Comparison of photoacoustic image reconstruction algorithms using the channelized Hotelling observer

Adam Petschke
Patrick J. La Rivière
Comparison of photoacoustic image reconstruction algorithms using the channelized Hotelling observer

Adam Petschke and Patrick J. La Rivière
The University of Chicago, Department of Radiology, 5841 South Maryland Avenue, Chicago, Illinois 60637

Abstract. We demonstrate the use of task-based image-quality metrics to compare various photoacoustic image-reconstruction algorithms, including a method based on the pseudoinverse of the system matrix, simple backprojection, and a method based on the Fourier transform. We use a three-dimensional forward model with a linear transducer array to simulate a photoacoustic imaging system. The reconstructed images correspond with two-dimensional slices of the object and are 128 x 128 pixels. In order to compare the algorithms, we use channelized Hotelling observers that predict the detection ability of human observers. We use two sets of channels: constant Q and difference of Gaussian spatial frequency channels. We look at three tasks, identification of a point source in a uniform background, identification of a 0.5-mm cube in a uniform background, and identification of a point source in a lumpy background. For the lumpy background task, which is the most realistic of the tasks, the method based on the pseudoinverse performs best according to both sets of channels. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JBO.18.2.026009]

Keywords: photoacoustic imaging; image reconstruction; channelized Hotelling observer; filtered backprojection; pseudoinverse; singular-value decomposition; Fourier-based reconstruction; medical imaging; biological imaging.

Paper 12477PRR received Jul. 24, 2012; revised manuscript received Jan. 14, 2013; accepted for publication Jan. 16, 2013; published online Feb. 5, 2013.

1 Introduction

Photoacoustic imaging is based on the detection of acoustic waves that are generated by thermoelastic expansion induced by optical absorption of laser light. The majority of photoacoustic imaging systems can be classified as either microscopy systems or computed tomography (CT) systems.1 In general, the microscopy-based systems do not require a reconstruction step as the lateral resolution is set by either a focused transducer or focused laser illumination. CT systems can be further subdivided into scanning systems (also called synthetic aperture) and array systems. From a mathematical point of view, synthetic apertures and arrays require the same reconstruction algorithm, although arrays tend to have smaller individual transducers so the directivity can be more uniform over the imaging area. Transducer arrays have a speed advantage as they do not need to be mechanically scanned, which can be important in many applications. In addition, the array-based imaging systems have the potential to be integrated with ultrasound imaging systems, and indeed many of the arrays in use in photoacoustic systems are taken from ultrasound systems.2,3,4 In this paper we consider a linear transducer array that is similar to the ones used in current photoacoustic and ultrasound imaging systems.

There are several reconstruction algorithms used in linear array imaging systems. Filtered backprojection is based on an analytical inversion of the solution to the photoacoustic wave equation.5 It is theoretically exact for a continuous, infinite linear transducer array, and when the object to be imaged is acoustically homogeneous and has no acoustic absorption. Unfiltered (or simple) backprojection, although not theoretically exact, is numerically equivalent to the delay-and-sum beamformer method employed in ultrasound imaging systems, and hence is very readily implemented in combined ultrasound-photoacoustic imaging systems. The Fourier method of image reconstruction is also based on inverting the solution to the photoacoustic wave equation, but this inversion is performed in the frequency domain.6 For sufficiently small datasets or given sufficiently powerful computer hardware, all of these simple reconstruction algorithms can be used in real-time imaging systems. There are more sophisticated reconstruction algorithms that can correct for acoustic inhomogeneity and absorption such as iterative methods.7 In general, these iterative algorithms produce high-quality images; unfortunately, using current computers they are not capable of real-time imaging.

The minimum-norm least-squares (MNLS) solution to reconstructing a photoacoustic image is given by the pseudoinverse of the system matrix, which relates the initial pressure distribution to the measured pressure wave. In this paper, we calculate the pseudoinverse using a singular-value decomposition (SVD). The pseudoinverse calculation involves dividing by the nonzero singular values of the system matrix. In practice, we have to regularize the problem by throwing out small singular values that would otherwise amplify the noise. Where to set the regularization threshold depends on the specific signal and noise levels. The noise in an imaging system can generally be divided into object-dependent and object-independent noise. In many imaging systems, such as x-ray CT, the object-dependent noise varies strongly across detectors. In photoacoustics, however, the object-dependent noise is relatively uniform across transducer elements (as long as the object is not too close to the transducer). The object-independent noise (such as electronic noise) is also uniform across the transducer elements.
The regularization level still depends on the relative strength of the signal and noise, but it is possible to calculate a few pseudoinverse matrices based on different singular-value truncations, store these in memory, and then decide in real time which one to use based on image quality. Due to the more uniform noise properties, we expect this method to work relatively better in photoacoustics than, say, x-ray CT, which would benefit from a more advanced regularization. Using this method the reconstruction step consists only of a matrix-vector multiplication, which is very fast. In a recent publication, the system matrix of a photoacoustic imaging system was measured experimentally, and the corresponding pseudoinverse was calculated ahead of time and used for real-time imaging.9 We perform a preliminary comparison between the filtered backprojection and pseudoinverse reconstruction algorithms. In a previous work,9 we performed a similar comparison between the filtered backprojection and pseudoinverse reconstruction algorithms. In that paper, we looked at various metrics such as resolution, streak length, and ideal observer signal-to-noise ratio (SNR) for a point source object; however, these metrics can be misleading. The resolution and streak length tell nothing of the noise properties of the algorithms, and the ideal observer SNR only measures how much information is lost for each algorithm in going from the raw data to the reconstructed image, which does not necessarily correlate with image quality as judged by human observer task performance.

In this paper we demonstrate the use in photoacoustics of so-called task-based metrics of image quality that seek to define image quality based not on physical measures but on the utility of an image for performing a predefined task. Such task-based image-quality assessment has gained widespread attention in nuclear medicine and x-ray imaging modalities.10 While the gold standard of task-based assessment is human observer studies using the receiver operating characteristic (ROC) methodology, these are tedious and labor-intensive. As an alternative, various model observers have been proposed that seek to emulate human observer performance in well-defined tasks, such as the detection of a known signal in a known or random background. The channelized Hotelling observer (CHO) has been shown to predict the performance of human observers quite well.10 It entails integrating images over a set of frequency-domain channels that mimic the frequency response of the human visual system and then applying a linear discriminant to the output. To assess the image quality of the reconstruction algorithms in this study, we use the CHO with two sets of channels that have been shown to predict the performance of human observers: constant Q11 and difference of Gaussian (DOG) channels.12

2 System Matrix

The photoacoustic equation can be written as1

$$\left(\nabla^2 - \frac{1}{v_i^2} \frac{\partial^2}{\partial t^2}\right) p(r, t) = -\frac{\beta}{C_P} \frac{\partial H(r, t)}{\partial t},$$

(1)

where $v_i$ is the speed of sound, $p(r, t)$ is the pressure, $\beta$ is the thermal coefficient of volume expansion, $C_P$ is the specific heat at constant pressure, and $H(r, t)$ is the heating function. The Green’s function solution for the pressure spectrum can be written as

$$\tilde{p}(r, \omega) = \frac{i \alpha \beta}{4\pi C_P} \int_{V_0} d^3 r_0 \tilde{H}(\mathbf{r}_0, \omega) e^{i(kr_0 - \omega t)},$$

(2)

where $k = \omega / v_i$. The geometry we are considering is shown in Fig. 1 and consists of a linear transducer array located on the $x$ axis with a cylindrical lens on its face and consisting of $N$ individual transducers, an object to be imaged in the half-space $z_0 > 0$, and a light source above the object pointing in the positive $y$ direction. The cylindrical lens rejects acoustic waves that do not originate near the imaging $(x_0 - z_0)$ plane. We assume uniform light illumination on the imaging plane to avoid modeling photon propagation. With this geometry, we can separate the heating function spectrum into $\tilde{H}(r_0, \omega) = \mu(r_0) \tilde{I}(\omega)$, where $\mu(r_0)$ is the absorption coefficient and $\tilde{I}(\omega)$ is the irradiance spectrum. The average pressure on each individual transducer is the pressure integrated over $x$ and $y$ divided by the area, where the width in the $x$ direction is $w$. The cylindrical acoustic lens imparts a phase of $e^{-ikd}e^{iky^2/(2f)}$ to the wave, where $d$ is the acoustic path length of the lens (refractive index times total thickness of the lens) and $f$ is the focal length. The lens has a height of $h$ and the transducer has a transfer function denoted by $T(\omega)$. This gives the measured voltage spectrum as

$$\tilde{V}_n(\omega) = \frac{i \alpha \beta \tilde{I}(\omega)}{4\pi \omega C_P} \int_{-w}^{w} \int_{-h}^{h} \int_{-w}^{w} \int_{-h}^{h} d^2 x d^2 y e^{iky^2/(2f)}$$

$$\times \frac{\mu(r_0) e^{-ikd} e^{iky^2/(2f)}}{|r - r_0|^2},$$

(3)

where $n$ denotes each individual transducer and runs from $-(N - 1)/2$ to $(N - 1)/2$. The middle transducer has $n = 0$, and the center of the middle transducer is located at $x = 0$. The extra $z_0 / |r - r_0|$ is an obliquity factor.

In this paper we consider a 32-element transducer with cylindrical lens focal length of 6 mm, height of 3 mm, individual element width of 0.15 mm, and center-to-center element spacing of 0.15 mm. The transducer transfer function is Gaussian with a
center frequency of 3.5 MHz and 50% bandwidth, and we assume the laser pulse is sufficiently short so that \( I(\omega)T(\omega) \approx I_0 T(\omega) \), where \( I_0 \) is a constant. Note that the transducer spacing satisfies the Nyquist sampling criterion. For the forward model, we use the full 3-D expression given by Eq. (3).

For the system matrix used to calculate the pseudoinverse, we use a 2-D approximation of Eq. (3) with \( y = y_0 = 0 \). For the discrete version, we use the notation \( g = H f \), where \( g \) is a vector representing the voltage measured by the transducers, \( H \) is the system matrix, and \( f \) is a vector representing the initial pressure distribution of the object being imaged. We use a 128 × 128 pixel basis with a pixel size of 37.5 μm. The temporal sampling rate is 100 ns with 71 samples per transducer, which produces a system matrix of size 2272 × 16384.

3 Image Reconstruction

3.1 Pseudoinverse

We performed an SVD on the system matrix, which was then used to calculate the pseudoinverse. The SVD factors the system matrix into \( H = USV^T \), where \( U \) and \( V \) are unitary matrices, the dagger represents conjugate transpose, and \( S \) is a diagonal matrix consisting of the singular values in decreasing order. The pseudoinverse is then given by \( H^+ = VS^T U^T \), where \( S^+ \) is a diagonal matrix consisting of the reciprocal of singular values above a threshold. The thresholding of the singular values serves as a regularization. Note that this method is also referred to as truncated SVD inversion. Figure 2 shows a plot of the magnitude of the singular values. The image, which is an estimation of the initial pressure \( f \), is then given by \( f = HH^+g \).

3.2 Filtered Backprojection and Simple Backprojection

The filtered backprojection formula\(^5\) can be written as

\[
\tilde{\mu}(k_x, k_z) = \frac{k_z}{k} \tilde{V}(k_x, \omega),
\]

where \( k^2 = k_x^2 + k_z^2 \). \( \tilde{V}(k_x, \omega) \) is constructed by taking the real part of the Fourier transform with respect to time, and taking the Fourier transform with respect to \( x \). The image is formed by employing bilinear interpolation and taking the inverse Fourier transform of \( \tilde{\mu}(k_x, k_z) \). Again, a Hamming window is used for apodization. We are ignoring multiplicative constants that would be required for quantitative results, as we did for the backprojection algorithms.

3.3 Fourier Method

The Fourier transform-based formula\(^6\) can be written as

\[
\tilde{\mu}(k_x, k_z) = \frac{k_z}{k} \tilde{V}(k_x, \omega),
\]

where \( k^2 = k_x^2 + k_z^2 \). \( \tilde{V}(k_x, \omega) \) is constructed by taking the real part of the Fourier transform with respect to time, and taking the Fourier transform with respect to \( x \). The image is formed by employing bilinear interpolation and taking the inverse Fourier transform of \( \tilde{\mu}(k_x, k_z) \). Again, a Hamming window is used for apodization. We are ignoring multiplicative constants that would be required for quantitative results, as we did for the backprojection algorithms.

3.4 Deconvolution

For the filtered backprojection, simple backprojection, and Fourier-based methods we deconvolve the transducer response using

\[
D(\omega) = \frac{\tilde{T}^*(\omega)}{|\tilde{T}(\omega)|^2 + \zeta},
\]

where the asterisk indicates complex conjugation and \( \zeta \) is a regularization parameter that varies from 10\(^{-6}\) to 10. For small values of \( \zeta \) the deconvolution is the inverse of the transducer response and noise amplification occurs. When \( \zeta = 10 \), \( D(\omega) \approx \tilde{T}^*(\omega) \), which is a matched filter. The matched filter lowers the noise at the expense of resolution. For all of the reconstructed images we perform envelope detection by taking the absolute value of the analytic function of the image.

Figures 3 and 4 show the image reconstruction of a point source and a 0.5-mm cube, respectively, for various singular-value truncations as well as the backprojection, filtered backprojection, and Fourier-based algorithms with different regularization parameters. Uncorrelated Gaussian noise is added to \( g \) before reconstruction to simulate electronic noise. The standard deviation of the noise is set to be equal to the peak signal level divided by 10. There is no background for these images. Figure 5 shows the image reconstruction of a lumpy background\(^7\) for the different algorithms. For each lumpy background image, there is a corresponding image with lumpy background and point source signal. All of the images in Fig. 5 contain the same background; in the next section the number of lumps and their locations are treated as random variables. The lumpy background is only present in a volume represented of lumps and their locations are treated as random variables.

A Hamming window was applied along the transducer elements for apodization. The simple backprojection algorithm is equivalent to the filtered backprojection algorithm except the pressure is not time-integrated or filtered with the Ram-Lak filter. Note that this filtered backprojection formula is built on a foundation developed by Norton in the context of reflection tomography.\(^8\) An alternative algorithm has been developed by Burgholzer et al. but has not been implemented in the present study.\(^9\)

Fig. 2 Magnitude of singular values of system matrix.
As mentioned in the Introduction, the CHO has been shown to predict human observer performance for various tasks. The choice of channels depends on the specific task to be performed. In this paper we use two sets of channels: constant $Q$ and DOG channels.\textsuperscript{12} Constant $Q$ channels have been shown to work well at predicting human performance when the noise is additive and Gaussian. DOG channels work well for additive Gaussian noise as well as for a lumpy background. The channel profiles are shown in Fig. 6.

In order to quantify image quality, we compute the channelized Hotelling SNR given by\textsuperscript{11}

$$\text{SNR}_{ch}^2 = \langle \hat{f}_{ch}^\dagger \rangle K_{ch}^{-1} \langle \hat{f}_{ch} \rangle,$$  \hfill (7)

where the brackets indicate expectation values, $\hat{f}_{ch} = U_{ch} \hat{f}$, $U_{ch}$ is a matrix that represents the spatial frequency filtering and summing of the channels, and $K_{ch}$ is the covariance matrix of $\hat{f}_{ch}$. Note that the covariance matrix is only a $5 \times 5$ matrix for the constant $Q$ channels and a $3 \times 3$ matrix for the DOG channels, which means the number of simulations required to make the matrix full rank is greatly reduced as compared to the full image covariance matrix. In this study we use 500 simulations to construct Eq. (7). For the uniform background tasks, uncorrelated Gaussian noise is added to $g$ before reconstruction, as discussed in the previous section. These tasks are referred to as signal known exactly, background known exactly (SKE/BKE). For the lumpy background\textsuperscript{17} task, the number of lumps and their locations are treated as random variables. The number of lumps, $L$, is modeled as a Gaussian random

---

**Fig. 3** Image reconstruction of a point source for singular-value truncations of (a) $10^5$; (b) $10^6$; (c) $10^7$; (d) $10^8$; filtered backprojection with regularization parameters of (e) $10^{-5}$; (f) $10^{-3}$; (g) $10^{-1}$; (h) $10^1$; backprojection with regularization parameters of (i) $10^{-3}$; (j) $10^{-3}$; (k) $10^{-3}$; (l) $10^1$; and the Fourier-based method with regularization parameters of (m) $10^{-3}$; (n) $10^{-3}$; (o) $10^{-3}$; (p) $10^1$.

**Fig. 4** Image reconstruction of a 0.5-mm cube for singular-value truncations of (a) $10^5$; (b) $10^6$; (c) $10^7$; (d) $10^8$; filtered backprojection with regularization parameters of (e) $10^{-5}$; (f) $10^{-3}$; (g) $10^{-1}$; (h) $10^1$; backprojection with regularization parameters of (i) $10^{-3}$; (j) $10^{-3}$; (k) $10^{-3}$; (l) $10^1$; and the Fourier-based method with regularization parameters of (m) $10^{-3}$; (n) $10^{-3}$; (o) $10^{-3}$; (p) $10^1$.

**Fig. 5** Image reconstruction of a lumpy background for singular-value truncations of (a) $10^7$; (b) $10^7$; (c) $10^8$; (d) $10^8$; filtered backprojection with regularization parameters of (e) $10^{-2}$; (f) $10^{-2}$; (g) $10^{1}$; (h) $10^{1}$; backprojection with regularization parameters of (i) $10^{-2}$; (j) $10^{-2}$; (k) $10^{1}$; (l) $10^{1}$; and the Fourier-based method with regularization parameters of (m) $10^{-2}$; (n) $10^{-2}$; (o) $10^{1}$; (p) $10^{1}$. A point source signal is also included in (b); (d); (f); (h); (j); (l); (n), and (p). The lumpy background is only present in a volume represented by a 0.5-mm cube at the center of the object.

### 4 Image Quality

As mentioned in the Introduction, the CHO has been shown to predict human observer performance for various tasks. The choice of channels depends on the specific task to be performed. In this paper we use two sets of channels: constant $Q$\textsuperscript{12} and DOG channels.\textsuperscript{12} Constant $Q$ channels have been shown to work well at predicting human performance when the noise is additive and Gaussian. DOG channels work well for additive Gaussian noise as well as for a lumpy background. The channel profiles are shown in Fig. 6.

In order to quantify image quality, we compute the channelized Hotelling SNR given by\textsuperscript{11}

$$\text{SNR}_{ch}^2 = \langle \hat{f}_{ch}^\dagger \rangle K_{ch}^{-1} \langle \hat{f}_{ch} \rangle,$$  \hfill (7)

where the brackets indicate expectation values, $\hat{f}_{ch} = U_{ch} \hat{f}$, $U_{ch}$ is a matrix that represents the spatial frequency filtering and summing of the channels, and $K_{ch}$ is the covariance matrix of $\hat{f}_{ch}$. Note that the covariance matrix is only a $5 \times 5$ matrix for the constant $Q$ channels and a $3 \times 3$ matrix for the DOG channels, which means the number of simulations required to make the matrix full rank is greatly reduced as compared to the full image covariance matrix. In this study we use 500 simulations to construct Eq. (7). For the uniform background tasks, uncorrelated Gaussian noise is added to $g$ before reconstruction, as discussed in the previous section. These tasks are referred to as signal known exactly, background known exactly (SKE/BKE). For the lumpy background\textsuperscript{17} task, the number of lumps and their locations are treated as random variables. The number of lumps, $L$, is modeled as a Gaussian random
variable with a mean and variance of 1250. That number was selected using $\frac{\langle L \rangle}{N} = 1/100$, where $N = 1.25 \times 10^5$ is the number of discrete object points in a 0.5-mm cube in our simulations. The location of the lumps is modeled as a uniform random variable over the volume represented by a 0.5-mm cube at the center of the object. Each lump is modeled as a point source with a magnitude that is 100 times lower than the point source signal magnitude. Uncorrelated Gaussian noise is also added to $g$ before reconstruction for this task. This task is referred to as SKE in a lumpy background and is a good model of tumor contrast agent imaging where the signal represents uptake of the contrast agent by the tumor and the lumpy background represents contrast agent uptake by the surrounding tissue.

We define the CHO efficiency as

$$\eta = \frac{\text{SNR}_{ch}^2}{\text{SNR}_{data}^2}; \quad (8)$$

where $\text{SNR}_{data}^2$ is given by

$$\text{SNR}_{data}^2 = (g^\dagger)K^{-1}(g). \quad (9)$$

In Figs. 7 through 9, we show the CHO efficiency for the three tasks. For the SKE/BKE tasks, the efficiency is highest

**Fig. 6** Channel profiles for (a) constant $Q$ channels and (b) DOG channels.

**Fig. 7** CHO efficiency for the (a) constant $Q$ channels and (b) DOG channels for the point source SKE/BKE task.

**Fig. 8** CHO efficiency for the (a) constant $Q$ channels and (b) DOG channels for the 0.5-mm cube SKE/BKE task.

**Fig. 9** CHO efficiency for the (a) constant $Q$ channels and (b) DOG channels for the point source SKE in a lumpy background task.
when the regularization is chosen so that the noise is lowest even though the resolution is worse as seen in Figs. 3 and 4. The performance of the algorithms is similar for the SKE/BKE tasks except that the Fourier-based algorithm performs best for the identification of the cube as determined by the difference of Gaussian channels. However, the SKE/BKE task is generally not a realistic model of actual imaging systems. For the SKE in a lumpy background task, which is more realistic, the method based on the pseudoinverse performs best according to both sets of channels. Also notice that for this task the optimal regularization is not necessarily to choose the lowest possible noise amplification. It is also interesting that for the three tasks studied in this paper, the simple backprojection performs as well as the filtered backprojection.

5 Conclusion

In this paper, we have used task-based image quality metrics to compare image reconstruction algorithms in photoacoustic tomography. The study was not intended to be the final word on algorithm selection but rather to introduce the use of task-based assessment and model observers to the field. As the field grows and matures, it could be valuable to deploy such tools for optimizing systems and algorithms for various tasks.

We have shown that for the case of a linear transducer array and an SKE in a lumpy background task, the image reconstruction algorithm based on the pseudoinverse performs best according to the CHOs we look at in this paper. Future work is planned to study the validity of the channels used in this paper by performing human observer experiments. Future work is also planned to study tasks that include signal variability. It should be noted that in this paper we assume uniform light illumination on the imaging plane. Depending on the specific system and task to be performed this may not be a realistic assumption, and in this case the laser irradiation geometry and photon propagation should be included in the forward model.

Another interesting conclusion of this study is that the simple backprojection algorithm performs as well as the filtered backprojection algorithm for the tasks we looked at. Since the simple backprojection algorithm is numerically equivalent to the delay-and-sum beamformer reconstruction employed in existing ultrasound imaging systems, combined ultrasound-photoacoustic imaging systems may be able to use this algorithm without loss of performance. This would be advantageous in terms of system design. The type of analysis presented in this paper could be used to determine if this conclusion is valid for the specific task the combined ultrasound-photoacoustic imaging system is being built for.

Acknowledgments

This work was supported by ACS Research Scholar Grant RSG-08-117-01-CCE to Patrick La Rivière.

References