

Multispectral Photoacoustic Imaging Artifact Removal and Denoising Using Time Series Model-Based Spectral Noise Estimation

Agne Kazakeviciute*, Chris J. H. Ho, and Malini Olivo

Abstract—The aim of this study is to solve a problem of denoising and artifact removal from *in-vivo* multispectral photoacoustic imaging when the level of noise is not known a priori. The study analyzes Wiener filtering in Fourier domain when a family of anisotropic shape filters is considered. The unknown noise and signal power spectral densities are estimated using spectral information of images and the autoregressive of the power 1 (AR(1)) model. Edge preservation is achieved by detecting image edges in the original and the denoised image and superimposing a weighted contribution of the two edge images to the resulting denoised image. The method is tested on multispectral photoacoustic images from simulations, a tissue-mimicking phantom, as well as *in-vivo* imaging of the mouse, with its performance compared against that of the standard Wiener filtering in Fourier domain. The results reveal better denoising and fine details preservation capabilities of the proposed method when compared to that of the standard Wiener filtering in Fourier domain, suggesting that this could be a useful denoising technique for other multispectral photoacoustic studies.

Index Terms—Photoacoustic imaging, artifact removal, denoising, Wiener filtering, Fourier domain, PSD estimation, AR(1) model.

I. INTRODUCTION

MULTISPECTRAL photoacoustic imaging (PAI) is well-known to be prone to various noise and artifacts. A type of noise in multispectral PAI may be similar to that present in hyperspectral imaging (HSI) such as pattern noise due to calibration error [1], while another type of noise, such as additive electronic noise, comes from system thermal noise or electromagnetic interference [2]. Multispectral PAI artifacts can be due to motion or image reconstruction. Motion-based artifacts arise due to movement of an object such as breathing or heart beat [3]. Reconstruction-based artifacts can arise due to limited angle issues in backprojection reconstruction algorithms. For instance, spatial undersampling can lead to

streak artifacts during image reconstruction due to limited number of elements in the transducer array. In these contexts, denoising multispectral PAI and removing artifacts is a crucial procedure for further image processing and analysis such as image segmentation, spectral unmixing or co-registration.

Great effort has been made to reduce reconstruction-based artifacts in multispectral PAI by making modifications to the reconstruction algorithms. For example, Paltauf et.al. [5] proposed special weight functions for direct, non-iterative back projection that reduces artifacts in multispectral PAI. Ma et.al. [6] introduced the mean-back-projection-iteration algorithm which minimizes the error between the observed signal and the signal calculated from the reconstructed image. Dean-Ben et.al. [4] proposed another weighted back-projection algorithm based on the estimation of acoustic scatterers distribution within the imaged object. In the work of Yao et. al. [7] total variation maximization (TVM) regularization was used in image reconstruction step. In the work of Wang et.al. [8] a constraint on image consistency was added to TVM approach, while in the work of Prakash et.al. [9] basis pursuit approach with sparsity-inducing L_1 regularization was used instead of TVM.

However, only reconstruction-based artifacts are targeted and in most cases artifacts are not removed completely but only reduced to a certain extent. Moreover, noise is not always guaranteed to be suppressed by the reconstruction-based artifact removal methods. It is known that, if the imaging system is tomographic, reconstruction algorithms and, especially, improved variations of them, are able to suppress noise to some level through the superposition of projection signals in the image domain (see [2]). However, in non-tomographic systems, the suppression does not give good results. Also, total variation regularization based on TVM or sparsity regularization based on L_1 works well only when the true image is, in fact, sparse or has low variation. As a result, this works well for many phantom data as usually phantoms are specially constructed to be sparse and containing only a few absorbers (such as in the works of [9] and [8], where sparse phantoms were selected to test the algorithm). However, it is not so clear why this should always apply to *in-vivo* multispectral photoacoustic imaging, especially to those cases, where complex tissue structures are imaged with high variation. Therefore, if the multispectral photoacoustic image does not comply with the required assumptions, it is highly likely that some amount of information will be lost after using TVM or L_1 related methods.

As for denoising, it is highly likely that, as in the case for

Manuscript received July 7, 2015; revised March 31, 2016; accepted April 3, 2016; date of current version April 4, 2016. Asterisk indicates corresponding author.

A. Kazakeviciute is with the Department of Statistical Science, University College London, London, WC1E 6BT, United Kingdom, and with Singapore Bioimaging Consortium, Agency for Science, Technology and Research, Singapore (e-mail: a.kazakeviciute.12@ucl.ac.uk).

C. J. H. Ho is with Singapore Bioimaging Consortium, Agency for Science, Technology and Research, Singapore.

M. Olivo is with Singapore Bioimaging Consortium, Agency for Science, Technology and Research, Singapore, and with the School of Physics, National University of Ireland, Galway, University Road, Galway, Ireland.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.org>.

Digital Object Identifier 00.0000/000.0000.00000000

HSI, a small image ROI obtained from multispectral PAI can be well approximated by a few pixels. Therefore, a wide range of HSI denoising methods could be applied to fit multispectral PAI domain, especially methods based on low rank assumption (see, e.g. [10], [11], [12]). However, many of them depend on local or global dictionary learning and cannot guarantee artifact removal that are hardly distinguished from image edges. Therefore, only a specific class of HSI denoising methods that target some common characteristics of both noise and artifacts could be of any improvement of overall image quality from multispectral PAI.

In this work, we are interested in removing white noise and artifacts from *in-vivo* multispectral PAI, where the level of noise and artifacts is not known a priori, and we will refer to them simply as noise, though, keeping in mind the very different nature they might have. We propose an analytical method based on Wiener filtering in the Fourier domain, where only Gaussian-like filters having an anisotropic elliptical shape are considered prior to minimizing MSE, and regularization is controlled by the two parameters of the filter.

In Wiener filtering, a filter is constructed to minimize the mean squared error (MSE) between the observed and denoised images. However, a crucial assumption of Wiener filtering is that the power spectral densities (PSDs) of noise and signal are known a priori or can be well estimated. While many PSD estimation methods exist in literature, they are usually applicable to time series data and not images. Therefore, image modality-specific assumptions have to be made so that PSDs of noise and signal could be estimated using tools for time series data. In the work of Kamel et.al. [13] autoregressive modeling was used to estimate noise in a 2D image based on spatial pixel dependence, while in the work of Seghouane et.al. [14] autoregressive modeling was used to estimate noise based on temporal pixel dependence in functional magnetic resonance imaging (fMRI). In this work, we use the autoregressive power 1 (AR(1)) model to estimate noise based on spectral pixel dependence in multispectral PAI. We demonstrated how estimates of noise and signal PSDs can be analytically derived, tested the method on phantom data as well as *in-vivo* mouse images and found that it is superior than the standard Wiener filtering in Fourier domain.

This work is organized as follows. In Section II we present the model and assumptions used. In Section III we give a brief introduction to the problem of Wiener deconvolution, solution of which in a special case gives us the well-known standard Wiener filtering in Fourier domain. In Section IV we describe the proposed method. Finally, in Section V we present the results of the proposed method on purely simulated data, phantom and *in-vivo* data, which are then compared to those of the standard Wiener filtering in Fourier domain. We end the work with a brief discussion in Section VI. The proof of analytical derivation of the noise power spectral density (PSD) is provided in Appendix.

II. MODEL AND ASSUMPTIONS

Suppose we are given a mean-corrected multispectral photoacoustic image $Y_{M \times N \times B}$, where $M \times N$ is the number of spatial coordinates and B is the number of spectral bands. We

assume that at each spatial location (i, j) , the following model holds:

$$y_{ij} = x_{ij} + v_{ij}, \quad (1)$$

where y is the observed pixel, x is the true signal and $v \sim N(0, \sigma^2)$ is noise. Since the original multispectral photoacoustic image has intensity values in interval $[0; 1]$, we use mean corrected multispectral PAI so that noise would be zero-centered. We assume that noise is independent of signal. We also assume that at each image background location (i, j, k) , where k denotes spectral band, the following relation holds:

$$y_{ijk} = b_0 + b y_{ij(k-1)} + v_{ijk},$$

where b_0, b are AR(1) model parameters, and $y_{i,j,k}$ is the (i, j, k) th pixel of the mean-corrected image Y . Finally, we assume that noise occurs with high frequencies and signal occurs with low frequencies, meaning that the signal has smooth intensity changes and the most dramatic intensity changes are caused by noise.

III. STANDARD WIENER FILTERING

In Wiener deconvolution, the following model is considered:

$$Y = h * X + V, \quad (2)$$

where X and V are not observable random functions of time, i.e. $X = X(t), V = V(t)$, X is independent of V , the expected value of V is zero, h is the known so-called blurring filter and $*$ denotes convolution. In Fourier domain, model (2) becomes

$$F(Y) = F(h)F(X) + F(V),$$

where F is a mapping from functional domain to the frequency (or, in other words, Fourier) domain. The goal of Wiener deconvolution is to guess what $X(t)$ with all t is, given observations $Y(t)$ with all t . That guess is the estimate $\hat{X}(t) = (g * Y)(t)$ with some function g . Therefore, the aim is to estimate $F(g)$ which is the function of coordinate λ in the frequency domain. The criteria of Wiener is that with every $\lambda \in \mathbb{R}$, $F(g)(\lambda)$ needs to minimize the following functional:

$$\begin{aligned} \epsilon(F(g)) &= E|F(X)(\lambda) - F(\hat{X})(\lambda)|^2 \\ &= E|F(X)(\lambda) - F(g)(\lambda)F(Y)(\lambda)|^2. \end{aligned} \quad (3)$$

Note that $|\cdot|$ here denotes the absolute value of a complex number: if z^2 is a complex number then $|z|^2$ is the real nonnegative number.

It is a well known fact (see, e.g. [15]) that the $F(g)$ that minimizes the cost function (3) is of the form

$$F(g) = \frac{F(h)^* s}{|F(h)|^2 s + n}, \quad (4)$$

where $s = E|F(X)|^2$ is the so-called signal power spectral density (PSD) and $n = E|F(V)|^2$ is the so-called noise PSD, $z^* = x - iy$ with $x, y \in \mathbb{R}$ and $zz^* = (x + iy)(x - iy) = x^2 + y^2 = |z|^2$.

Let λ_{ij} denote the spatial position in the frequency domain. We can see that our model (1) can be regarded as Wiener deconvolution with unit blurring filter h . Then at every position λ_{ij} we can use the same Fourier domain estimate (4)

$$F(g)(\lambda_{ij}) = \frac{s(\lambda_{ij})}{s(\lambda_{ij}) + n(\lambda_{ij})} \quad (5)$$

which minimizes the following cost function:

$$\epsilon(F(g)) = \sum_{i=1}^M \sum_{j=1}^N E|F(X)(\lambda_{ij}) - F(g)(\lambda_{ij})F(Y)(\lambda_{ij})|^2. \quad (6)$$

We will refer to filter (5) as to the standard Wiener filtering in Fourier domain.

IV. PROPOSED METHOD

A. Filter Form

We note that using standard Wiener filtering in Fourier domain (5) does not help us to remove high frequency components which we assume to be noise. To account for this, we construct the following filter:

$$F(g)(\lambda_{ij}) = e^{-((i-c_x)^2/\gamma_1 + (j-c_y)^2/\gamma_2)}. \quad (7)$$

Note that (if $\gamma_1 \neq \gamma_2$) filter (7) is an anisotropic elliptical filter which is centered at position $\lambda_{c_x c_y} = (c_x, c_y)$ and whose parameters γ_1, γ_2 control the two radii of an ellipse. Note that in practice we always set $\lambda_{c_x c_y}$ to be the center of the frequency domain of an image and therefore c_x, c_y are known and the filter $F(g)$ is fully determined by the two radii γ_1, γ_2 .

There are four extreme cases: (a) both γ_1, γ_2 tend to ∞ ; (b) both γ_1, γ_2 tend to 0; (c) γ_1 tends to 0, γ_2 tends to ∞ and (d) γ_1 tends to ∞ and γ_2 tends to 0. In the case (a) the filter $F(g)$ at every position λ_{ij} tends to 1 and the cost function (6) reduces to $\sum_{i=1}^M \sum_{j=1}^N n_{ij}$. In the case (b) $F(g)$ tends to 0 for all λ_{ij} , except the $\lambda_{c_x c_y}$, where it still equals 1. There, the cost function (6) reduces to $\sum_{i=1}^M \sum_{j=1}^N E|F(X)(\lambda_{ij}) - F(Y)(\lambda_{c_x c_y})|^2$. In the case (c) the cost function (6) reduces to $\sum_{i=1}^M \sum_{j=1}^N E|F(X)(\lambda_{ij}) - e^{-1}F(Y)(\lambda_{i c_y})|^2$. On the other hand, in the case (d) the cost function (6) reduces to $\sum_{i=1}^M \sum_{j=1}^N E|F(X)(\lambda_{ij}) - e^{-1}F(Y)(\lambda_{c_x j})|^2$. Minimizing (6) with filter of the form (7) does not give a trivial solution because the optimal point might be achieved somewhere in the middle of these four extreme cases. Therefore, the proposed method gives us a plausible solution.

B. Estimating γ_1, γ_2

Since filter $F(g)$ is fully determined by the parameters γ_1, γ_2 , the cost function (6) is actually $\epsilon(\gamma_1, \gamma_2)$, that is, a function of γ_1, γ_2 . Obviously, we search for such γ_1, γ_2 so that the cost function (6) is minimized. We can expand this as

$$\begin{aligned} \epsilon(\gamma_1, \gamma_2) &= \sum_{i=1}^M \sum_{j=1}^N [E|(1 - F(g)(\lambda_{ij}))F(X)(\lambda_{ij}) \\ &\quad - F(g)(\lambda_{ij})F(V)(\lambda_{ij})|^2] \\ &= \sum_{i=1}^M \sum_{j=1}^N [E[(1 - F(g)(\lambda_{ij}))F(X)(\lambda_{ij}) \\ &\quad - F(g)(\lambda_{ij})F(V)(\lambda_{ij})] \times \\ &\quad \times [(1 - F(g)(\lambda_{ij}))F(X)(\lambda_{ij}) - F(g)(\lambda_{ij})F(V)(\lambda_{ij})]^*] \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^M \sum_{j=1}^N [(1 - F(g)(\lambda_{ij}))^2 E|F(X)(\lambda_{ij})|^2 \\ &\quad + F(g)(\lambda_{ij})^2 E|F(V)(\lambda_{ij})|^2] \\ &= \sum_{i=1}^M \sum_{j=1}^N \left[\left(1 - e^{-((i-c_x)^2/\gamma_1 + (j-c_y)^2/\gamma_2)} \right)^2 s_{ij} + \right. \\ &\quad \left. + e^{-2((i-c_x)^2/\gamma_1 + (j-c_y)^2/\gamma_2)} n_{ij} \right], \end{aligned} \quad (8)$$

where $s_{ij} = E|F(X)(\lambda_{ij})|^2, n_{ij} = E|F(V)(\lambda_{ij})|^2$ are the so-called power spectral densities (PSDs) of signal and noise, respectively. Denote $a_i = (i - c_x)^2/\gamma_1$ and $a_j = (j - c_y)^2/\gamma_2$. For $a_i = a_j = 0$, i.e. when $i = c_x, j = c_y$, any γ_1, γ_2 are optimal. For $a_i \neq 0, a_j \neq 0$, we cannot solve (8) analytically and we use computer to minimize this non-convex function.

C. Estimating PSDs

The remaining major challenge is the estimation of PSDs s and n . Studies that considered deconvolution problem roughly fall into two categories: (a) the ones that used clean images and added simulated noise in which case s and n were known exactly (see, e.g. [16], [17]); (b) the ones that estimated them using some method (such as Welch's method, periodogram, averaging, taking median from a set of samples (see, e.g. [18], [19]), making assumption that signal or noise are distributed equally across samples. There are also studies that considered autoregressive modeling to estimate noise from an image. For these studies, it is usual to treat as samples the spatial pixels in a 2D image (see, e.g. [13]), or temporal pixels in different 2D bands of a 3D temporal image, such as functional magnetic resonance imaging (fMRI) (see, e.g. [14]). However, at first look, none of the above strategies fits the denoising problem in multispectral PAI. First of all, the actual s and n are unknown. Secondly, we have only one sample (that is, image) but it is not a temporal image, that is, different bands do not correspond to different time. Moreover, using autoregressive modeling on spatial information of a 2D band would not work in our case because pixels are not functions of spatially neighboring pixels. Therefore, the task to give good estimates of s and n is not a straightforward one.

We note, that in order to estimate s and n , it is enough to estimate one of them and the sum $s + n$. Then the second one can be found by subtracting the first from the sum.

1) Estimating $s+n$

The first task of estimating $s + n$ is much easier. We propose to exploit spectral information and treat each spectral band of the multispectral photoacoustic image as a sample. Then, for a fixed position λ_{ij} , we have the following set of samples: $\{F(y)_1(\lambda_{ij}), \dots, F(y)_B(\lambda_{ij})\}$, where λ_{ij} denotes the position (i, j) in the frequency domain, $F(y)_k(\lambda_{ij}), k = 1, \dots, B$ is the Fourier transform of a pixel y at position (i, j, k) and B is the number of spectral bands. We note that

$$\begin{aligned} s + n &= E|F(X)|^2 + E|F(V)|^2 \\ &= E|F(X) + F(V)|^2 = E|F(Y)|^2, \end{aligned}$$

where the second equation holds because signal is assumed to be independent of noise and the third equation holds because

of our model. In practice we need to estimate $s + n$ for each position λ_{ij} and the true expectation is replaced by sample average. This leads to the following estimation:

$$\widehat{s + n} = \frac{1}{B} \sum_{k=1}^B |F(y)_k(\lambda_{ij})|^2. \quad (9)$$

2) Estimating n

Now what is left to do, is to estimate s or n . In this work we will estimate n . The standard methods to estimate PSD of noise are usually applied to time series data, where we have, say, n observations of time y_{t_1}, \dots, y_{t_n} , and each observation is directly related to the previous one. The standard model for time series is the autoregressive power p (AR(p)) model:

$$y_{t_i} = b_0 + \sum_{i=1}^p b_i y_{t_i-p} + \epsilon_{t_i}, \quad (10)$$

where $t_i, i = 1, \dots, n$ denote time points, $\epsilon_{t_i} \sim N(0, \sigma^2)$ is the noise term at the time point t_i and b_0, \dots, b_p are the parameters of the AR(p) model. While model (10) is a standard model for time series, it is not straightforward how this could be applied to image domain. If we treated each pixel as an observation, this would mean that each pixel in an image is a function of spatially neighboring pixels. This was assumed, for example, in the work of Kamel et.al. [13] but this may not hold for multispectral PAI. To be able to apply AR model to our case here we make the assumption about pixel dependence not on spatially neighboring pixels but on spectrally neighboring ones instead:

$$y_{ijk} = b_0 + b y_{ij(k-1)} + v_{ijk}, \quad (11)$$

where y_{ijk} is the background pixel of a mean-corrected multispectral photoacoustic image at position (i, j, k) , v_{ijk} is the noise at position (i, j, k) and b_0, b are parameters of the model. Here $i = 1, \dots, M, j = 1, \dots, N, k = 1, \dots, B$. We note that (11) is autoregressive power 1 (AR(1)) model. If (11) holds in spatial domain, it holds also in frequency domain and we have

$$F(y)_k(\lambda_{ij}) = c_0 + cF(y)_{k-1}(\lambda_{ij}) + F(v)_k(\lambda_{ij}), \quad (12)$$

where λ_{ij} denotes the position (i, j) in frequency domain, $F(y)_k(\lambda_{ij}), k = 1, \dots, B$ denotes the Fourier transform of y_{ijk} and $c_0 = F(b_0), c = F(b)$.

PROPOSITION 4.1: Let $F(y)_0(\lambda_{ij}) = 1$. Then the ML estimate of n is

$$\hat{n} = \frac{1}{2MNB} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - \hat{c}_0|^2, \quad (13)$$

where

$$\hat{c}_0 = \frac{1}{MNB} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B [F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij})].$$

The proof is provided in Appendix.

The estimate $\hat{s}(\lambda_{ij})$ is then calculated by setting

$$\hat{s}(\lambda_{ij}) = \max(\widehat{(s+n)}(\lambda_{ij}) - \hat{n}(\lambda_{ij}), 0), \quad (14)$$

where $\widehat{(s+n)}(\lambda_{ij})$ is as defined by (9).

D. Preserving Edges

Even though edge preservation is not the main focus of this work, it is important that edges are preserved in the denoised image. Here, we apply a simple approach to edge preservation. After image is denoised, the MATLAB built-in function `edge('canny')` is used to detect image edges. Since the denoised image is noise-free, it is able to detect true image edges and automatic built-in parameter value search give promising results. The 3D edge map is then created, from which we can obtain two copies of edge pixels: one from the original image and one from the denoised image. We then set edge pixels of the final image as the following affine combination of the two copies:

$$X^{edge} = w_{old} Y_{old}^{edge} + w_{new} Y_{new}^{edge},$$

where X^{edge} is a 3D tensor containing edge pixel values in the final image, Y_{old}^{edge} is a 3D tensor containing edge pixel values in the original image and Y_{new}^{edge} is a 3D tensor containing edge pixel values in the denoised image. The weights w_{old}, w_{new} are probabilistic weights, i.e. $0 \leq w_{old}, w_{new} \leq 1$ and $w_{old} + w_{new} = 1$. The weight w_{new} is selected to minimize the following non-convex cost function:

$$\hat{w}_{new} = \arg \min_{w_{new}} \frac{\text{MRMSE}(X_{w_{new}})}{\text{SNR}_2(X_{w_{new}})},$$

where $X_{w_{new}}$ is the final 3D image, depending on the weight w_{new} and MRMSE, SNR_2 measures are as defined in Results Section V. The conceptual illustration of the proposed method is displayed in Figure 1.

V. RESULTS

A. Quality Measures

The method was compared with the standard Wiener filtering in Fourier domain described in Section IV, where the filter of the form (5) was applied, using the same estimates (13) and (14) as for the proposed method. Since both methods should enhance signal-to-noise ratio (SNR), one of the quality measures was also SNR, defined in two ways. The first way is the most popular while working with the $[0, 1]$ gray-scale images. For a given 2D spectral band $X_k, k = 1, \dots, B$, it is calculated as:

$$\text{SNR}_1(X_k) = 10 \log_{10} \frac{\bar{X}_k}{\sigma^2},$$

where $\bar{X}_k = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N X_{ijk}$ is the mean of the k th spectral band X_k and σ^2 is the variance of the noise. We estimated σ^2 as the variance of the image background and we used the MATLAB built-in function `imfreehand` to extract the background. The mean value over the spectral bands is also reported and named as MSNR_1 :

$$\text{MSNR}_1(X) = \frac{1}{B} \sum_{k=1}^B \text{SNR}_1(X_k),$$

where X is a 3D image. The second definition of signal-to-noise ratio is based on our proposal. For a given 3D image X

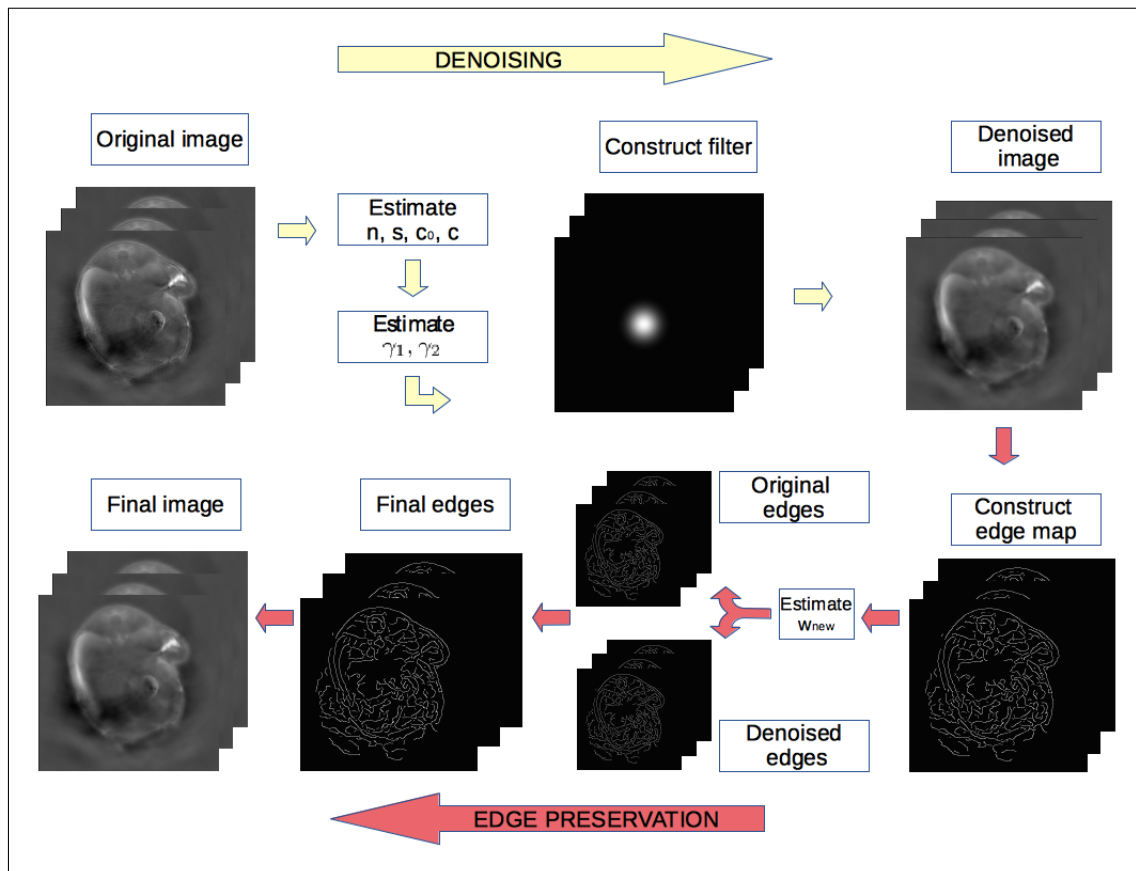


Fig. 1. Conceptual illustration of the proposed method. First, the AR(1) model parameter c is estimated from the original 3D image resulting in the estimation of depending parameter c_0 and estimates of signal and noise PSDs. Then, using estimated PSDs, the filter shape parameters γ_1, γ_2 are estimated and the filter with an anisotropic elliptical shape is constructed. Using this filter, the filtering of original image is done in Fourier domain and noise is removed. The 3D edge map is then constructed using the denoised image. Using this edge map, the optimal weight w_{new} is found. Using this edge map, two copies of edge pixels are extracted together with their weights (one copy from the denoised image and one copy from the original image) and a combination of affine transformations is used to calculate the final version of edge pixels. Finally, the image edges in the denoised image are replaced with the final version of edges and final image is obtained.

it is calculated as

$$\text{SNR}_2(X) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \frac{\hat{s}(\lambda_{ij})}{\hat{n}(\lambda_{ij})},$$

where \hat{s}, \hat{n} are the proposed estimates of s and n , respectively. The difference between MSNR_1 and SNR_2 is that in MSNR_1 signal estimation is based on spatial information, while in SNR_2 signal is estimated using spectral information.

To test the similarity of denoised images to the original image, two measures were taken. First measure is a relative mean-squared error defined as

$$\text{RMSE}(X_k) = \frac{\|Y_k - X_k\|_F}{\|Y_k\|_F},$$

where Y_k and X_k are the k th spectral band of the original and final denoised images, respectively. The mean RMSE value is also reported and named MRMSE:

$$\text{MRMSE}(X) = \frac{1}{B} \sum_{k=1}^B \text{RMSE}(X_k).$$

The second similarity measure that we used is the so-called structured similarity (SSIM), the long detailed description of

which can be found in [20]. The default values of parameters were used, except that for the image intensity range which was set to be $[0; 1]$. The mean SSIM value is also reported and named MSSIM:

$$\text{MSSIM}(X) = \frac{1}{B} \sum_{k=1}^B \text{SSIM}(X_k).$$

B. Multispectral PAI Parameters and Computational Cost

The MSOT system, inVision 128 (iThera Medical GmbH, Neuberberg, Germany) was used to acquire all phantom and *in-vivo* images [21]. The phantom is made of polyurethane, cylindrical in shape with a diameter of 2 cm, which is specially designed to mimic the shape, size and optical properties of the mouse. Typical wavelength-dependent absorption and reduced scattering coefficient values of the internal organs of a mouse at 800 nm are approximately 0.7 cm^{-1} and 8 cm^{-1} respectively, according to literature [22]. The laser excitation is provided by an optical parametric oscillator (OPO) laser and falls in the near-infrared range from 650 to 980 nm, while the ultrasound detection is achieved by a ring-shaped 128-element transducer array spanning a circular arc of 270° with a central frequency of 5 MHz. This gives a in-plane spatial resolution of around

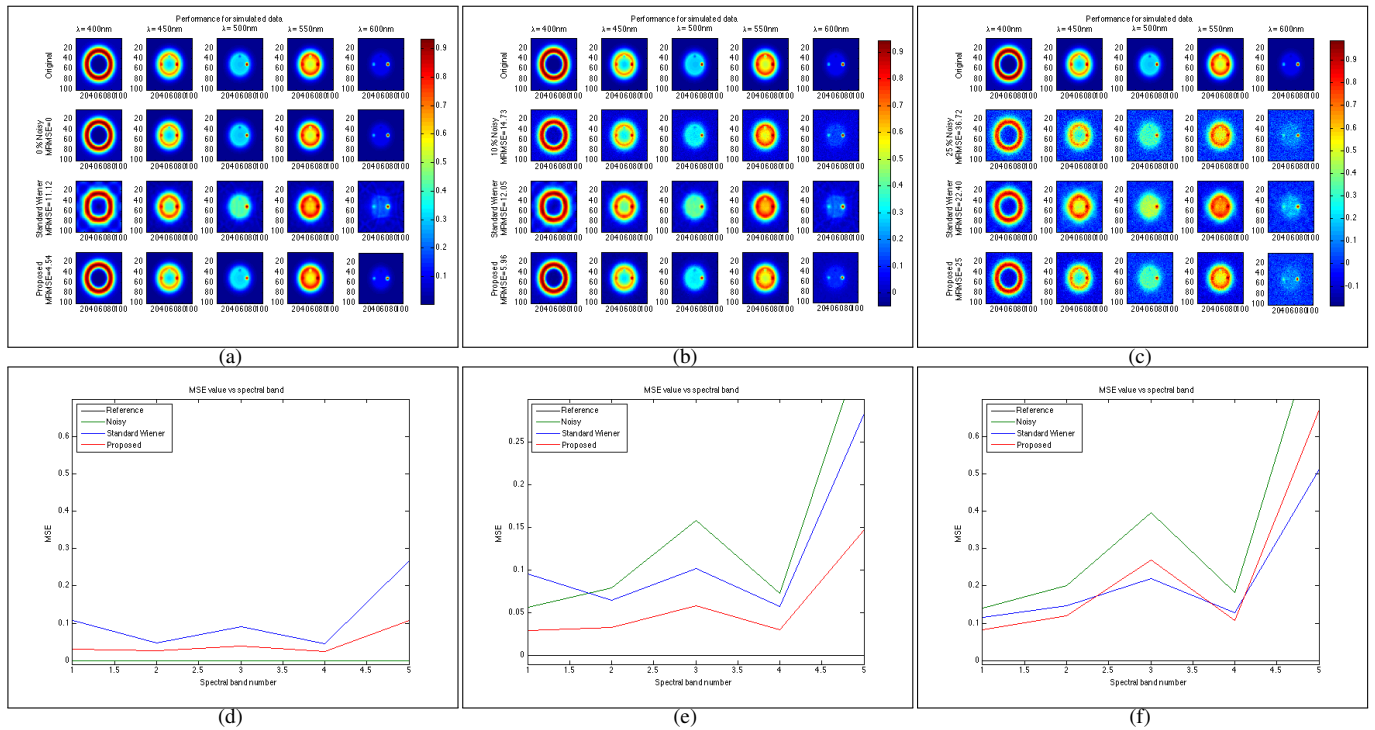


Fig. 2. Visual comparison of the proposed and standard Wiener methods in terms of the capability to remove white noise with different noise levels. (a)-(c) Visual comparison and MRMSE for 0%, 10% and 25% noise level, respectively. (d)-(f) MSE over spectral bands comparison for 0%, 10% and 25% noise level, respectively.

150 μm . Each of the transverse 2D image slices formed from the laser pulses are stacked to form a 3D tomographic image. Image reconstruction is then performed using a model-based approach for offline analysis.

For all three types of data: simulated data, phantom data and *in-vivo* data, the AR(1) model parameter c was found by maximizing the convex log-likelihood function $L(\sigma^2, c_0, c)$ using unconstrained minimization with a starting value equal to 0 and 10^{-15} precision. Since minimization with MATLAB built-in functions supports only real parameter values, the minimization procedure was divided into two steps, one for the real part of the parameter and the second for its complex part. Then $\hat{c}_0, \hat{n}, \hat{s}$ were calculated and reported for each image separately. The cost function (8) was minimized using constrained minimization with lower bound equal to $[0, 0]$ for $[\gamma_1, \gamma_2]$, the starting value equal to $[1, 1]$ and 10^{-15} precision. The weight w_{new} was calculated by using constrained minimization with lower bound equal to 0, upper bound equal to 1, the starting value equal to 0.5 and the 10^{-3} precision.

The method, when performed on a Mac computer with 1.8 GHz Intel Core i5 processor and 4GB memory, takes less than 2 minutes to solve the three minimization problems. Within these 2 minutes, around 1.6 s, on average, is for estimating the AR(1) model parameter c with 10^{-15} precision, around 2.6 s, on average, is for estimating filter parameters γ_1, γ_2 with 10^{-15} precision, and the rest of the time is for estimating w_{new} with 10^{-3} precision. Time measurements are based on an average image size of $500 \times 500 \times 9$ pixels.

C. Simulated Data

To investigate the smearing and white noise removal capabilities of the proposed method, we used purely simulated multispectral PA data which mimics the optical properties of biological tissue. For this, we used constructed absorption and scattering dictionaries based on *in-vivo* data [23] of the main absorbers: Hb, HbO₂, melanin, water and fat, as well as the main scattering media: brain, skin, breast, bone and fatty tissues (see Figure 3).

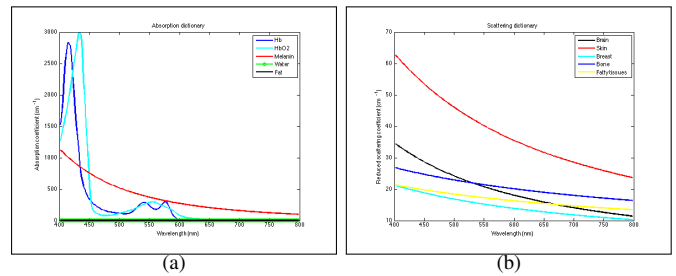


Fig. 3. Absorption and scattering dictionaries used for simulated data.

These absorption and scattering spectral signatures were then weighted by simulated weights which followed non-negativity and sum-to-one constraints to obtain the original object. The multispectral PA data was then simulated using a meshing method presented in [24] with 160801 nodes, 160000 elements, 1600 sources and 1600 detectors. The grid size of the solution was set to 100×100 . The resulting ‘observed’ image was of dimensions $100 \times 100 \times 5$, where the number 5 represents the number of spectral bands (400 nm, 450 nm, 500 nm, 550 nm

and 600 nm) chosen, based on spectral differentiation among the absorbers in Figure 3.

To investigate both smearing effects and white noise removal capabilities of the proposed method, we added different Gaussian noise levels to the original observed image. Usually, *in-vivo* multispectral photoacoustic images suffer from low-level noise [24] but it has been stated in [7] that noise level in a multispectral photoacoustic image could be as high as 20%. Therefore, our investigation concentrated on 0%, 10% and 25% Gaussian noise levels. The summary statistics for parameter estimates are provided in Table I together with the quality measures and visual comparison in Figure 2.

TABLE I
SUMMARY STATISTICS FOR SIMULATED DATA

Noise level (%)	0	10	25
\hat{c}	$0.2702 + 0.1554i$	$0.1373 - 0.2392i$	$-0.0840 - 0.2691i$
\hat{c}_0	$-0.0548 - 0.0315i$	$-0.0271 + 0.0473i$	$0.0171 + 0.0548i$
Mean \hat{s}	2.5209	3.0241	8.7007
Mean \hat{n}	1.3198	2.4213	7.8414
$\hat{\gamma}_1$	257.2687	346.4	$2.06 * 10^3$
$\hat{\gamma}_2$	253.2458	327.33	$2.06 * 10^3$
\hat{w}_{new}	0.0038	0.0996	0.1989

As we can see from Table I, in all cases the AR(1) model parameter $\hat{c} < 1$ which proves the point that the background of the simulated multispectral photoacoustic image, when mean-corrected, can be successfully considered as a stationary process. Also, the estimated signal-to-noise ratio decreases when the noise level increases, which indicates that the proposed SNR estimation is valid. We can see from Figure 2 (a) and (d) that in presence of no noise, both the proposed method and standard Wiener method suffers from smearing effects. However, we can see that the smearing effects of the proposed method are nearly 2.5 times smaller than those of the standard Wiener one (MRMSE=4.54 for the proposed filter versus MRMSE=11.12 for the standard Wiener filter). In fact, the stronger smearing effects of the standard Wiener method can be visually discerned (see, for example, Figure 2 (a), line 3, column 5). In the presence of 10% noise, both methods substantially improve the quality of the image with the proposed method performing about 2 times better than the standard Wiener filter. In the presence of 25% noise, the quality of the image improved by both filters is very similar, with the standard Wiener method performing slightly better on the last spectral band.

In addition, we also performed similar experiments to investigate the capabilities of both methods to remove white noise in the presence of much higher noise levels. Figure 4 summarizes such findings.

It turns out, that up to about 2% noise level, both filters suffer from smearing effects, with the proposed method always having smaller smearing effects than those of the standard Wiener method. More interestingly, the standard Wiener method starts performing better than the proposed method with the noise level greater than about 25%. These findings allow us to conclude that the choice of the filter depends on the prior knowledge of the noise level present in an image: the proposed method for noise level up to 25% and the standard Wiener method for the noise level greater than 25%. Therefore, since in *in-vivo* multispectral PAI the noise level is known not to exceed 25%

of the image intensity, the proposed denoising method is a better choice.

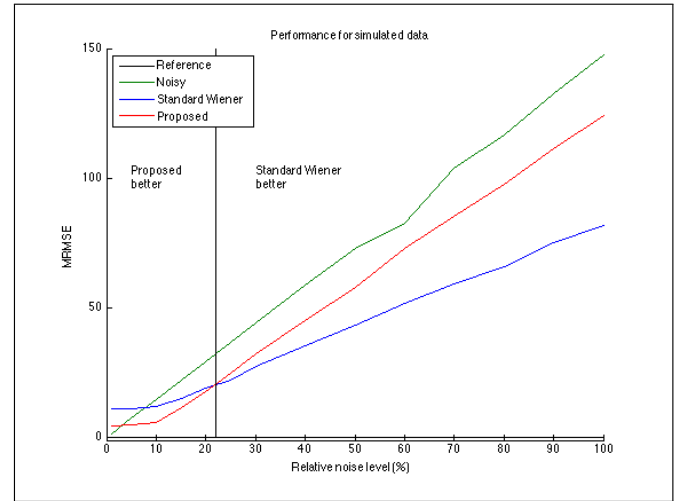


Fig. 4. Comparison of white noise removal capabilities of the proposed method and standard Wiener method.

D. Phantom Data

The proposed method (7) was tested on MSOT phantom image which consisted of 10 spectral bands ranging from 650 nm to 750 nm. Typical wavelength-dependent absorption and reduced scattering coefficient values of the internal organs of a mouse at 800 nm are approximately 0.7 cm⁻¹ and 8 cm⁻¹ respectively, according to literature. The summary statistics are provided in Table II together with the quality measures in Table III. The performance of quality measures SNR₁, MSE and SSIM across 10 spectral bands is shown in Figure 5. Figure 6 displays the visual performance of the proposed method for both background and object areas in the image. A more zoomed-in view of the original image, image denoised using the standard Wiener filtering in Fourier domain as well as the image denoised using the proposed method is provided in Figure 7.

As we can see from Table II, the AR(1) model parameter $\hat{c} \approx 0.54 + 0.54i$ which proves the point that the background of the MSOT phantom image, when mean-corrected, can be successfully considered as a stationary process which is an intrinsic assumption of Wiener filtering and which is needed to justify the use of AR(1) model. We can also see that $\hat{\gamma}_1$ is very similar to $\hat{\gamma}_2$ which indicates that nearly round shape of the filter was favored by the algorithm. This means that the filter is nearly proportional to that of the standard Gaussian filter. Nevertheless, we can see from Table III that the proposed method outperforms the standard Wiener filtering in Fourier domain in terms of all the quality measures.

Moreover, as we can see from Figure 5, the proposed method outperforms the standard Wiener filtering not only on average but continuously across different spectral bands. Figure 6 clearly shows that the background of the image is substantially smoothed (see sub-figure (a), surface plots) when using the proposed method, while the important edges in the image are preserved (see sub-figure (b), edge plots). A reader can take a more zoomed-in look at this in Figure 7.

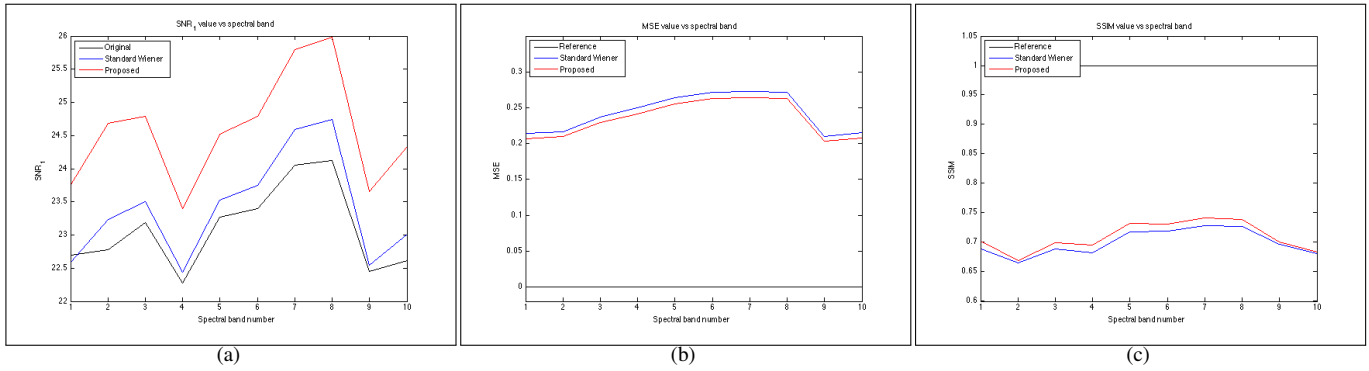


Fig. 5. SNR₁, MSE and SSIM values for phantom data across 10 spectral bands.

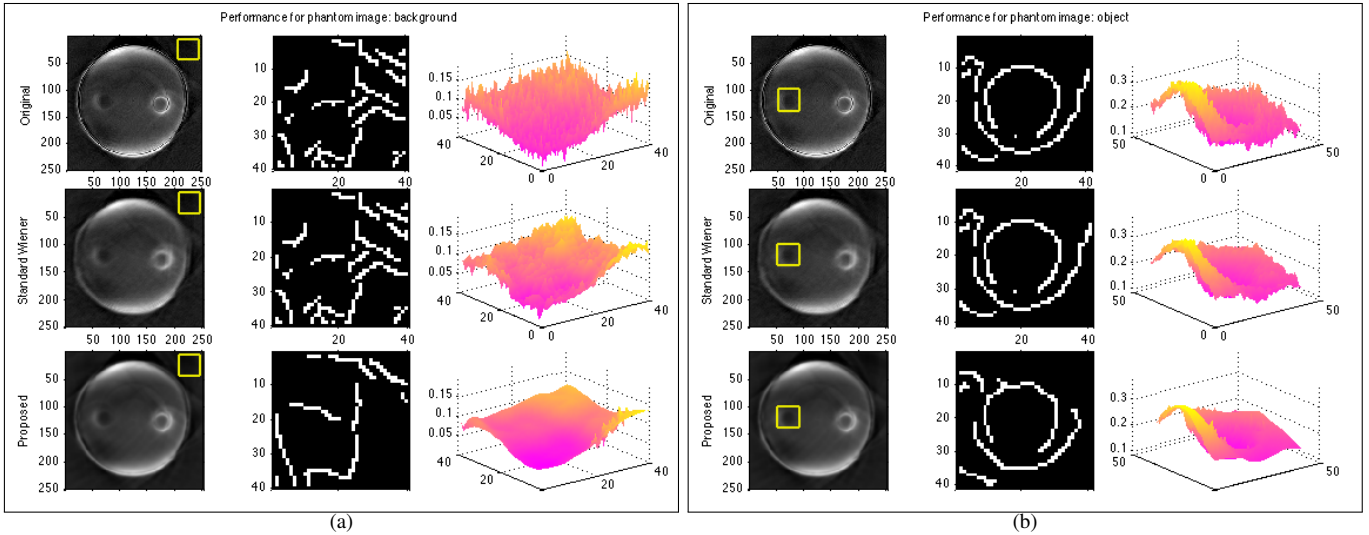


Fig. 6. Visual performance of the proposed method on phantom data. (a) Performance for background. (b) Performance for object. The edges in local patches are detected with MATLAB built-in function *edge* and default parameter values. Surface plots of background and object patches (the location of which is marked by a yellow square) are also included. Note that only the first spectral band is displayed, for simplicity.

TABLE II
SUMMARY STATISTICS FOR
PHANTOM DATA

\hat{c}	$0.5443 + 0.5373i$
\hat{c}_0	$-0.0540 - 0.0533i$
Mean \hat{s}	4.2107
Mean \hat{n}	2.0187
$\hat{\gamma}_1$	$1.3121 * 10^3$
$\hat{\gamma}_2$	$1.007 * 10^3$
\hat{w}_{new}	0.2541

TABLE III
EVALUATION MEASURES. THE
HIGHER MSNR₁, SNR₂,
MSSIM VALUES, THE BETTER.
THE LOWER MRMSE VALUE,
THE BETTER. THE BEST RESULTS
ARE MARKED IN BOLD

Measure	Image	Value
MSNR ₁	original	23.0838
MSNR ₁	standard Wiener	23.3918
MSNR ₁	proposed	24.5689
SNR ₂	original	2.0858
SNR ₂	standard Wiener	2.4413
SNR ₂	proposed	2.5406
MRMSE	standard Wiener	0.2251
MRMSE	proposed	0.2202
MSSIM	standard Wiener	0.6995
MSSIM	proposed	0.7094

Table V.

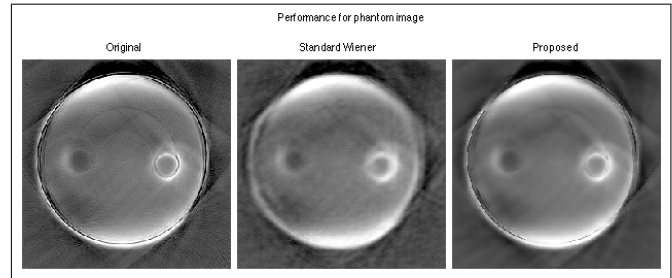


Fig. 7. Visual performance of the proposed method on phantom image.

E. In-Vivo Data

The proposed method (7) was also tested on *in-vivo* MSOT images of mouse body, liver and brain. *In-vivo* MSOT images consisted of 7 spectral bands, ranging from 740 nm to 900 nm, for mouse body, 9 spectral bands, ranging from 680 nm to 900 nm, for mouse liver, and 10 spectral bands, ranging from 680 nm to 900 nm, for mouse brain. The summary statistics are provided in Table IV together with the quality measures in

The performance of quality measures SNR₁, MSE and SSIM across spectral bands is displayed in Figure 9. Figure 10 is for visual performance of the proposed method for both background and object areas in the image. A more zoomed-in view of the original image, image denoised using standard Wiener filtering in Fourier domain as well as the image denoised using the proposed method is provided in Figure 8.

As we can see from Table IV, the AR(1) model parameter $\hat{c} \approx 0.39 + 0.4i$ for mouse body, $\hat{c} \approx 0.88 + 0.85i$ for mouse liver and $\hat{c} \approx 0.67 + 0.67i$ for mouse brain, which again proves

the point that the backgrounds of *in-vivo* multispectral photoacoustic images, when mean-corrected, can be successfully considered as stationary processes. We can see that again nearly round shape of the filter was favored by the algorithm. This indicates that the edges of *in-vivo* mouse images are distributed similarly to those of the phantom image which justifies the suitability of the phantom for this study. We can see from Table V that the proposed method slightly outperforms the standard Wiener filtering in Fourier domain in terms of all the quality measures and for all three types of images. Moreover, we can see from Figure 9 that better performance is visible more or less across different spectral bands. Similarly as for phantom data, Figure 10 clearly shows that the background of the image is substantially smoothed (see sub-figures (a), (c), (d), both surface and edge plots) when using the proposed method, while the important edges in the object are preserved (see sub-figures (b), (d), (f), especially edge plots). A reader can take a more zoomed-in look at this in Figure 8.

TABLE IV
SUMMARY STATISTICS FOR *In-Vivo* DATA

	Mouse body	Mouse liver	Mouse brain
\hat{c}	0.3910 + 0.4021 <i>i</i>	0.8763 + 0.8497 <i>i</i>	0.6715 + 0.6683 <i>i</i>
\hat{c}_0	-0.0557 - 0.0573 <i>i</i>	-0.1034 - 0.1003 <i>i</i>	-0.0686 - 0.0683 <i>i</i>
Mean \hat{s}	28.3599	373.0694	75.5633
Mean \hat{n}	12.7495	61.2683	23.7770
$\hat{\gamma}_1$	1.6378 * 10 ³	1.5227 * 10 ³	1.3059 * 10 ³
$\hat{\gamma}_2$	1.4643 * 10 ³	1.7377 * 10 ³	1.3107 * 10 ³
\hat{w}_{new}	0.0248	0.0318	0.0834

TABLE V
EVALUATION RESULTS. THE HIGHER MSNR₁, SNR₂, MSSIM VALUES, THE BETTER. THE LOWER MRMSE VALUE, THE BETTER. THE BEST RESULTS ARE MARKED IN BOLD

Measure	Image	Mouse body	Mouse liver	Mouse brain
MSNR ₁	original	25.1408	26.1307	27.7497
MSNR ₁	standard Wiener	25.1833	26.3477	28.1611
MSNR ₁	proposed	25.3576	26.5282	28.3412
SNR ₂	original	2.2244	6.0891	3.1780
SNR ₂	standard Wiener	2.3138	6.1591	3.3177
SNR ₂	proposed	2.3145	6.1623	3.3245
MRMSE	standard Wiener	0.0426	0.0349	0.0298
MRMSE	proposed	0.0364	0.0293	0.0273
MSSIM	standard Wiener	0.9331	0.9155	0.8879
MSSIM	proposed	0.9453	0.9276	0.8917

VI. DISCUSSION

In this work, the problem of artifact removal and denoising of multispectral photoacoustic images was considered. This study focused towards *in-vivo* images, where the level of noise is not known a priori. The regularized Wiener filtering in Fourier domain was applied, where only a family of filters having an anisotropic elliptical shape was considered and regularization was achieved by the two radii of an ellipse. Although in the presented experiments nearly round filter shapes were favored, the method remains flexible and can be applied to the cases, where edges are distributed mostly along one of the two axis in the frequency domain.

To estimate power spectral densities of both noise and signal, the autoregressive power 1 (AR(1)) model was adopted that exploits the spectral information of multispectral PAI. Minimizing the negative log-likelihood function of the unknown

parameters c_0, c of the AR(1) model for all experiments gave estimates of c less than 1. According to the general theory of autoregressive models, this indicates that the signal is stationary and justifies our choice of using the AR(1) model. An important point to note is that one needs to mean-correct a multispectral photoacoustic image before applying the AR(1) model. By doing this, noise can be guaranteed to be zero-centered and the mean trend can be successfully removed from an image.

The filter preserved image edges by creating the edge images of the original and the denoised image and superimposing a weighted contribution of the two edge images to the resulting final image. Probabilistic weights were used. The motivation behind using this method was that (a) it can be applied in a post-processing step, since it is impossible to track edge coordinates when the edge map undergoes Fourier transformation; (b) it is easy to implement, with low computational cost; (c) sufficient edge preservation results were obtained already with this method. However, we do acknowledge that more advanced edge preservation methods exist and edge enhancement or preservation is, in fact, a separate research area. Therefore, finding the best edge preserving method that would achieve the best denoising results when combined with our proposed method remains out of the scope for this work.

Another important point to note is that only high-frequency noise can be successfully removed by this method and low-frequency noise that is hardly differentiable from an object will remain. Therefore, for limited-angle artifacts, only the high-frequency components of them will be removed. An interesting follow-up study would be to treat the filter centers as additional parameters, or to increase the number of ellipses in the filter to more than one. In this way, multiple regions of noise and artifacts could be captured without relying on the assumption that they occur only in high frequency regions. However, we leave this for future investigations.

In addition, we work with reconstructed images in the frequency domain. In this way, we reduce the time spent in algorithmic computations, as compared to that in the signal domain. This has been demonstrated in the work of Yao *et al.* [7], where many CPUs were used for parallel computation of their proposed reconstruction algorithm. Moreover, working in the image domain is beneficial as we are able to exploit the relations between pixels in a spatial 2D space. Without doubt, the proposed algorithm can be applied to the signal domain as well or in other words, to data obtained before the image reconstruction step. To achieve this, a filter needs to be simply modified to fit 1D cases. We note, however, that working in a 1D space is an easier option because an elliptical shape is not possible for 1D space and only a ball-shaped filter could be used (a ‘ball’ in \mathbb{R} is simply an interval). This means that we do not exploit spatial relations when working in the signal domain.

In this work we demonstrated that the proposed method has better denoising and fine details preservation capabilities when compared to those of the standard Wiener filtering in Fourier domain which suggests that this could be a useful denoising technique for other *in-vivo* multispectral photoacoustic studies and that the method offers room for further improvement and extension towards other imaging modalities as well.

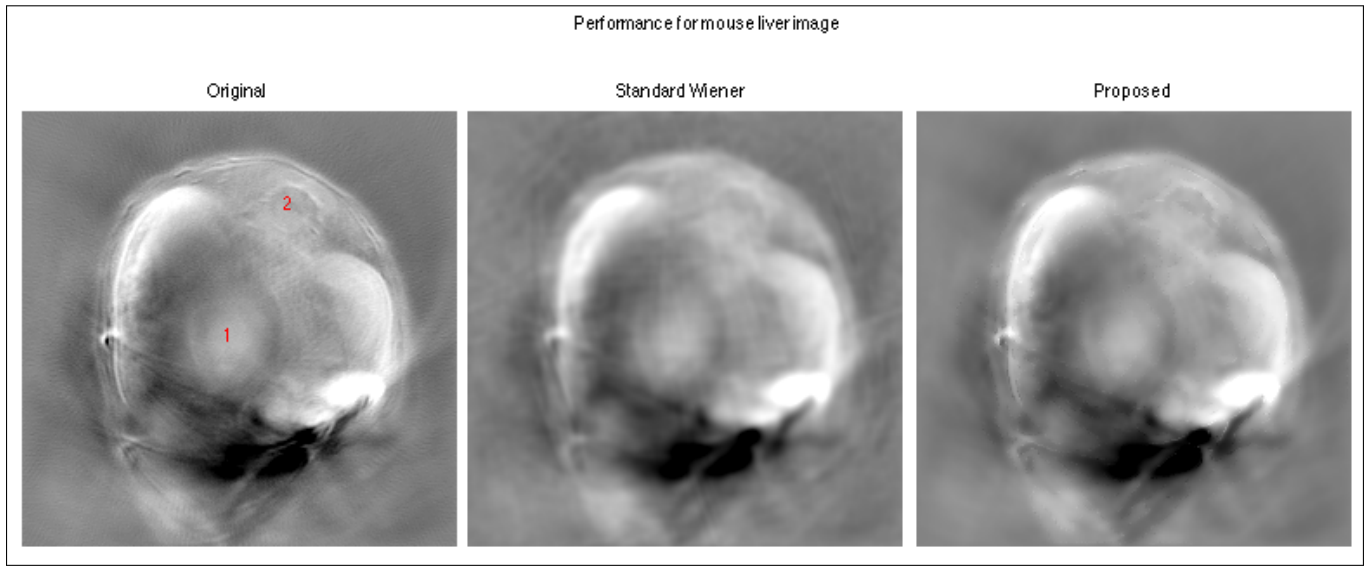


Fig. 8. Visual performance of the proposed method for mouse liver image. ‘1’ denotes stomach and ‘2’ denotes spinal cord.

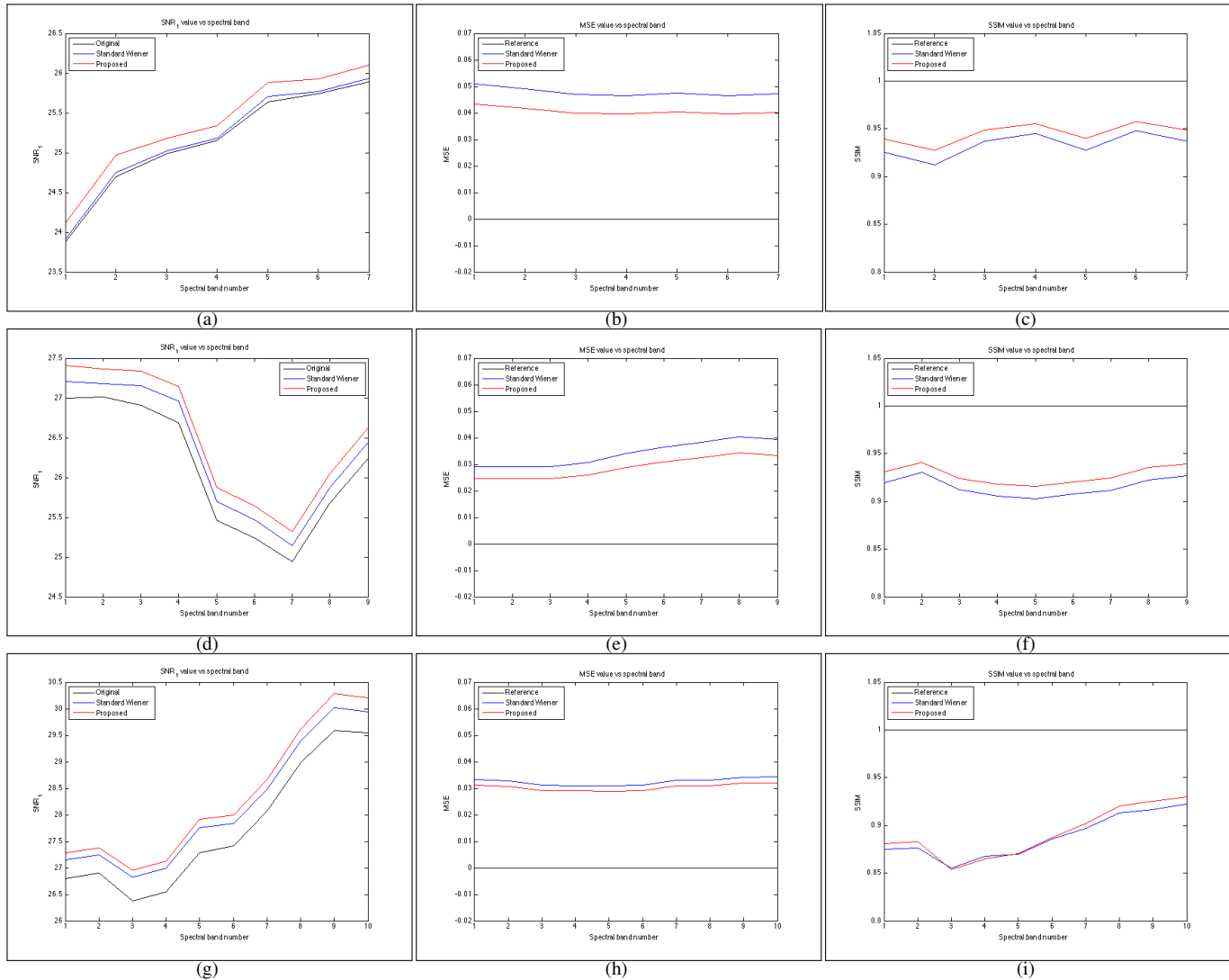


Fig. 9. SNR_1 , MSE and SSIM measures across different spectral bands. (a)-(c) Mouse body. (d)-(f) Mouse liver. (g)-(i) Mouse brain.

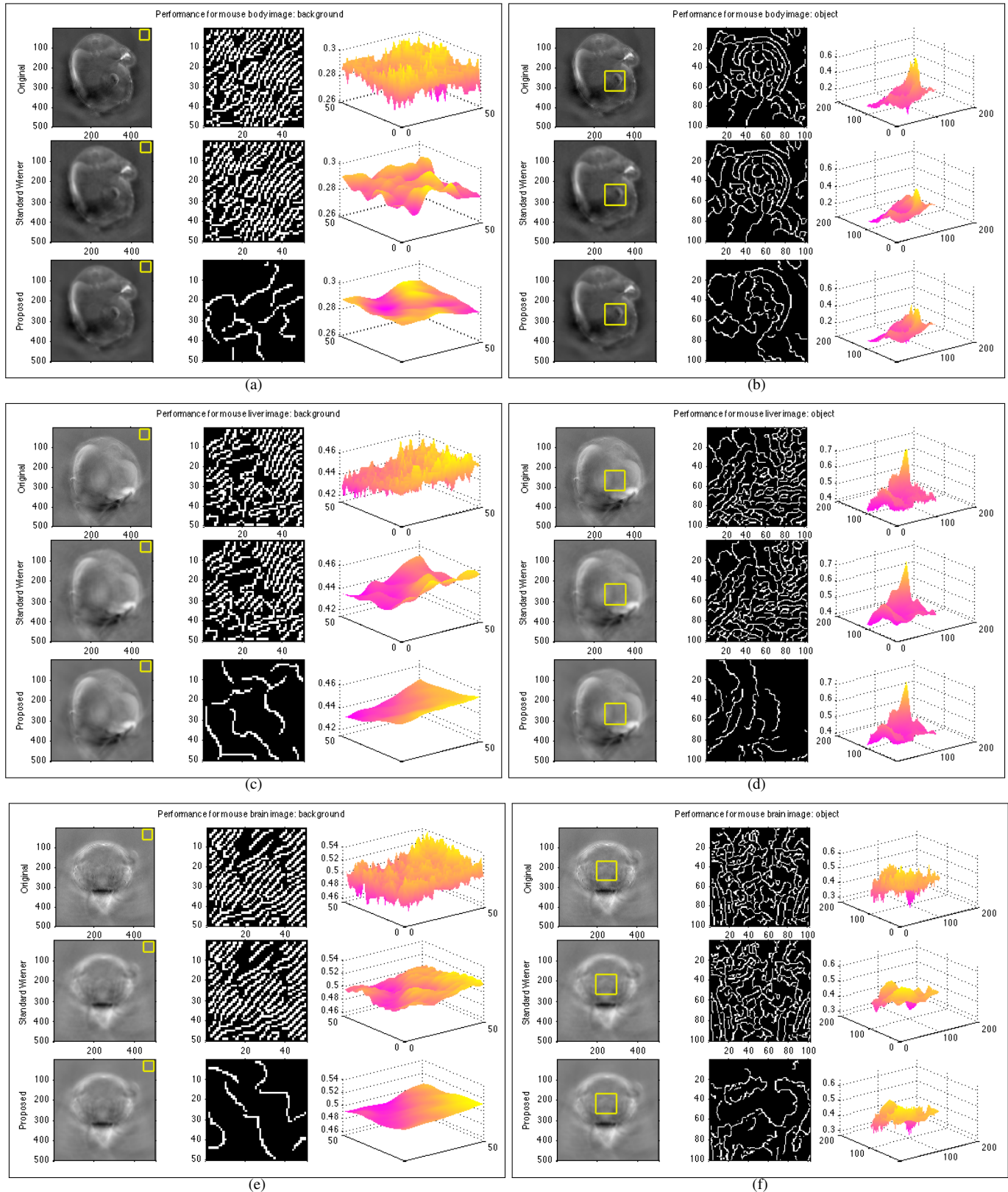


Fig. 10. Visual performance of the proposed method. (a)-(b) Mouse body. (c)-(d) Mouse liver. (e)-(f) Mouse brain. (a), (c), (e) The performance for image background. (b), (d), (f) The performance for image object. The edges in a small local background patch are detected with MATLAB built-in function *edge* and default parameter values. Surface plots of background and object patches (the location of which is marked by a yellow squares) are also included. Note that only the first spectral band of each image is displayed, for simplicity.

APPENDIX
PROOF OF PROPOSITION 4.1

If (12) holds and $F(y)_0(\lambda_{ij}) = 1$, we have the following:

$$\begin{aligned} F(y)_1(\lambda_{ij}) &= c_0 + c + F(v)_1(\lambda_{ij}), \\ F(y)_2(\lambda_{ij}) &= c_0 + cF(y)_1(\lambda_{ij}) + F(v)_2(\lambda_{ij}), \\ &\vdots \\ F(y)_k(\lambda_{ij}) &= c_0 + cF(y)_{k-1}(\lambda_{ij}) + F(v)_k(\lambda_{ij}), \\ &\vdots \\ F(y)_B(\lambda_{ij}) &= c_0 + cF(y)_{B-1}(\lambda_{ij}) + F(v)_B(\lambda_{ij}). \end{aligned}$$

Then

$$\begin{aligned} F(v)_1(\lambda_{ij}) &= F(y)_1(\lambda_{ij}) - c - c_0, \\ F(v)_2(\lambda_{ij}) &= F(y)_2(\lambda_{ij}) - cF(y)_1(\lambda_{ij}) - c_0, \\ &\vdots \\ F(v)_k(\lambda_{ij}) &= F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0, \\ &\vdots \\ F(v)_B(\lambda_{ij}) &= F(y)_B(\lambda_{ij}) - cF(y)_{B-1}(\lambda_{ij}) - c_0. \end{aligned}$$

Now we exploit the assumption that at each position (i, j, k) noise v_{ijk} is distributed normally with mean 0 and variance σ^2 . This implies that we add some relation between different time series, assuming that noise across different time series has the same variance. We construct the log-likelihood function:

$$\begin{aligned} \log L(\sigma^2, c_0, c) &= -MNB \log(2\pi\sigma^2) \\ &- \frac{1}{2\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0|^2. \end{aligned}$$

We can estimate σ^2 by maximizing the log-likelihood, i.e. by minimizing the negative log-likelihood:

$$\begin{aligned} \hat{\sigma}^2 &= \arg \min_{\sigma^2} [-\log L(\sigma^2, c_0, c)] \\ &= \arg \min_{\sigma^2} [MNB \log(2\pi\sigma^2) \\ &+ \frac{1}{2\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0|^2]. \end{aligned}$$

We calculate the derivative:

$$\begin{aligned} \frac{\partial [-\log L(\sigma^2, c_0, c)]}{\partial \sigma^2} &= \frac{MNB}{\sigma^2} \\ &- \frac{1}{2(\sigma^2)^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0|^2. \end{aligned}$$

Setting this to 0 we get

$$\begin{aligned} \frac{MNB}{\sigma^2} &= \\ &= \frac{1}{2(\sigma^2)^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0|^2, \end{aligned} \quad (15)$$

$\hat{\sigma}^2 =$

$$= \frac{1}{2MNB} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0|^2. \quad (16)$$

However, c_0 can be expressed via c as

$$\hat{c}_0 = \arg \min_{c_0} [-\log L(\sigma^2, c_0, c)].$$

Let's denote $F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0 = U + iV$, where U is the real part and V is the imaginary part of the complex number. Similarly, let $c_0 = A + iB$. Then the partial derivatives are

$$\begin{aligned} \frac{\partial [-\log L(\sigma^2, c_0, c)]}{\partial A} &= \frac{1}{\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B U, \\ \frac{\partial [-\log L(\sigma^2, c_0, c)]}{\partial B} &= \frac{1}{\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B iV. \end{aligned}$$

Equating these to zero we get

$$\frac{1}{\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B U = 0, \quad , \frac{1}{\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B iV = 0.$$

If these are equal to 0, their sum are also equal to zero and we get

$$\begin{aligned} \frac{1}{\sigma^2} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B (U + iV) &= 0, \\ \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B [F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - c_0] &= 0, \\ \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B [F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij})] &= MNBc_0, \\ \hat{c}_0 &= \frac{1}{MNB} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B [F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij})]. \end{aligned} \quad (17)$$

We note that estimating n is equivalent to estimating σ^2 because $n(\lambda_{ij}) = \text{Var}[F(V)(\lambda_{ij})] = \sigma^2$. Therefore,

$$\hat{n} = \frac{1}{2MNB} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^B |F(y)_k(\lambda_{ij}) - cF(y)_{k-1}(\lambda_{ij}) - \hat{c}_0|^2,$$

where \hat{c}_0 is as defined in (17).

ACKNOWLEDGMENT

This work was supported by the Singapore Bioimaging Consortium, Agency for Science, Technology and Research, Singapore and Department of Statistical Science, University College London, United Kingdom. Authors would also like to thank anonymous reviewers for constructive comments that helped to improve this work.

REFERENCES

- [1] N. Acito, M. Diani, G. Corsini, *Signal-dependent noise modeling and model parameter estimation in hyperspectral images*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 49 (8), 2011, pp. 2957–2971.
- [2] S. Tzoumas, A. Rosenthal, C. Lutzweiler, D. Razansky, V. Ntziachristos, *Spatiotemporal denoising framework for multispectral photoacoustic imaging based on sparse signal representation*, Medical Physics, Vol. 41(11), 2014, pp. 113301-1–113301-13.
- [3] X. L. Den-Ben, D. Razansky, *Adding fifth dimension to photoacoustic imaging: volumetric time-resolved spectrally enriched tomography*, Light: Science and Applications, Vol. 3, 2014, pp. 1–5.
- [4] X.L. Den-Ben, V. Ntziachristos, D. Razansky, *Artefact reduction in photoacoustic tomographic imaging by estimating the distribution of acoustic scatterers*, Journal of Biomedical Optics, Vol. 17(11), 2012, pp. 110504–110504.
- [5] G. Paltauf, R. Nuster, P. Burgholzer, *Weight factors for limited angle photoacoustic tomography*, Physics in Medicine and Biology, Vol. 54(11), 2009, pp. 3303.
- [6] S. Ma, S. Yang, H. Guo, *Limited-view photoacoustic imaging based on linear-array detection and filtered mean-backprojection-iterative reconstruction*, Journal of Applied Physics, Vol. 106, 2009, pp. 123104.
- [7] L. Yao, H. Jiang, *Enhancing finite element-based photoacoustic tomography using total variation minimization*, Applied Optics, Vol. 50 (25), 2011, pp. 5031–5041.
- [8] K. Wang, E.Y. Sidky, M.A. Anastasio, A.A. Oraevsky, X. Pan, *Limited data image reconstruction in photoacoustic tomography by constrained, total variation minimization*, Photon Plus Ultrasound: Imaging and Sensing, Vol. 7899, 2011, pp. 78993U-1–78993U-6.
- [9] J. Prakash, A.S. Raju, C.B. Shaw, M. Pramanik, P.K. Yalavarthy, *Basis pursuit deconvolution for improving model-based reconstructed images in photoacoustic tomography*, Biomedical Optics Express, Vol. 5(5), 2014, pp. 1363–1377.
- [10] Y.Q. Zhao, J. Yang, *Hyperspectral image denoising via sparse representation and low-rank constraint*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 53 (1), 2015, pp. 296–308.
- [11] B. Moore, R. Nadakuditi, J. Fessler, *Improved robust pca using low-rank denoising with optimal singular value shrinkage*, IEEE Workshop on Statistical Signal Processing, 2014, pp. 13–16.
- [12] F. Palsson, M. Ulfarsson, J.R. Sveinsson, *Hyperspectral image denoising using a sparse low rank model and dual-tree complex wavelet transform*, IEEE Conference on Geoscience and Remote Sensing Symposium, 2014, pp. 3670–3673.
- [13] N. Kamel, S. Kafa, *Image SNR estimation using the autoregressive modeling*, International Conference on Intelligent Advanced Systems (ICIAS), 2010, pp. 1–5.
- [14] A.K. Seghouane, A. Shah, *Consistent estimation of the fMRI hemodynamic response function in AR(1) noise*, IEEE International Symposium on Biomedical Imaging (ISBI), 2015, pp. 114–117.
- [15] M.R. Banham, A.K. Katsaggelos, *Digital image restoration*, IEEE Signal Processing Magazine, 1997, pp. 24–41.
- [16] T. Hou, Y. Wang, Y. Guo, *Despeckling medical ultrasound images based on blind deconvolution and maximum-likelihood estimation*, IEEE Conference on Bioinformatics and Biomedical Engineering (iCBBE), 2010, pp. 1–4.
- [17] R. Prager, A. Gee, G. Treece, N. Kingsbury, J. Lindop, H. Gomersall, H.C. Shin, *Deconvolution and Elastography based on Three-Dimensional Ultrasound*, IEEE International Ultrasonics Symposium Proceedings, 2008, pp. 548–557.
- [18] F. Lin, C. Jin, *An improved Wiener deconvolution filter for high-resolution electron microscopy images*, Micron, Vol. 50, 2013, pp. 1–6.
- [19] S. Wan, B.I. Raju, M.A. Srinivasan, *Robust deconvolution of high-frequency ultrasound images using higher-order spectral analysis and wavelets*, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, Vol. 50 (10), 2003, pp. 1286–1295.
- [20] Z. Wang, E.P. Simoncelli, A.C. Bovik, *Multi-scale structural similarity for image quality assessment*, IEEE Conference on Signals, Systems and Computers, Vol. 2, 2003, pp. 1398–1402.
- [21] C.J.H. Ho, G. Balasundaram, W. Driessen, R. McLaren, C.L. Wong, U.S. Dinish, A.B.E. Atta, V. Ntziachristos, M. Olivo, *Multifunctional photosensitizer-based contrast agents for photoacoustic imaging*, Scientific Reports, Vol. 4 (5342), 2014, pp. 1–6.
- [22] A.D. Krainov, A.M. Mokeeva, E.A. Sergeeva, P.D. Agrba, M.Yu. Kirillin, *Optical properties of mouse biotissues and their optical phantoms*, Optics and Spectroscopy, Vol. 115(2), 2013, pp. 193–200.
- [23] S.L. Jacques, *Corrigendum: Optical properties of biological tissues: a review*, Physics in Medicine and Biology, Vol. 58, 2013, pp. R37–R61.
- [24] T. Saratoon, T. Tarvainen, B.T. Cox, S.R. Arridge, *A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation*, Inverse Problems, Vol. 29, 2013, pp. 1–19.