Estimating relative chromophore concentrations from multiwavelength photoacoustic images using independent component analysis

Lu An* and **Benjamin T. Cox**
University College London, Department of Medical Physics and Biomedical Engineering, London, United Kingdom

Abstract. Independent component analysis (ICA) is an unmixing method based on a linear model. It has previously been applied in *in vivo* multiwavelength photoacoustic imaging studies to unmix the components representing individual chromophores by assuming that they are statistically independent. Numerically simulated and experimentally acquired two-dimensional images of tissue-mimicking phantoms are used to investigate the conditions required for ICA to give accurate estimates of the relative chromophore concentrations. A simple approximate fluence correction was applied to reduce but not completely remove the nonlinear fluence distortion, as might be possible in practice. The results show that ICA is robust against the residual effect of the partially corrected fluence distortion. ICA is shown to provide accurate unmixing of the chromophores when the absorption coefficient is within a certain range of values, where the upper absorption threshold is comparable to the absorption of blood. When the absorption is increased beyond these thresholds, ICA abruptly fails to unmix the chromophores accurately. The ICA approach was compared to a linear spectroscopic inversion (SI) with known absorption spectra. In cases where the mixing matrix with the specific absorption spectra is ill-conditioned, ICA is able to provide accurate unmixing when SI results in large errors. © *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.23.7.076007]

Keywords: spectral unmixing; independent component analysis; quantitative photoacoustic tomography; photoacoustic imaging.

Paper 170442PRRR received Jul. 9, 2017; accepted for publication Jun. 15, 2018; published online Jul. 10, 2018.

1 Introduction

Biomedical photoacoustic imaging has been a rapidly growing field of research in the last two decades, with significant development in the instrumentation, reconstruction algorithms, and exogenous contrast agents. This has enabled a range of clinical applications of photoacoustic imaging, such as breast imaging,[1,2] cardiovascular imaging,[3–5] and skin imaging,[6,7] as well as preclinical studies using endogenous contrast, such as hemoglobin, for imaging the vasculature in the mouse brain[8–10] and in tumor models[11–13] or exogenous contrast agents[14,15] and reporter genes[16,17] for molecular or cellular imaging. To acquire a photoacoustic image, the tissue is illuminated with nanosecond laser pulses, and the optical energy is absorbed by the chromophores in the tissue. This leads to a small temperature and pressure rise, generating broadband acoustic waves that propagate to the surface of the tissue where they are detected and used to reconstruct images of the pressure rise.

By acquiring multiple images using excitation light with different wavelengths, photoacoustic imaging can be extended to a spectroscopic technique that can in principle quantify the concentration of the chromophores and hence reveal functional information about the subject. However, the image contrast depends on the absorbed optical energy, which is nonlinearly related to the chromophore concentrations, because it is a product of the local light fluence and the absorption coefficient. The light fluence is not only spatially and spectrally nonuniform

but also depends on the unknown chromophore concentrations. Therefore, it presents the key challenge for quantitative photoacoustic imaging. The most straightforward method of accounting for the nonlinear fluence distortion is to divide the multiwavelength photoacoustic images by an approximation of the fluence at each wavelength, which can be obtained by for example assuming homogeneous optical properties.[18,19] In this way, the effect of the fluence is approximately "cancelled out," such that the images represent a linear sum of chromophore concentrations that can be unmixed using linear spectroscopic inversion (SI), as in conventional optical spectroscopy, which uses the pseudoinverse of the matrix of the known specific absorption spectra. Linear methods with approximate fluence correction are easy to implement and fast to compute; hence, they are attractive for *in vivo* studies. However, the accuracy of SI is limited for nonsuperficial imaging depths,[20] because it depends on the accuracy of the fluence correction, which is only an approximation of the true fluence distribution in practice. Improved quantification accuracy in some scenarios has been demonstrated using more sophisticated inversion schemes, including methods involving light source modulation,[21] sparse signal decomposition,[22] fluence modeling using light transport equations,[23–25] or generating spectral libraries.[26] However, these methods are more challenging to implement for *in vivo* images due to issues such as model mismatch, increased level of complexity, and/or high computational demand.

One way to enable quantification in a wider range of applications would be to retain the simplicity of a linear model but increase the robustness to errors in the fluence correction. This

---

*Address all correspondence to: Lu An, E-mail: lu.an.13@ucl.ac.uk

may, in some cases, be achieved using independent component analysis (ICA),[27–29] which is a linear source separation method that decomposes the multiwavelength image data into components representing the individual chromophores by assuming that they are mutually statistically independent. ICA has the advantages of being fast, easy to implement, and computationally inexpensive. It was first proposed for unmixing photoacoustic images by Glatz et al.,[28] who showed that it can provide spectral unmixing with less cross talk error than SI. It has also been shown in some cases that applying ICA to the logarithm of the photoacoustic images can result in more accurate estimations of the relative concentrations compared to using ICA on the images without taking the logarithm.[30] However, since the components of interest are represented as a ratio to the background chromophore concentration, this approach requires that the chromophore concentration in the background tissue is known. ICA has been used in an *in vivo* multiwavelength photoacoustic imaging study of near-infrared fluorescent protein-expressing glioma cells,[31] as well as in studies using exogenous probes that bind to injured cells[32] or cells undergoing apoptosis.[33] In the above-mentioned studies, the main purpose of applying ICA was to aid the visualization of the probe by distinguishing it from the background tissue, rather achieving quantitatively accurate representation of the relative chromophore concentrations. As a result, despite having been applied in several *in vivo* studies, the quantitative accuracy of ICA has been not rigorously assessed for photoacoustic imaging. Therefore, the aim of this paper is to investigate the conditions required for ICA to provide quantitatively accurate unmixing, which required an approximate fluence correction step. Simulated multiwavelength photoacoustic images are used to assess the robustness of ICA against inaccuracies in the fluence correction and demonstrate the effect of retaining different numbers of dimensions as a preprocessing step to ICA. The results of ICA are also compared to SI for experimentally acquired images of a tissue-mimicking phantom. Furthermore, the performances of ICA and SI are analyzed for inversions using the experimental images with varying spectral ranges. We demonstrate both the advantages of ICA and its limitations, by including examples of cases where it provides accurate quantification, as well as cases where it breaks down.

2 Principles of Independent Component Analysis

Assuming that the acoustic reconstruction of the initial pressure rise due to the absorption of optical energy is perfect, so there are no image artifacts, a set of photoacoustic images with M voxels acquired at N wavelengths, whose contrast originates from K chromophores, can be described using matrix notation as

$$\mathbf{P} = \mathbf{\Phi} \circ [\mathbf{A}(\mathbf{\Gamma} \circ \mathbf{C})], \tag{1}$$

where $\circ$ denotes the Hadamard product, $\mathbf{P}$ is a $(N \times M)$ matrix with each row corresponding to the reconstruction of the initial pressures at one of the wavelengths, $\mathbf{\Phi}$ is the $(N \times M)$ fluence matrix with each row representing the spatially varying fluence at one of the wavelengths, $\mathbf{A}$ is the $(N \times K)$ mixing matrix with each column representing the wavelength-dependent specific absorption coefficient of one of the chromophores, $\mathbf{C}$ is the $(K \times M)$ concentrations matrix with each row representing the concentration distribution of one of the chromophores, and

$\mathbf{\Gamma}$ is a $(K \times M)$ matrix, where all rows are identical and equal to the spatially varying Grüneisen parameter.

Both ICA and SI are unmixing methods based on a linear model. The initial photoacoustic pressure is, however, nonlinearly related to the concentrations, due to the spectrally and spatially varying fluence matrix in Eq. (1). An approximation to linearity can be achieved by dividing $\mathbf{P}$ element-wise by an approximate estimate of the fluence, $\check{\mathbf{\Phi}}$, such that

$$\frac{\mathbf{P}}{\check{\mathbf{\Phi}}} = \check{\mathbf{P}} \approx \mathbf{A}(\mathbf{\Gamma} \circ \mathbf{C}). \tag{2}$$

In SI, the mixing matrix with the specific absorption coefficients is known. Hence, the chromophore concentrations can be estimated using the pseudoinverse of the mixing matrix, $\mathbf{A}^{\dagger}$

$$\mathbf{A}^{\dagger}\check{\mathbf{P}} \approx \mathbf{\Gamma} \circ \mathbf{C}. \tag{3}$$

SI cannot separate $\mathbf{\Gamma}$ from the concentrations in general, as $\mathbf{\Gamma}$ does not vary with wavelength and scales with the concentrations. In *in vivo* imaging studies, $\mathbf{\Gamma}$ is often assumed to be a spatially constant scaling factor, even though it may vary significantly between different tissue types.[34]

Unlike in SI, the mixing matrix is not known in ICA. Instead, ICA aims to decompose the multiwavelength photoacoustic images into the source components, which represent the individual chromophores, based on the assumption that they are statistically mutually independent of each other. Statistical independence is defined as follows: two events, X and Y, are independent of each other if the probability of X occurring does not in any way influence the probability of Y occurring. In quantitative photoacoustic imaging, this requires that the knowledge of the concentration of a chromophore at a location is not affected by the information provided about another chromophore at the same location. In other words, if the distributions of certain probes are unrelated to other chromophores, such as blood and lipids, then the concentration of the probe does not reveal any information about the amount of blood in the same pixel; hence, they are independent from each other. Example of such independent probes can be found in the imaging of reporter genes, such as genetically encoded fluorescent proteins[18] or tyrosinase-expressing cells,[35] and exogenous contrast agents that accumulate in regions unrelated to the blood distribution, such as in lymphatic vessels and lymph nodes.[36] Examples of chromophores that are unlikely to be independent include oxy- and deoxyhemoglobin, as their distributions are typically highly correlated. Therefore, ICA is not suitable for unmixing oxy- and deoxyhemoglobin for the estimation of blood oxygenation.

In ICA, the independent components are found from the multiwavelength images by searching for a mixing matrix, $\mathbf{W}$, whose inverse can be multiplied by the matrix with the mixed signals to give the matrix of the source components, $\mathbf{S}$

$$\mathbf{W}^{\dagger}\check{\mathbf{P}} = \mathbf{S}, \tag{4}$$

where the rows of $\mathbf{S}$ have maximal statistical independence. The unmixed component in each row of $\mathbf{S}$ corresponds to one of the chromophores, such that $\mathbf{S} \approx \mathbf{\Gamma} \circ \mathbf{C}$, provided that the true spatial distributions of the chromophore concentrations are independent.

In the widely used ICA algorithm FastICA,[27] the search for the most independent components is based on the central limit theorem, which states that the probability distribution of the sum of multiple independent variables will always be more Gaussian than the distribution of one of the independent variables alone. The Gaussianity of the probability distribution of the output components is measured using negentropy, which is a normalized version of entropy. Hence, the independent source components are found by iteratively searching for a mixing matrix that maximizes the negentropy of each row of $\mathbf{S}$.

Unlike SI, ICA does not recover the concentrations for different chromophores to a common scale, because the magnitude of each independent source is unknown. It is clear from the ICA model in Eq. (4) that since neither $\mathbf{W}$ nor $\mathbf{S}$ is fixed, the magnitude of each source component can be arbitrarily changed by multiplying any row in $\mathbf{S}$ and dividing the corresponding column in $\mathbf{W}^\dagger$ with the same constant. This means that while the relative concentration to other spatial locations for each chromophore can be recovered, the concentration of one chromophore cannot be compared to that of another chromophore.

3 Generating Multiwavelength Photoacoustic Images

If the fluence was estimated exactly, such that $\check{\Phi} = \Phi$, the approximate model in Eq. (2) would be an exact equality, and if the chromophore distributions were independent, both ICA and SI would achieve accurate unmixing of the chromophores, such that the outputs of both inversions are exactly equal to $\Gamma \circ \mathbf{C}$. In practice, however, $\check{\Phi}$ is only an approximation of Φ, and both methods will lead to errors. Both experimental and numerically simulated multiwavelength photoacoustic images of tissue-mimicking phantoms were generated to investigate the accuracy of ICA for various levels of absorption and heterogeneity in the tissue structure, as well as the effect of dimension reduction and wavelength selection, in order to understand the extent to which ICA can be used as a reliable quantification method. The same data sets are also unmixed with SI for comparison. The experimental and numerical phantoms were designed such that the following criteria were satisfied:

1. The spatial distributions of the chromophores are independent, so that when the fluence adjustment is perfect, ICA results in accurate unmixing.

2. The specific absorption spectra of the chosen chromophores are such that the mixing matrix $\mathbf{A}$ is full rank, so that when the fluence adjustment is perfect, SI also results in accurate unmixing.

Hence, using these phantoms, it is possible to see how the inaccuracies in an approximate fluence adjustment affect the performance of the unmixing methods.

3.1 Experimental Photoacoustic Image Acquisition

Eight capillary tubes (Paradigm Optics, Morcap 83) with inner diameter 590 μm and wall thickness 66.5 μm were used to construct the tissue-mimicking phantom. The tubes were arranged horizontally in two vertical lines as shown in Fig. 1, which shows a cross section of the phantom in the experimental setup. The left column of tubes was filled with four different concentrations of copper(II) chloride dihydrate (CuCl$_2 \cdot$ 2H$_2$O)
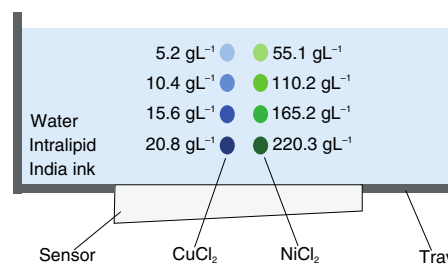

Fig. 1 A cross section of the experimental phantom which consisted of eight tubes filled with different concentrations of CuCl$_2$ (left column) and NiCl$_2$ (right column). The center of the tubes was positioned at depths 1.0, 2.7, 4.4, and 6.1 mm in a scattering and absorbing background solution of water, 1% Intralipid, and India ink. The distance between the water surface, where the phantom was illuminated, and the acoustic sensor at the bottom of the tray was 7.3 mm.

dissolved in deionized water. The concentrations were in the ratios 1:2:3:4, where the uppermost tube had the lowest concentration of 5.2 gL^{-1} while the bottom tube had the highest concentration of 20.8 gL^{-1}. The right column of tubes was filled with solutions of nickel(II) chloride hexahydrate (NiCl$_2 \cdot$ 6H$_2$O), also with concentration ratio of 1:2:3:4. These contrast agents simulate different absorbers in the tissue and will be referred to as CuCl$_2$ and NiCl$_2$. The absorption spectra of all chromophores are shown in Fig. 2. Since the specific absorption coefficient of NiCl$_2$ is approximately one order of magnitude lower than that of CuCl$_2$ [see Fig. 2(a)], the concentrations of NiCl$_2$ were set to be ~10 times higher than CuCl$_2$ to give similar optical absorption. Therefore, the uppermost and bottom tubes in the right column had NiCl$_2$ concentrations of 55.1 and 220.3 gL^{-1}, respectively. The average absorption of all tubes was 0.25 mm^{-1} at 810 nm.

The tubes were submerged at depths between 1.0 and 6.1 mm in a background solution containing India ink (951 black, Winsor & Newton) and 1% Intralipid diluted in deionized water, such that they provide an absorption of ~0.013 mm^{-1} at 810 nm and a scattering coefficient of ~1 mm^{-1}, which are comparable to realistic values in soft tissue.[38] The Grüneisen parameter of CuCl$_2$ and NiCl$_2$ is both known to scale with their concentrations such that

$$\Gamma_i = \Gamma_{H_2O}(1 + \beta_i c_i), \quad i = CuCl_2 \quad \text{or} \quad NiCl_2, \qquad (5)$$

where the β_i coefficients are 5.80×10^{-3} Lg^{-1} and 2.25×10^{-3} Lg^{-1} for CuCl$_2$ or NiCl$_2$, respectively,[39] c_i denotes the concentrations of CuCl$_2$ or NiCl$_2$, and Γ_{H_2O} is the Grüneisen parameter of water, which is 0.12 Lg^{-1} at 22°C.[40] The phantom was illuminated at the surface using pulsed laser radiation (SpitLight 600 OPO, InnoLas Laser GmbH) of 6-ns duration. The beam diameter was ~10 mm and the pulse energy was 7 to 13 mJ at the surface of the phantom depending on wavelength. A small fraction of the excitation light was reflected into an integrating sphere with a photodiode to measure the laser pulse energy, which was used to normalize the photoacoustic signals. A planar Fabry–Perot polymer film interferometer (FPI)[41] situated at the bottom of the tray containing the phantom was used for ultrasound detection. The acoustic pressure waves that propagate to the sensor modulate the thickness and hence the reflectivity of the FPI. Thus, by scanning the surface of the FPI with an interrogation laser and recording the reflected intensity, a time-varying spatial mapping of the photoacoustic signals
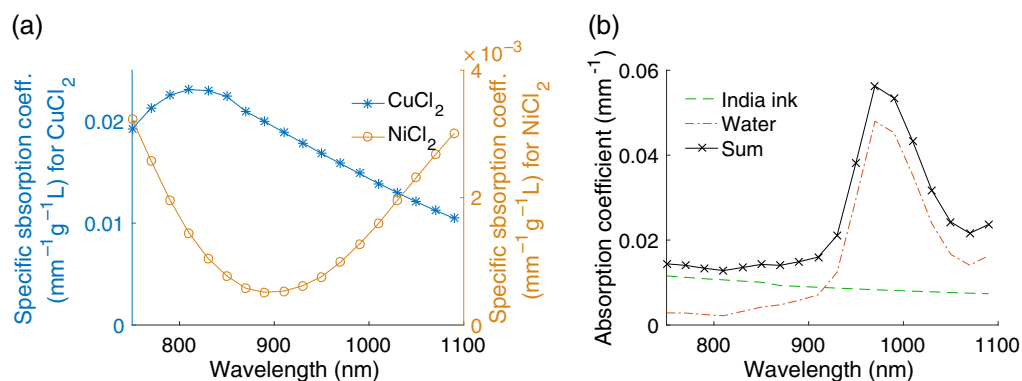
Fig. 2 (a) The specific absorption coefficient of $CuCl_2$ (asterisks) is approximate one order of magnitude larger than that of $NiCl_2$ (circles). (b) The crosses indicate the absorption of the background solution, which is the sum of the absorption coefficients of water (dash and dotted curve) and India ink (dashed curve). The absorption of the Intralipid is negligible in this wavelength range. The absorption spectra of $CuCl_2$, $NiCl_2$, and India ink are based on spectrophotometer measurements (Lambda 750S, PerkinElmer), and the absorption spectra of water were published by Kou et al.[37]

was generated. The sensor was interrogated along a 20-mm line with 10-μm step size to reconstruct two-dimensional (2-D) images of the cross section of the tubes at 18 equally spaced wavelengths between 750 and 1090 nm.

3.2 Numerically Simulated Photoacoustic Images

A total of 37 sets of simulated 2-D multiwavelength photoacoustic images based on the same structure and chromophores as the experimental phantom were generated to evaluate the performance of ICA for different levels of absorption. The dimensions of the simulated images were 7.4×20.0 mm^2 and spacing between the elements was 20 μm. The main differences to the experimental phantom are the absence of the tube walls and the fact that the India ink and Intralipid are also present within the tubes. The concentration ratios of 1:2:3:4 between the tubes in each column, as well as the 10-fold ratio between $CuCl_2$ and $NiCl_2$ concentrations were kept constant for all data sets, while the absolute concentrations of $CuCl_2$ and $NiCl_2$ were varied in different ways in three case studies:

Case I consists of 15 data sets, where the concentrations inside the tubes were increased such that the average absorption coefficient of all tubes increased from 0.05 to 0.75 mm^{-1} at 810 nm in equal steps. The ink concentration was kept constant for all data sets such that its absorption coefficient is the same as shown in Fig. 2(b).

Case II consists of 11 data sets, where the concentration of the ink was increased such that the absorption coefficient of the background solution at 810 nm was increased from 0.003 to 0.20 mm^{-1}. The $CuCl_2$ and $NiCl_2$ concentrations inside the tubes were kept constant at the same values as the experimental phantom, such that the average absorption of the tubes was 0.25 mm^{-1} at 810nm.

Case III aims to investigate the impact of spatially inhomogeneous absorption in the background. A region containing $CuCl_2$ which surrounded two of the tubes was included, as shown in Fig. 3. The $CuCl_2$ concentration in this region was increased such that its absorption coefficient was increased from 0.015 to 0.362 mm^{-1} at 810 nm in 11 data sets. The concentrations of the chromophores outside this additional region were
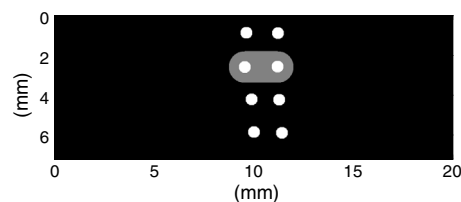
Fig. 3 The structure of the numerical phantom. The white circular regions represent the tubes and the gray region represents the additional region with increasing $CuCl_2$ concentration in Case III.

kept constant at the same values as the experimental phantom.

The absorption coefficient was calculated for the same 18 wavelengths as the experimental phantom for each data set based on the chromophore concentrations and their known specific absorption coefficients. The reduced scattering coefficient was kept the same as the experimental phantom for all data sets. Based on the distribution of the absorption and reduced scattering coefficients, the fluence was modeled using the diffusion approximation to the radiative transfer equation with the MATLAB software package TOAST++[42] for a 10-mm wide Gaussian light source at the top of the phantom. The initial pressure was found by multiplying the fluence by the optical absorption coefficient and the Grüneisen parameter.

The photoacoustic wave propagation from the initial pressure distribution was modeled using the MATLAB toolbox k-Wave[43] based on a k-space pseudospectral method. The time-varying photoacoustic pressure was recorded at the bottom of the numerical phantom.

3.3 Image Processing

Both the experimental and simulated images were reconstructed using the time-reversal method,[44] which backpropagates the recorded time series of the photoacoustic pressure into the image domain to form an image of the initial pressure. Gaussian noise with standard deviation equal to 3% of the peak intensity of data set in Case I with the same concentrations as the experimental phantom was added to all simulated images.
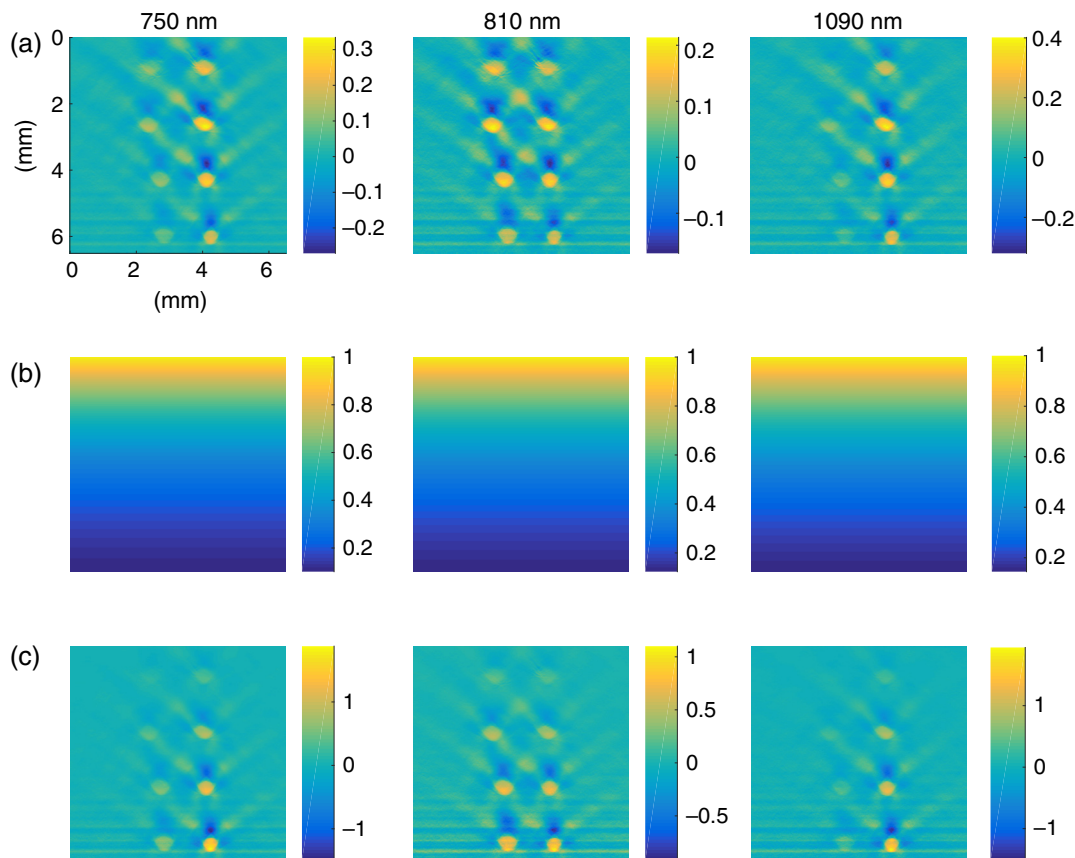
Fig. 4 (a) The experimental photoacoustic images of the tube phantom acquired at three different wavelengths, (b) the fluence estimations based on the 1-D exponential decay, and (c) the fluence adjusted photoacoustic images. The fluence adjusted images are more similar to the expected absorption coefficients than the raw images.

An approximate fluence adjustment was performed by dividing the reconstructions of the initial pressure by the estimation of the fluence, $\check{\Phi}$. The $\check{\Phi}$ was calculated based on the simplifying assumptions that the medium is a semi-infinite optically homogeneous slab illuminated by infinitely wide plane waves, such that the one-dimensional (1-D) solution to the diffuse approximation can be applied

$$\Phi(z,\lambda) = \Phi_0 \, \exp[-\mu_{\text{eff}}(\lambda)z], \tag{6}$$

where z is the depth from the illuminated surface, Φ_0 is the fluence at the illuminated surface, and μ_{eff} is the effective absorption coefficient given by $\mu_{\text{eff}} = \sqrt{3\mu_a(\mu_a + \mu_s')}$, where the μ_a and μ_s' are assumed to be known and equal to that of the background solution. This exponential fluence adjustment is chosen in this study because it can be applied straightforwardly in practice for *in vivo* photoacoustic images, with μ_{eff} estimated as an average parameter for the tissue. A similar correction can also be applied for imaging systems with circular illumination, in which case the fluence can be approximated as a radially symmetric function.[45] While the fluence correction step is necessary for obtaining accurate estimates of the relative chromophore concentrations, it is still an approximate method and does not fully remove the fluence distortion.

The fluence adjustment overamplifies the noise in the regions away from the light source where the values of $\check{\Phi}$ are small. To avoid this, a 6.5×6.5-mm^2 region of interest in the fluence

adjusted experimental images was decomposed using ICA and SI. The raw experimental images, as well as the fluence estimations and the fluence adjusted images, are shown in Fig. 4. As expected based on the specific absorption coefficients of the chromophores, the raw images at 750 and 1090 nm show higher intensity for the right column of tubes containing NiCl$_2$, while the image acquired at 810 nm shows higher intensity for the left column of tubes containing CuCl$_2$. The streak patterns extending from the tubes and the negative values in the reconstruction are artifacts due to the limited detection aperture of the planar sensor. The bottom row of figures shows that after the fluence adjustment, the image intensity at the tubes increases with depth, which is expected as the chromophore concentrations are higher inside the deeper tubes. Since the negative values in the images lack physical meaning, they were set to zero before the unmixing.

The FastICA algorithm requires further preprocessing of the multiwavelength fluence adjusted images before unmixing. The images are mean subtracted and whitened using principal component analysis, and the dimensions of the image data are reduced such that only the three principal components (PCs) with the largest eigenvalues of the covariance matrix of $\check{P}$ are unmixed with FastICA. Since the ordering of the estimated independent components is arbitrary, the output components are identified to the corresponding chromophores by calculating the sum of squared differences between the normalized columns in the estimated mixing matrix and the normalized known absorption spectra of the chromophores of interest.

For comparison, the fluence adjusted images were also unmixed using SI, based on the known specific absorption spectra of $CuCl_2$, $NiCl_2$, and the background solution.

## 4 Unmixing Using Independent Component Analysis and Spectroscopic Inversion

4.1 Accuracy as a Function of Absorption

The fluence adjusted and preprocessed simulated photoacoustic images at 18 wavelengths were decomposed using ICA and SI. The unmixed components were normalized to their average value at the tube with the highest concentration of the relevant chromophore. The normalized components were compared to the expected components, which are the true concentrations of the relevant chromophores multiplied by the Grüneisen parameter. Three types of errors are defined to provide a quantitative assessment of the accuracy of the unmixing methods:

1. The concentration error, δ_c, is defined as the average error at the pixels in the tubes where the chromophore of interest is present.

2. The average error at the tubes where the relevant chromophore is absent is defined as the "false positive" error, δ_f.

3. The background error, δ_b, is the average error outside the eight tubes.

For example, in the estimated $CuCl_2$ component, δ_c is the average error at the left column of tubes, δ_f is the average error at the right column of tubes, and δ_b is the average error
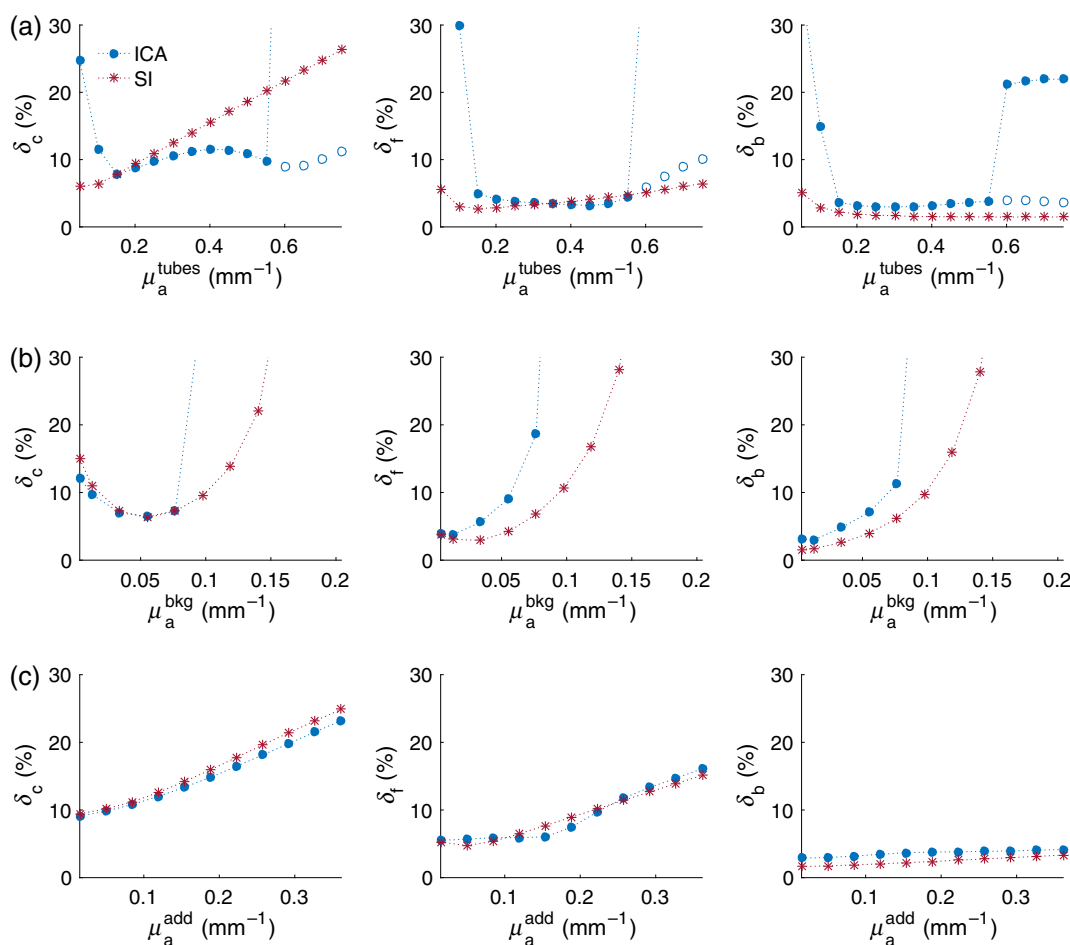
Fig. 5 The left, center, and right columns of plots show the δ_c, δ_f, and δ_b errors, respectively, for the chromophore components unmixed from the simulated images at 18 wavelengths from (a) Case I, (b) Case II, and (c) Case III using ICA (dots) and SI (asterisks). The x-axis indicates the average absorption of the eight tubes for Case I, the absorption of the background solution for Case II, and the absorption of the additional region with $CuCl_2$ for Case III, all at 810 nm. For Case I, the δ_c error for ICA is $\lesssim 10\%$ for μ_a^{tubes} between 0.15 and 0.55 mm^{-1}, while the δ_c error for SI increases approximately constantly with μ_a^{tubes}. The circles indicate the ICA errors for manual selection of the corresponding components. Case II shows that the errors associated with ICA and SI are both relatively low for physiologically realistic range of absorption in the background tissue. The δ_c and δ_f errors in Case III increase with the absorption in the additional region for both ICA and SI. The δ_c, δ_f, and δ_b errors in the plot represent the average error at the pixels in the relevant regions. The variability (standard deviation) of the δ_c and δ_f errors between the individual tubes typically scales with the size of the errors and is on average 3% for errors <10% for both ICA and SI.
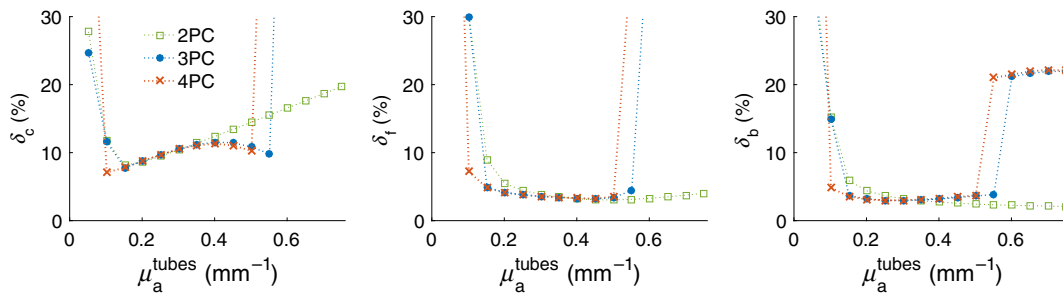
Fig. 6 The δ_c, δ_f, and δ_b errors of the unmixed components when two, three, or four PCs are further processed with ICA. Using two PCs (squares) result in larger δ_c for higher absorption, but no abrupt increase in errors is observed. The range of absorption, in which ICA leads to accurate unmixing, is shifted to lower absorption when four PCs (crosses) are used instead of three (dots).

outside the tubes. For Case III, the errors in the additional region are included in the definition of δ_c and δ_f.

The three types of errors are plotted in Fig. 5 as functions of the average absorption of the tubes at 810 nm, μ_a^{tubes}, for Case I, the absorption of the background solution at 810 nm, μ_a^{bkg}, for Case II, and the absorption of the additional region at 810 nm, μ_a^{add}, for Case III. In Case I, the δ_c error for SI increases with μ_a^{tubes} at a nearly constant rate for the whole range of absorptions between 0.05 and 0.75 mm^{-1}, as shown with asterisks. This is expected as the estimated fluence is based on an optically homogeneous region, and therefore, the accuracy of $\check{\Phi}$ is reduced for increasing μ_a^{tubes}. The δ_f and δ_b of SI errors remain below 7% for this entire range of absorption. The errors of ICA are represented with dots and all three types of errors are >10% for $\mu_a^{\text{tubes}} < 0.15$ mm^{-1} in Case I. The large errors at low absorption may be due to the lower signal-to-noise ratio (SNR), which leads to the first three PCs containing a smaller fraction of the total variance of the data (see Sec. 4.2). When μ_a^{tubes} is increased beyond 0.15 mm^{-1}, the errors of ICA are comparable to SI for the data points with relatively low μ_a^{tubes}, and as μ_a^{tubes} is further increased, ICA results in smaller δ_c than SI. However, when $\mu_a^{\text{tubes}} > 0.55$ mm^{-1}, the performance of ICA abruptly deteriorates as the errors increase to large values beyond the scale of the plots. This is because beyond this threshold, the accuracy of the fluence estimation is sufficiently low to lead to nonlinearities so severe that the independent components found by ICA cannot be correctly identified to the relevant chromophores based on the estimated absorption spectra, hence resulting in a significant increase in errors. This absorption level is comparable to that of blood, which is 0.46 mm^{-1} at 810 nm[38] (assuming a total hemoglobin concentration of 150 gL^{-1} and 100% oxygenation). The blue circles show the errors that would have been obtained if the correct components were manually selected. This manual selection may not be possible in practical applications because prior knowledge of the expected chromophore distribution may not be available.

In Case II, accurate unmixing with all three types of errors $\lesssim 10\%$ was obtained for $\mu_a^{\text{bkg}} < 0.06$ mm^{-1} at 810 nm using ICA and <0.10 mm^{-1} using SI. When μ_a^{bkg} is increased further, the errors for both unmixing methods increase rapidly. However, these thresholds are significantly higher than the absorption coefficient of the common types of biological tissue, which is typically ≈ 0.01 mm^{-1} at 810 nm.[38]

The performance of both ICA and SI is dependent on the accuracy of the fluence adjustment, which in this study is mainly determined by the level of spatial inhomogeneity in the optical

properties of the phantom. This is highlighted by the fact that both methods fail to produce accurate results in the presence of the additional region with increasing concentration of CuCl$_2$ in Case III, where the errors increase with μ_a^{add}.

4.2 Component Selection

The thresholds above which the independent components cannot be identified to the correct chromophores can potentially be shifted toward higher absorption levels if the multiwavelength image data are decomposed into fewer independent components. This can be realized by keeping fewer PCs in the preprocessing stage to reduce the dimensions of the data.

Figure 6 shows the δ_c, δ_f, and δ_b errors from unmixing the images in Case I with ICA using two, three, or four PCs. The results show that if only two PCs were processed with ICA, the unmixed independent components can be correctly identified as CuCl$_2$ or NiCl$_2$ for the full range of absorption levels investigated in Case I, and no abrupt increase in error is observed. However, using only two PCs also results in larger unmixing errors for the higher absorption levels. When four PCs are used, the error trends are similar to using three PCs, but a slight lower shift of the range of the absorption in which ICA results in accurate unmixing is observed. This shift may be explained by the fact that, when the main features in the image have lower contrast, a smaller fraction of the total variance of the data is included in the first three PCs.

4.3 Experimental Images and the Condition Number of the Mixing Matrix

The experimental images were acquired at 18 wavelengths over a spectral range where the absorption spectra of the chromophores have distinct features. However, in *in vivo* imaging applications, some chromophores in the tissue may have flatter and/or less unique absorption spectra. To investigate how the unmixing methods deal with poorly conditioned absorption spectra, ICA and SI were applied on data sets with reduced spectral ranges. The spectral range was reduced in 15 steps by removing the images of the longest wavelengths, such that in the final inversion (the inversion with the fewest wavelengths) only the images at the wavelengths 750, 770, and 790 nm were used. The condition number of the mixing matrix is plotted as a function of the longest wavelength, λ_{max}, used in the inversion in Fig. 7. The figure shows that the condition number increases significantly when the range of wavelengths is limited to $\lambda_{\text{max}} < 830$ nm.
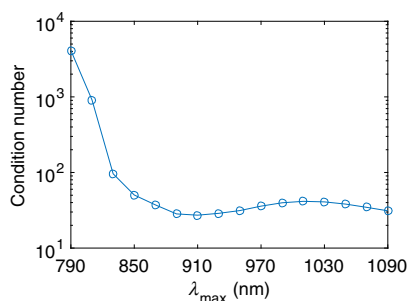
Fig. 7 The condition number of the mixing matrix **A** with the known specific absorption spectra as a function of the longest wavelength, λ_{max}, used in the inversion. The shortest wavelength is fixed at 750 nm.

The unmixed components were normalized and the errors δ_c, δ_f, and δ_b are defined in the same way as for the simulated images. The three types of errors are plotted as a function of λ_{max} used in ICA and SI in Fig. 8. When the full range of wavelengths between 750 and 1090 nm is used, both methods result in accurate unmixing with the δ_c, δ_f, and δ_b errors equal to 11%, 4%, and 3% for ICA and 10%, 5%, and 3% for SI, respectively. This result is in agreement with the simulated data which showed high accuracy for lower absorption levels for both ICA and SI. The high accuracy is maintained for the inversions using $\lambda_{max} > 930$ nm. When λ_{max} is further reduced, the accuracy of SI rapidly deteriorates, while the errors associated with ICA remain relatively low. An example of the unmixed components is shown in Fig. 9, where $\lambda_{max} = 850$ nm. Both the $CuCl_2$ and $NiCl_2$ components are estimated accurately with ICA.

On the other hand, SI is able to unmix the $NiCl_2$ component with relatively low errors, but its estimate of the $CuCl_2$ component is noisy and contains large errors.

Two key factors are likely to affect the unmixing results when the spectral range is reduced: first, due to the spectral variation in the absorption of the contrast agents in the tubes and the background solution, the fluence correction is more accurate for some wavelength combinations than others. This results in variations in the accuracy of SI when certain wavelengths are excluded. Second, the condition number of the known mixing matrix increases by several orders of magnitude when $\lambda_{max} < 830$ nm, as shown in Fig. 7. Therefore, the fact that the inversion is more ill-conditioned is likely to be the dominant cause of the large errors of SI for the inversions where $\lambda_{max} < 830$ nm. However, both these factors affect ICA less significantly. As shown with the simulated image data in Sec. 4.1, ICA is more robust against nonlinearities caused by poor fluence correction, and therefore, the errors of ICA remain low when wavelength range changes. Furthermore, the low errors of ICA when $\lambda_{max} < 830$ nm suggest that unmixing based on statistical independence can potentially be used to obtain accurate separation of the chromophores when SI performs poorly due to ill-conditioning of the mixing matrix. This is possible because ICA does not rely on the known spectra for unmixing and can, therefore, tolerate ill-conditioning better than SI.

5 Summary and Discussion

Accurate unmixing of chromophores from multiwavelength photoacoustic images in the absence of accurate knowledge of the fluence is a challenging task. This study has shown that a simple exponential fluence correction, which is straightforwardly
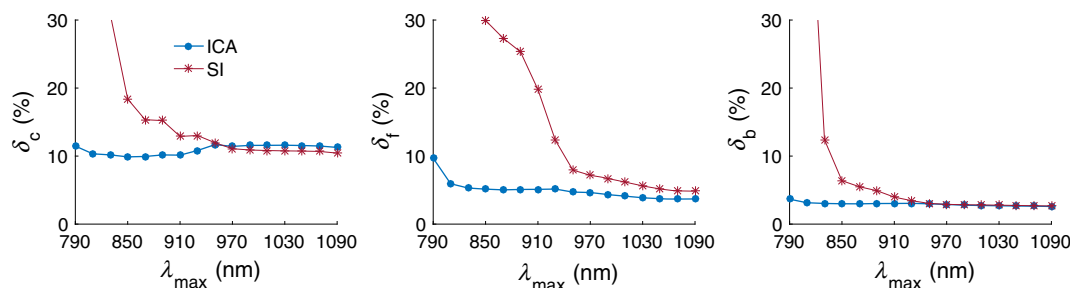
Fig. 8 The δ_c, δ_f, and δ_b errors of the unmixed components using different spectral ranges for ICA and SI. The λ_{max} is shown in the x-axis and the errors are shown in the y-axis. The SI errors increase significantly when $\lambda_{max} < 930$ nm, while the errors of ICA remain relatively low.
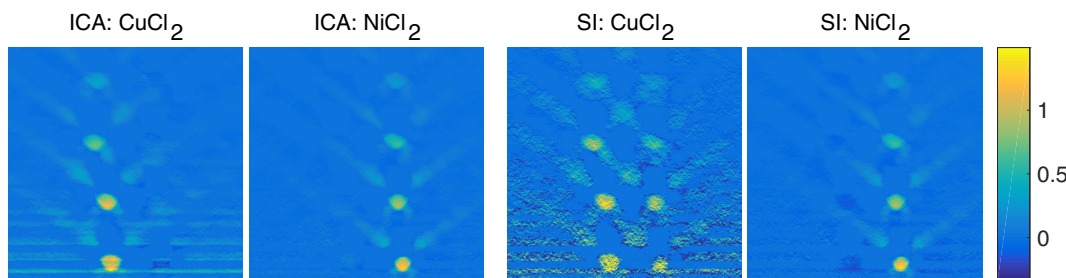
Fig. 9 The estimated components corresponding to $CuCl_2$ and $NiCl_2$ using ICA and SI when the spectral range is between 750 and 850 nm. The components are normalized to the average of the estimated concentration at bottom tube containing the relevant chromophore. The $CuCl_2$ component estimated using SI contains large cross talk errors, while both components are estimated relatively accurately with ICA.

applicable in practice, allows linear models—ICA and SI—to be used to unmix the chromophores to useful degree of accuracy under certain conditions. Experimental and numerical phantoms were designed for which the assumptions underlying both ICA (statistical independence) and SI (full rank molar absorption coefficient matrix) were true. Either approach would, therefore, give accurate results for a perfect fluence correction. When using the approximate, but practical, fluence correction [Eq. (6)], it was shown that ICA results in smaller unmixing errors compared to SI in two circumstances: first, when the absorption level of the contrast agents is ~0.4 to 0.5 mm^{-1}, ICA is more robust against nonlinearities caused by inaccurate fluence estimation than SI, because it allows the mixing matrix to vary in order to produce the most independent components. Second, when the mixing matrix with the known absorption spectra is more ill-conditioned, ICA provides significantly more accurate results as it avoids using the fixed mixing matrix by searching for the chromophore components based on their statistical independence instead.

When the average absorption of the features of interest is higher than a certain threshold (which in this particular phantom is at 0.55 mm^{-1}), leading to greater differences between the estimated and the true fluence, the errors of ICA were shown to abruptly increase as the output components corresponding to the chromophores of interest could not be identified by comparing the estimated spectra with the known. This suggests that, with sufficient SNR, ICA can provide accurate unmixing for absorptions up to a threshold. The threshold will likely be dependent on the absorption spectra, the spatial structure of the chromophores, and the accuracy of the fluence adjustment. Given the potential presence of these upper absorption thresholds under which ICA provides relatively low errors, it would be ideal to use lower concentrations of contrast agents. This would require the imaging system to have high sensitivity to obtain sufficient SNR.

ICA is a suitable unmixing method that is robust to the small errors in the fluence correction that are unavoidable in practical scenarios, provided that the chromophores are known to be mutually statistically independent. This assumption is valid for the chromophores in the phantoms used in this study and in applications such as unmixing some exogenous contrast agents from the tissue in the background. However, the independence criterion is not always fulfilled for all tissue chromophores, hence limiting the range of applications of ICA.

The approximate fluence correction based on spatially homogenous optical properties cannot account for highly absorbing structures in the vicinity of the regions of interest. This is demonstrated in Case III, where the additional absorbing region is shown to cause errors for both ICA and SI. Despite this being an extreme situation where the chromophores of interest are completely surrounded by the additional absorber, it illustrates that the general applicability of linear methods relying on a simple exponential fluence correction is limited in biological tissues with large highly absorbing regions.

The number of components that will be estimated is fixed in SI and determined by the dimensions of the mixing matrix, which is equal to the number of chromophores. This study showed that in ICA, it is not straightforward to choose how many independent components the multiwavelength images should be decomposed into. Since larger dimensionality leads to difficulty in identifying the components as chromophores, one should ideally retain the minimum number of PCs that

contain a sufficient fraction of the total variance to explain the data well. In this study, PCs representing >75% of the variance needed to be kept for further processing with ICA to provide accurate results. However, there are no general guidelines based on theoretical principles for the optimal choice of dimension reduction.

As discussed in Sec. 2, the magnitude of the independent components estimated using ICA is arbitrary because both **W** and **S** are unknown. However, some prior knowledge often exists for the absorption spectra of the chromophores. These known specific absorption spectra can potentially be used to fix the magnitude of the independent components, such that the relative concentration between different chromophores is scaled correctly, and hence reducing ambiguities of ICA. One simple approximate scaling method would be to divide the estimated independent components by a scaling factor equal to the ratio between the mean of each estimated spectrum and the corresponding known spectrum.

6 Conclusion

In conclusion, ICA offers a fast and simple alternative to unmixing multiwavelength photoacoustic images into components representing individual chromophores, provided that the spatial distributions of the chromophore concentrations are statistically independent. When a first-order fluence adjustment has been applied and the absorption is within certain ranges and relatively spatially homogeneous, ICA can provide accurate quantification of the relative chromophore concentrations. The results of ICA depend on the choice of dimensions retained in the preprocessing step, as accurate results require that the components can be identified as the correct chromophores. It was shown that ICA outperforms SI when mixing matrix is ill-conditioned, and that ICA is more robust to errors in the fluence correction compared to SI.

Disclosures

The authors have no relevant financial interests in this article and no potential conflicts of interest to disclose.

References

1. R. A. Kruger et al., "Photoacoustic angiography of the breast," *Med. Phys.* **37**(11), 6096–6100 (2010).
2. S. Manohar et al., "The Twente photoacoustic mammoscope: system overview and performance," *Phys. Med. Biol.* **50**(11), 2543–2557 (2005).
3. T. J. Allen et al., "Spectroscopic photoacoustic imaging of lipid-rich plaques in the human aorta in the 740 to 1400 nm wavelength range," *J. Biomed. Opt.* **17**(6), 061209 (2012).
4. B. Wang et al., "In vivo intravascular ultrasound-guided photoacoustic imaging of lipid in plaques using an animal model of atherosclerosis," *Ultrasound Med. Biol.* **38**(12), 2098–2103 (2012).
5. K. Jansen et al., "Intravascular photoacoustic imaging of human coronary atherosclerosis," *Opt. Lett.* **36**(5), 597–599 (2011).
6. 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).
7. C. P. Favazza, L. A. Cornelius, and L. V. Wang, "In vivo functional photoacoustic microscopy of cutaneous microvasculature in human skin," *J. Biomed. Opt.* **16**(2), 026004 (2011).

8. X. Wang et al., "Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain," *Nat. Biotechnol.* **21**(7), 803–806 (2003).

9. J. Laufer et al., "Three-dimensional noninvasive imaging of the vasculature in the mouse brain using a high resolution photoacoustic scanner," *Appl. Opt.* **48**, D299–D306 (2009).

10. E. W. Stein, K. Maslov, and L. V. Wang, "Noninvasive, in vivo imaging of blood-oxygenation dynamics within the mouse brain using photoacoustic microscopy," *J. Biomed. Opt.* **14**(2), 020502 (2009).

11. G. Ku et al., "Imaging of tumor angiogenesis in rat brains in vivo by photoacoustic tomography," *Appl. Opt.* **44**(5), 770–775 (2005).

12. Y. Lao et al., "Noninvasive photoacoustic imaging of the developing vasculature during early tumor growth," *Phys. Med. Biol.* **53**(15), 4203–4212 (2008).

13. J. Laufer et al., "In vivo preclinical photoacoustic imaging of tumor vasculature development and therapy," *J. Biomed. Opt.* **17**(5), 056016 (2012).

14. A. De La Zerda et al., "Carbon nanotubes as photoacoustic molecular imaging agents in living mice," *Nat. Nanotechnol.* **3**(9), 557–562 (2008).

15. P.-C. Li et al., "In vivo photoacoustic molecular imaging with simultaneous multiple selective targeting using antibody-conjugated gold nanorods," *Opt. Express* **16**, 18605–18615 (2008).

16. L. Li et al., "Photoacoustic imaging of lacZ gene expression in vivo," *J. Biomed. Opt.* **12**(2), 020504 (2007).

17. C. Qin et al., "Tyrosinase as a multifunctional reporter gene for photoacoustic/MRI/PET triple modality molecular imaging," *Sci. Rep.* **3**, 1490 (2013).

18. D. Razansky et al., "Multispectral opto-acoustic tomography of deep-seated fluorescent proteins in vivo," *Nat. Photonics* **3**(7), 412–417 (2009).

19. S. Kim et al., "In vivo three-dimensional spectroscopic photoacoustic imaging for monitoring nanoparticle delivery," *Biomed. Opt. Express* **2**(9), 2540–2550 (2011).

20. R. Hochuli, P. C. Beard, and B. Cox, "Accuracy of approximate inversion schemes in quantitative photoacoustic imaging," *Proc. SPIE* **8943**, 89435V (2014).

21. S. Choi et al., "Wavelength-modulated differential photoacoustic spectroscopy (WM-DPAS) for noninvasive early cancer detection and tissue hypoxia monitoring," *J. Biophotonics* **9**(4), 388–395 (2016).

22. A. Rosenthal, D. Razansky, and V. Ntziachristos, "Quantitative optoacoustic signal extraction using sparse signal representation," *IEEE Trans. Med. Imaging* **28**, 1997–2006 (2009).

23. B. T. Cox, S. R. Arridge, and P. C. Beard, "Estimating chromophore distributions from multiwavelength photoacoustic images," *J. Opt. Soc. Am. A* **26**, 443–455 (2009).

24. G. Bal and K. Ren, "On multi-spectral quantitative photoacoustic tomography in diffusive regime," *Inverse Prob.* **28**(2), 025010 (2012).

25. T. Saratoon et al., "A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation," *Inverse Prob.* **29**(7), 075006 (2013).

26. S. Tzoumas et al., "Eigenspectra optoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," *Nat. Commun.* **7**, 12121 (2016).

27. A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Network* **13**, 411–430 (2000).

28. J. Glatz et al., "Blind source unmixing in multi-spectral optoacoustic tomography," *Opt. Express* **19**(4), 3175–3184 (2011).

29. L. An and B. Cox, "Independent component analysis for unmixing multi-wavelength photoacoustic images," *Proc. SPIE* **9708**, 970851 (2016).

30. X. L. Deán-Ben et al., "Estimation of optoacoustic contrast agent concentration with self-calibration blind logarithmic unmixing," *Phys. Med. Biol.* **59**, 4785–4797 (2014).

31. N. C. Deliolanis et al., "Deep-tissue reporter-gene imaging with fluorescence and optoacoustic tomography: a performance overview," *Mol. Imaging Biol.* **16**(5), 652–660 (2014).

32. A. Taruttis et al., "Multispectral optoacoustic tomography of myocardial infarction," *Photoacoustics* **1**(1), 3–8 (2013).

33. A. Buehler et al., "High resolution tumor targeting in living mice by means of multispectral optoacoustic tomography," *EJNMMI Res.* **2**(1), 14 (2012).

34. D.-K. Yao et al., "Photoacoustic measurement of the Gruneisen parameter of tissue," *J. Biomed. Opt.* **19**(1), 017007 (2014).

35. A. P. Jathoul et al., "Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter," *Nat. Photonics* **9**(4), 239–246 (2015).

36. C. Kim et al., "Sentinel lymph nodes and lymphatic vessels: noninvasive dual-modality in vivo mapping by using indocyanine green in rats—volumetric spectroscopic photoacoustic imaging and planar fluorescence imaging," *Radiology* **255**(2), 442–450 (2010).

37. L. Kou, D. Labrie, and P. Chylek, "Refractive indices of water and ice in the 0.65- to 2.5-μm spectral range," *Appl. Opt.* **32**(19), 3531–3540 (1993).

38. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**, R37–R61 (2013).

39. J. Laufer, E. Zhang, and P. Beard, "Evaluation of absorbing chromophores used in tissue phantoms for quantitative photoacoustic spectroscopy and imaging," *IEEE J. Sel. Top. Quantum Electron.* **16**(3), 600–607 (2010).

40. L. V. Wang and H.-I. Wu, *Biomedical Optics: Principles and Imaging*, John Wiley and Sons, Inc., Hoboken, New Jersey (2009).

41. E. Zhang, J. Laufer, and P. Beard, "Backward-mode multiwavelength photoacoustic scanner using a planar Fabry–Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues," *Appl. Opt.* **47**, 561–577 (2008).

42. M. Schweiger and S. Arridge, "The Toast++ software suite for forward and inverse modeling in optical tomography," *J. Biomed. Opt.* **19**, 040801 (2014).

43. B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**, 021314 (2010).

44. B. E. Treeby, E. Z. Zhang, and B. T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," *Inverse Prob.* **26**(11), 115003 (2010).

45. D. Razansky and V. Ntziachristos, "Hybrid photoacoustic fluorescence molecular tomography using finite-element-based inversion," *Med. Phys.* **34**(11), 4293–4301 (2007).

Biographies for the authors are not available.