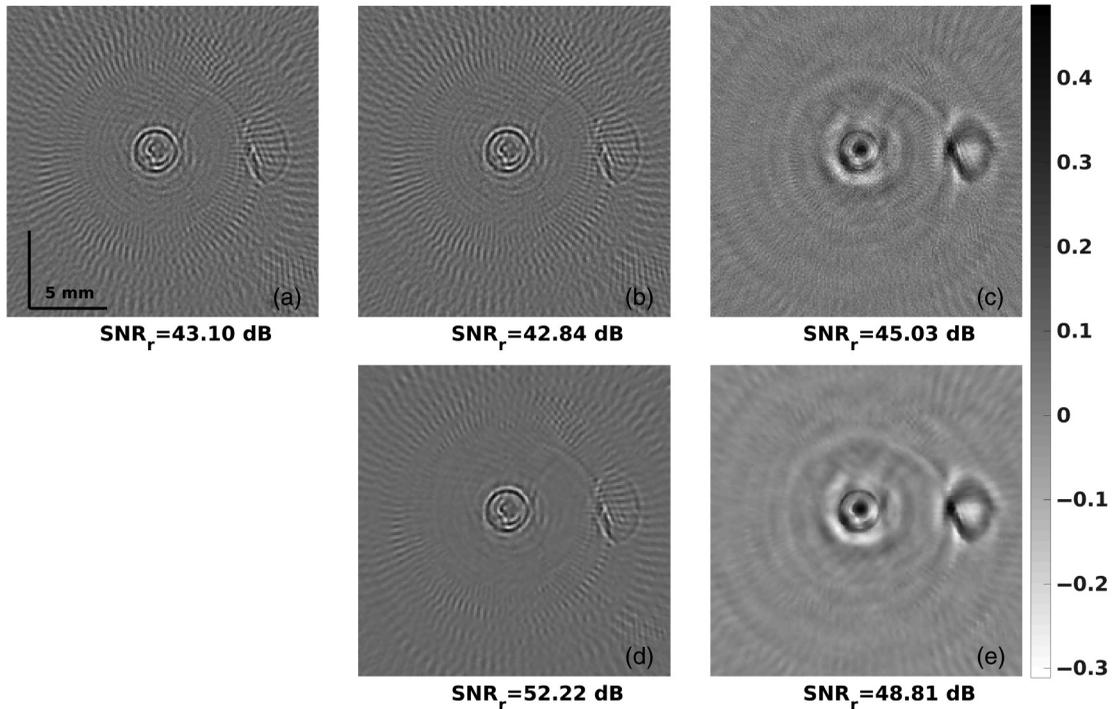


Fig. 14 The reconstructed initial pressure distribution with experimental tube phantom data using (a) LBP, (b) LTO, and (c) TV regularization. The guided filter results with input image being (b) and (c) are correspondingly given in (d) and (e). The photograph of the ink tubes is shown in Fig. 2(b). The figure of merit, SNR_r (in dB), of each reconstructed image is correspondingly given below. The computational time required for generating (a), (b), (c), (d), and (e) is 0.5, 10.22, 9.10, 0.01, and 0.01 s, respectively.

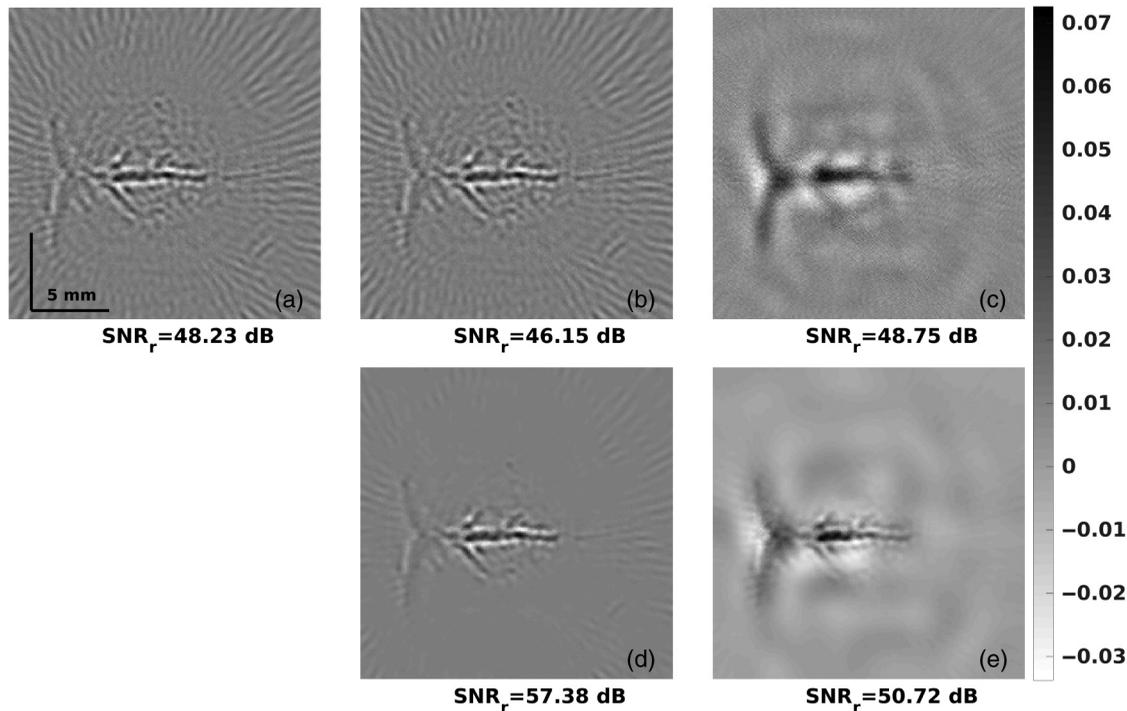


Fig. 15 The reconstructed initial pressure distribution of *in-vivo* rat brain (lateral plane) using (a) LBP, (b) LTO, and (c) TV regularization. The guided filter results with input image being (b) and (c) are correspondingly given in (d) and (e). The photograph of the rat brain with and without skin is shown in Figs. 2(d) and 2(e), respectively. The figure of merit, SNR_r (in dB), of each reconstructed image is correspondingly given below. The computational time required for generating (a), (b), (c), (d), and (e) is 0.5, 10.55, 6.68, 0.01, and 0.01 s, respectively.

In this work, the BP image, which is typically used as an initial guess for most model-based iterative reconstruction algorithms, was utilized as the guiding image. This may not be an optimal choice and exploration of optimal choice of guide and/or input image along with development of automated methods for an optimal choice will be taken up as future extension to this work.

It may be feasible to utilize a model-free solution either as a guiding or input image, such as delay-and-sum,³⁰ but these tend to have arbitrary units for the initial pressure requiring normalization for each case making the comparison of reconstruction results among discussed methods difficult. Also, for this work, the limited data refer to the data availability in comparison to the state-of-the-art clinical PAT setups. The current PAT scanners with advances in instrumentation are capable of acquiring boundary data with the number of detectors being 512 and number of time samples for each detector being 2048.³⁶ Here, the maximum number of detectors utilized was 100 (maximum time samples were 512), making the boundary data available for image reconstruction still limited compared to these setups.

The figures of merit, such as RMSE and CNR, were utilized extensively in the literature to assess the performance of the reconstructed image quality in PA imaging.^{36,77,78,81} The task-based image quality metrics can better assess the reconstructed image quality,^{91,92} but use of RMSE and CNR as metrics has been a common practice across tomographic problems.^{36,77,78,81} As the main aim of this work has been to introduce a image-guided filter as a fast postprocessing step to improve the image characteristics of the reconstructed image, which was clearly demonstrated through numerical and experimental phantom as well as *in-vivo* cases, the other

metrics were not considered. The figures of merit—RMSE and CNR—are also robust to different realizations of noise. For example, the results presented in Table 2 differed by less than 1% within 10 noise realizations for obtaining the required SNR of boundary data.

The bipolar nature of the objects (even though expected in uni-polar) in the reconstructed images (Figs. 4 and 9–15) can be corrected using the Hilbert transform (which essentially eliminates the negative initial pressure values).^{82,93,94} Utilization of this might leads to inaccurate representation of the reconstructed images, overshadowing the performance of postprocessing methods utilized here (including wavelet thresholding, NLM, and the proposed guided filter). As the dynamic range of the reconstructed image can be utilized as a metric for knowing the performance of algorithms used in this work, these transforms that alter the raw performance of these methods were not attempted here.

The postprocessing reconstructed images to improve their utility are common among other imaging modalities, such as emission computed tomography.³⁷ The guided filter approach, even though it is a postprocessing step, attempts to filter the input image using another image as a guiding image. Here, this guiding image is chosen to be the BP image, which is also the initial guess for the iterative reconstruction methods. The methods of this type are very simple to integrate into the iterative image reconstruction framework of PAT, as the computational complexity is insignificant, and can work with any set of images in terms of improving them. This work showed that standard denoising filters, such as wavelet thresholding as well as NLM based ones, have limited utility in PA imaging. The proposed guided filter can also be seen as an adaptive sharpening

filter, which increases the local contrast with removal of noise based on guiding image.^{95,96} The reference sharpness level of the guiding image acts as edge preserving to result in this adaptive sharpening (see Fig. 7). Approaches such as the proposed guided filter that improves the reconstructed images are of high value as they leverage the best features of the input and guiding images to provide high quality filtered image. The developed algorithms were provided as an open source⁹⁷ for enthusiastic users.

6 Conclusions

The PA image reconstruction in case of limited data is ill-conditioned, often necessitating the usage of regularization to provide meaningful results. There are smooth and nonsmooth regularization schemes that are available, which can promote different characteristics in the reconstructed images. A simple guided filtering approach was proposed here that combines the best features in the input and guiding images (obtained from different reconstruction methods). Utilizing both numerical and experimental data, it was shown that the guided filtering approach improves the reconstructed images substantially (as high as 11.23 dB in terms of SNR of reconstructed image) with very little to no computational burden. It was also shown that the performance of the proposed guided filter approach even with sparse data (half of the original) is superior compared to others. As this approach can be applied as a post-processing step for the image reconstruction, it is easily integratable into the existing reconstruction algorithmic framework.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

References

- L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**(6075), 1458–1462 (2012).
- M. Pramanik et al., "Design and evaluation of a novel breast cancer detection system combining both thermoacoustic (TA) and photoacoustic (PA) tomography," *Med. Phys.* **35**(6), 2218–2223 (2008).
- L. V. Wang, "Tutorial on photoacoustic microscopy and computed tomography," *IEEE J. Sel. Top. Quantum Electron.* **14**(1), 171–179 (2008).
- P. K. Upputuri and M. Pramanik, "Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review," *J. Biomed. Opt.* **22**(4), 041006 (2017).
- L. Li et al., "Single-impulse panoramic photoacoustic computed tomography of small-animal whole-body dynamics at high spatiotemporal resolution," *Nat. Biomed. Eng.* **1**(5), 0071 (2017).
- Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *J. Biomed. Opt.* **21**(6), 061007 (2016).
- A. E. Cerussi et al., "Sources of absorption and scattering contrast for near-infrared optical mammography," *Acad. Radiol.* **8**(3), 211–218 (2001).
- S. Srinivasan et al., "Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured in vivo by near-infrared breast tomography," *Proc. Natl. Acad. Sci. U. S. A.* **100**(21), 12349–12354 (2003).
- K. Jansen et al., "Intravascular photoacoustic imaging of human coronary atherosclerosis," *Opt. Lett.* **36**(5), 597–599 (2011).
- S. Jiao et al., "Photoacoustic ophthalmoscopy for in vivo retinal imaging," *Opt. Express* **18**(4), 3967–3972 (2010).
- S. A. Ermilov et al., "Laser photoacoustic imaging system for detection of breast cancer," *J. Biomed. Opt.* **14**(2), 024007 (2009).
- M. Heijblom, W. Steenbergen, and S. Manohar, "Clinical photoacoustic breast imaging: the Twente experience," *IEEE Pulse* **6**(3), 42–46 (2015).
- S. J. Ford et al., "Structural and functional analysis of intact hair follicles and pilosebaceous units by volumetric multispectral photoacoustic tomography," *J. Invest. Dermatol.* **136**(4), 753–761 (2016).
- J.-T. Oh et al., "Three-dimensional imaging of skin melanoma in vivo by dual-wavelength photoacoustic microscopy," *J. Biomed. Opt.* **11**(3), 034032 (2006).
- J. Yao et al., "High-speed label-free functional photoacoustic microscopy of mouse brain in action," *Nat. Methods* **12**(5), 407–410 (2015).
- K. Wang et al., "Investigation of iterative image reconstruction in three-dimensional photoacoustic tomography," *Phys. Med. Biol.* **57**(17), 5399–5423 (2012).
- A. Rosenthal, V. Ntziachristos, and D. Razansky, "Acoustic inversion in photoacoustic tomography: a review," *Curr. Med. Imaging Rev.* **9**(4), 318–336 (2013).
- C. Huang et al., "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE Trans. Med. Imaging* **32**(6), 1097–1110 (2013).
- K. Wang et al., "An imaging model incorporating ultrasonic transducer properties for three-dimensional photoacoustic tomography," *IEEE Trans. Med. Imaging* **30**(2), 203–214 (2011).
- Y. Xu et al., "Reconstructions in limited-view thermoacoustic tomography," *Med. Phys.* **31**(4), 724–733 (2004).
- S. R. Arridge et al., "On the adjoint operator in photoacoustic tomography," *Inverse Probl.* **32**(11), 115012 (2016).
- S. Arridge et al., "Accelerated high-resolution photoacoustic tomography via compressed sensing," *Phys. Med. Biol.* **61**(24), 8908–8940 (2016).
- X. L. Dean-Ben, V. Ntziachristos, and D. Razansky, "Acceleration of photoacoustic model-based reconstruction using angular image discretization," *IEEE Trans. Med. Imaging* **31**(5), 1154–1162 (2012).
- X. L. Dean-Ben et al., "Accurate model-based reconstruction algorithm for three-dimensional photoacoustic tomography," *IEEE Trans. Med. Imaging* **31**(10), 1922–1928 (2012).
- A. Buehler et al., "Model-based photoacoustic inversions with incomplete projection data," *Med. Phys.* **38**(3), 1694–1704 (2011).
- S. Gutta et al., "Modeling errors compensation with total least squares for limited data photoacoustic tomography," *IEEE J. Sel. Top. Quantum Electron.* in press (2019).
- G. Paltauf et al., "Iterative reconstruction algorithm for photoacoustic imaging," *J. Acoust. Soc. Am.* **112**(4), 1536–1544 (2002).
- J. Prakash et al., "Basis pursuit deconvolution for improving model-based reconstructed images in photoacoustic tomography," *Biomed. Opt. Express* **5**(5), 1363–1377 (2014).
- C. B. Shaw et al., "Least squares QR-based decomposition provides an efficient way of computing optimal regularization parameter in photoacoustic tomography," *J. Biomed. Opt.* **18**(8), 080501 (2013).
- N. Awasthi et al., "Vector extrapolation methods for accelerating iterative reconstruction methods in limited-data photoacoustic tomography," *J. Biomed. Opt.* **23**(7), 071204 (2018).
- C. Li and L. V. Wang, "Photoacoustic tomography and sensing in biomedicine," *Phys. Med. Biol.* **54**(19), R59–R97 (2009).
- L. Zeng et al., "High antinoise photoacoustic tomography based on a modified filtered backprojection algorithm with combination wavelet," *Med. Phys.* **34**(2), 556–563 (2007).
- S. H. Holan and J. A. Viator, "Automated wavelet denoising of photoacoustic signals for circulating melanoma cell detection and burn image reconstruction," *Phys. Med. Biol.* **53**(12), N227–N236 (2008).
- N. A. Rejesh, H. Pullagurra, and M. Pramanik, "Deconvolution-based deblurring of reconstructed images in photoacoustic/thermoacoustic tomography," *J. Opt. Soc. Am. A* **30**(10), 1994–2001 (2013).
- S. Gutta et al., "Deep neural network-based bandwidth enhancement of photoacoustic data," *J. Biomed. Opt.* **22**(11), 116001 (2017).
- R. A. Kruger et al., "Dedicated 3D photoacoustic breast imaging," *Med. Phys.* **40**(11), 113301 (2013).
- R. M. Lewitt and S. Matej, "Overview of methods for image reconstruction from projections in emission computed tomography," *Proc. IEEE* **91**(10), 1588–1611 (2003).
- J. Chen et al., "Blind-deconvolution optical-resolution photoacoustic microscopy in vivo," *Opt. Express* **21**(6), 7316–7327 (2013).
- J. Prakash et al., "Model-resolution-based basis pursuit deconvolution improves diffuse optical tomographic imaging," *IEEE Trans. Med. Imaging* **33**(4), 891–901 (2014).

40. H. B. Mitchell, *Image Fusion: Theories, Techniques and Applications*, Springer Science and Business Media, Heidelberg (2010).
41. S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.* **22**(7), 2864–2875 (2013).
42. K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(6), 1397–1409 (2013).
43. G. L. Zeng, *Medical Image Reconstruction: A Conceptual Tutorial*, Springer, New York (2010).
44. M. Tao, J. Yang, and B. He, "Alternating direction algorithms for total variation deconvolution in image reconstruction," TR0918, Department of Mathematics, Nanjing University (2009).
45. Y. Wang et al., "A new alternating minimization algorithm for total variation image reconstruction," *SIAM J. Imaging Sci.* **1**(3), 248–272 (2008).
46. J. Prakash and P. K. Yalavarthy, "A LSQR-type method provides a computationally efficient automated optimal choice of regularization parameter in diffuse optical tomography," *Med. Phys.* **40**(3), 033101 (2013).
47. B. E. Treeby and B. T. Cox, "k-wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**(2), 021314 (2010).
48. Y. Hristova, P. Kuchment, and L. Nguyen, "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media," *Inverse Probl.* **24**(5), 055006 (2008).
49. C. C. Paige and M. A. Saunders, "LSQR: an algorithm for sparse linear equations and sparse least squares," *ACM Trans. Math. Software* **8**(1), 43–71 (1982).
50. M. E. Kilmer and D. P. O'Leary, "Choosing regularization parameters in iterative methods for ill-posed problems," *SIAM J. Matrix Anal. Appl.* **22**(4), 1204–1221 (2001).
51. C. Brezinski, G. Rodriguez, and S. Seatzu, "Error estimates for the regularization of least squares problems," *Numer. Algorithms* **51**(1), 61–76 (2009).
52. L. Reichel, G. Rodriguez, and S. Seatzu, "Error estimates for large-scale ill-posed problems," *Numer. Algorithms* **51**(3), 341–361 (2009).
53. M. Bhatt, A. Acharya, and P. K. Yalavarthy, "Computationally efficient error estimate for evaluation of regularization in photoacoustic tomography," *J. Biomed. Opt.* **21**(10), 106002 (2016).
54. M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.* **19**(9), 2345–2356 (2010).
55. A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imaging Vision* **20**(1), 89–97 (2004).
56. J. Eckstein and D. P. Bertsekas, "On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.* **55**(1), 293–318 (1992).
57. J. C. Goswami and A. K. Chan, *Fundamentals of Wavelets: Theory, Algorithms, and Applications*, John Wiley and Sons, Hoboken, New Jersey (2011).
58. K. R. Castleman, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, New Jersey (1979).
59. Y. Qiang, "Image denoising based on Haar wavelet transform," in *Proc. of Int. Conf. on Electronics and Optoelectronics*, Vol. 3, pp. 129–132 (2011).
60. A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, pp. 60–65 (2005).
61. R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Imaging Processing Using MATLAB*, Prentice Hall, Upper Saddle River, New Jersey (2004).
62. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990).
63. Z. Farbman et al., "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graphics* **27**(3), 67 (2008).
64. A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(2), 228–242 (2008).
65. K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2341–2353 (2011).
66. A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," *ACM Trans. Graphics* **23**(3), 689–694 (2004).
67. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth Int. Conf. on Computer Vision*, pp. 839–846, IEEE (1998).
68. V. Aurich and J. Weule, "Non-linear Gaussian filters performing edge preserving diffusion," in *Mustererkennung*, G. Sagerer, S. Posch, and F. Kummert, Eds., pp. 538–545, Springer, Berlin, Germany (1995).
69. S. M. Smith and J. M. Brady, "Susana new approach to low level image processing," *Int. J. Comput. Vision* **23**(1), 45–78 (1997).
70. C. Xiao and J. Gan, "Fast image dehazing using guided joint bilateral filter," *Visual Comput.* **28**(6–8), 713–721 (2012).
71. J.-H. Kim et al., "Optimized contrast enhancement for real-time image and video dehazing," *J. Visual Commun. Image Represent.* **24**(3), 410–425 (2013).
72. H.-J. Seo and P. Milanfar, "Robust flash denoising/deblurring by iterative guided filtering," *EURASIP J. Adv. Signal Process.* **2012**, 3 (2012).
73. B. Ham, M. Cho, and J. Ponce, "Robust guided image filtering using nonconvex potentials," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(1), 192–207 (2018).
74. G. L. Zeng, "A fast method to emulate an iterative POCS image reconstruction algorithm," *Med. Phys.* **44**(10), e353–e359 (2017).
75. D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Vol. 4, Cambridge University Press, Cambridge (2006).
76. D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* **81**(3), 425–455 (1994).
77. N. Gandhi et al., "Photoacoustic-based approach to surgical guidance performed with and without a da Vinci robot," *J. Biomed. Opt.* **22**(12), 121606 (2017).
78. P. P. Pai, A. De, and S. Banerjee, "Accuracy enhancement for noninvasive glucose estimation using dual-wavelength photoacoustic measurements and kernel-based calibration," *IEEE Trans. Instrum. Meas.* **67**(1), 126–136 (2018).
79. B. T. Cox and B. E. Treeby, "Artifact trapping during time reversal photoacoustic imaging for acoustically heterogeneous media," *IEEE Trans. Med. Imaging* **29**(2), 387–396 (2010).
80. J. Prakash et al., "Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction," *IEEE J. Sel. Top. Quantum Electron.* **20**(2), 74–82 (2014).
81. X. Song et al., "Automated region detection based on the contrast-to-noise ratio in near-infrared tomography," *Appl. Opt.* **43**(5), 1053–1062 (2004).
82. L. Li et al., "Multiview Hilbert transformation in full-ring transducer array-based photoacoustic computed tomography," *J. Biomed. Opt.* **22**(7), 076017 (2017).
83. Y. Lou et al., "Generation of anatomically realistic numerical phantoms for photoacoustic breast imaging," *Proc. SPIE* **9708**, 97084O (2016).
84. Y. Lou et al., "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *J. Biomed. Opt.* **22**(4), 041015 (2017).
85. S. K. Kalva and M. Pramanik, "Experimental validation of tangential resolution improvement in photoacoustic tomography using modified delay-and-sum reconstruction algorithm," *J. Biomed. Opt.* **21**(8), 086011 (2016).
86. The A Insitute, "American national standard for safe use of lasers (ansi z136. 1-2000)," (2000).
87. P. K. Upputuri and M. Pramanik, "Pulsed laser diode based photoacoustic imaging of biological tissues," *Biomed. Phys. Eng. Express* **1**(4), 045010 (2015).
88. Y. Matsumoto et al., "Label-free photoacoustic imaging of human palmar vessels: a structural morphological analysis," *Sci. Rep.* **8**(1), 786 (2018).
89. M. Xu et al., "Photoacoustic characteristics of lipid-rich plaques under ultra-low temperature and formaldehyde treatment," *Chin. Opt. Lett.* **16**(3), 031702 (2018).
90. S. K. Kalva and M. Pramanik, "Modified delay-and-sum reconstruction algorithm to improve tangential resolution in photoacoustic tomography," *Proc. SPIE* **10064**, 100643W (2017).
91. A. Petschke and P. J. La Rivière, "Comparison of photoacoustic image reconstruction algorithms using the channelized hotelling observer," *J. Biomed. Opt.* **18**(2), 026009 (2013).

92. K. Wang et al., "Discrete imaging models for three-dimensional optoacoustic tomography using radially symmetric expansion functions," *IEEE Trans. Med. Imaging* **33**, 1180–1193 (2014).
93. G. Li et al., "Multiview Hilbert transformation for full-view photoacoustic computed tomography using a linear array," *J. Biomed. Opt.* **20**(6), 066010 (2015).
94. Q. Sheng et al., "A constrained variable projection reconstruction method for photoacoustic computed tomography without accurate knowledge of transducer responses," *IEEE Trans. Med. Imaging* **34**(12), 2443–2458 (2015).
95. H. S. Kam, M. Hanmandlu, and W. H. Tan, "An improved image enhancement combining smoothing and sharpening," in *Conf. on Convergent Technologies for Asia-Pacific Region*, Vol. 1, pp. 36–40 (2003).
96. I. V. Safonov et al., "Adaptive sharpening of photos," *Proc. SPIE* **6807**, 68070U (2008).
97. N. Awasthi et al., "Image-guided filtering for improving photoacoustic tomographic image reconstruction," <https://sites.google.com/site/sercmig/home/patguided> (10 March 2018).

Navchetan Awasthi received MTech degree in computational science from Indian Institute of Science (IISc), Bangalore in 2016. He also received BTech degree in electronics and communication engineering from National Institute of Technology (NIT), Jalandhar in 2011. He is currently a PhD student in the Department of Computational and Data Sciences at Indian Institute of Science, Bangalore. His research interests include inverse problems in biomedical optics and medical image reconstruction.

Sandeep Kumar Kalva received his Masters degree in biomedical engineering from the Indian Institute of Technology (IIT), Hyderabad, India, in 2015. He is currently a PhD student in the School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore. His research area is on functional photoacoustic imaging for various clinical applications.

Manojit Pramanik received his PhD in biomedical engineering from Washington University in St. Louis, Missouri, USA. Currently, he is an assistant professor of the School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore. His research interests include the development of photoacoustic/thermoacoustic imaging systems, image reconstruction methods, clinical application areas such as breast cancer imaging, molecular imaging, contrast agent development, and Monte Carlo simulation of light propagation in biological tissue.

Phaneendra K. Yalavarthy received his MSc degree in engineering from the Indian Institute of Science, Bangalore and PhD degree in biomedical computation from Dartmouth College, Hanover, NH, USA, in 2007. He is an associate professor with the Department of Computational and Data Sciences at Indian Institute of Science, Bangalore. His research interests include medical image computing, medical image analysis, and biomedical optics. He is a senior member of IEEE and serves as an associate editor of IEEE Transactions on Medical Imaging.