

A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization

Dong Han¹, Jie Tian^{1,*}, Shouping Zhu¹, Jinchao Feng², Chenghu Qin¹,
Bo Zhang³, and Xin Yang¹

¹ Medical Image Processing Group, Institute of Automation Chinese Academy of Sciences,
Beijing, 100190, China

² The College of Electronic Information & Control Engineering, Beijing University of Technology China

³ Sino-Dutch Biomedical and Information Engineering School of Northeastern University,
Shenyang, 110004, China

*tian@ieee.org

Abstract: Through the reconstruction of the fluorescent probe distributions, fluorescence molecular tomography (FMT) can three-dimensionally resolve the molecular processes in small animals *in vivo*. In this paper, we propose an FMT reconstruction algorithm based on the iterated shrinkage method. By incorporating a surrogate function, the original optimization problem can be decoupled, which enables us to use the general sparsity regularization. Due to the sparsity characteristic of the fluorescent sources, the performance of this method can be greatly enhanced, which leads to a fast reconstruction algorithm. Numerical simulations and physical experiments were conducted. Compared to Newton method with Tikhonov regularization, the iterated shrinkage based algorithm can obtain more accurate results, even with very limited measurement data.

©2010 Optical Society of America

OCIS codes: (170.6960) Tomography; (170.3010) Image reconstruction Techniques; (170.6280) Spectroscopy, fluorescence and luminescence; (170.3880) Medical and biological imaging.

References and links

1. V. Ntziachristos, J. Ripoll, L. V. Wang, and R. Weissleder, "Looking and listening to light: the evolution of whole-body photonic imaging," *Nat. Biotechnol.* **23**(3), 313–320 (2005).
2. R. Weissleder, and U. Mahmood, "Molecular imaging," *Radiology* **219**(2), 316–333 (2001).
3. J. K. Willmann, N. van Bruggen, L. M. Dinkelborg, and S. S. Gambhir, "Molecular imaging in drug development," *Nat. Rev. Drug Discov.* **7**(7), 591–607 (2008).
4. J. Tian, J. Bai, X. P. Yan, S. Bao, Y. Li, W. Liang, and X. Yang, "Multimodality molecular imaging," *IEEE Eng. Med. Biol. Mag.* **27**(5), 48–57 (2008).
5. C. H. Contag, and M. H. Bachmann, "Advances in *in vivo* bioluminescence imaging of gene expression," *Annu. Rev. Biomed. Eng.* **4**(1), 235–260 (2002).
6. V. Ntziachristos, "Fluorescence molecular imaging," *Annu. Rev. Biomed. Eng.* **8**(1), 1–33 (2006).
7. X. Song, D. Wang, N. Chen, J. Bai, and H. Wang, "Reconstruction for free-space fluorescence tomography using a novel hybrid adaptive finite element algorithm," *Opt. Express* **15**(26), 18300–18317 (2007), <http://www.opticsinfobase.org/abstract.cfm?URI=OE-15-26-18300>.
8. Y. Lu, X. Zhang, A. Douraghy, D. Stout, J. Tian, T. F. Chan, and A. F. Chatzioannou, "Source Reconstruction for Spectrally-resolved Bioluminescence Tomography with Sparse a priori Information," *Opt. Express* **17**(10), 8062–8080 (2009), <http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-17-10-8062>.
9. P. Mohajerani, A. A. Eftekhar, J. Huang, and A. Adibi, "Optimal sparse solution for fluorescent diffuse optical tomography: theory and phantom experimental results," *Appl. Opt.* **46**(10), 1679–1685 (2007).
10. D. Wang, X. Song, and J. Bai, "Adaptive-mesh-based algorithm for fluorescence molecular tomography using an analytical solution," *Opt. Express* **15**(15), 9722–9730 (2007), <http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-15-15-9722>.
11. W. Bangerth, and A. Joshi, "Adaptive finite element methods for the solution of inverse problems in optical tomography," *Inverse Probl.* **24**(3), 034011 (2008).
12. N. Cao, A. Nehorai, and M. Jacobs, "Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm," *Opt. Express* **15**(21), 13695–13708 (2007), <http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-15-21-13695>.

13. I. F. Gorodnitsky, and B. D. Rao, "Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.* **45**(3), 600–616 (1997).
14. E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.* **59**(8), 1207–1223 (2006).
15. I. Daubechies, M. Defrise, and C. DeMol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.* **57**(11), 1413–1457 (2004).
16. F. Gao, H. Zhao, L. Zhang, Y. Tanikawa, A. Marjono, and Y. Yamada, "A self-normalized, full time-resolved method for fluorescence diffuse optical tomography," *Opt. Express* **16**(17), 13104–13121 (2008), <http://www.opticsinfobase.org/oe/abstract.cfm?uri=oe-16-17-13104>.
17. Y. Tan, and H. Jiang, "DOT guided fluorescence molecular tomography of arbitrarily shaped objects," *Med. Phys.* **35**(12), 5703–5707 (2008).
18. A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Adaptive finite element based tomography for fluorescence optical imaging in tissue," *Opt. Express* **12**(22), 5402–5417 (2004), <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-12-22-5402>.
19. J. Feng, K. Jia, G. Yan, S. Zhu, C. Qin, Y. Lv, and J. Tian, "An optimal permissible source region strategy for multispectral bioluminescence tomography," *Opt. Express* **16**(20), 15640–15654 (2008), <http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-16-20-15640>.
20. C. Qin, J. Tian, X. Yang, K. Liu, G. Yan, J. Feng, Y. Lv, and M. Xu, "Galerkin-based meshless methods for photon transport in the biological tissue," *Opt. Express* **16**(25), 20317–20333 (2008), <http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-16-25-20317>.
21. M. Schweiger, S. R. Arridge, M. Hiraoka, and D. T. Delpy, "The finite element method for the propagation of light in scattering media: Boundary and source conditions," *Med. Phys.* **22**(11), 1779–1792 (1995).
22. R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *J. R. Stat. Soc., B* **58**, 267–288 (1996).
23. M. Elad, B. Matalon, and M. Zibulevsky, "Image denoising with shrinkage and redundant representations," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York* **2**, 1924–1931 (2006).
24. H. Minc, *Nonnegative matrices*, (Wiley, New York, 1988).
25. M. T. Chu, and J. L. Watterson, "On a multivariate eigenvalue problem, Part I: Algebraic theory and a power method," *SIAM J. Sci. Comput.* **14**(5), 1089–1106 (1993).
26. A. X. Cong, and G. Wang, "A finite-element-based reconstruction method for 3D fluorescence tomography," *Opt. Express* **13**(24), 9847–9857 (2005), <http://www.opticsinfobase.org/oe/abstract.cfm?URI=oe-13-24-9847>.

1. Introduction

In vivo small animal optical molecular imaging has become an important and rapidly developing method for biomedical research, and has been widely utilized for cancer detection, drug discovery and gene expression visualization, etc [1–4]. When small animals are labeled with optical molecular probes, the corresponding biological information can be indirectly visualized by measuring the emitted light photons of the probes over the surface, thus realizing non-invasive investigation of the small animals [5]. Among optical molecular imaging modalities, fluorescence molecular imaging has become a popular and promising technique. This is partly due to the wealth of the fluorescent probes [6]. However, because the photons in the NIR or visible spectrum are highly scattered by tissue, the traditional planar fluorescence imaging cannot reflect the probe distributions accurately. Fluorescence molecular tomography (FMT), on the contrary, can three-dimensionally resolve the molecular processes by measuring the photons over the animal surface and reconstructing the distribution of the fluorescent probes, which makes it a hot spot in recent years.

FMT is often an ill-posed problem since only the photon distribution over the surface is measurable [7]. This can be alleviated by increasing the measurement data sets. However, even if sufficient measurements can be obtained, the problem may still be ill-conditioned, which means that it is unstable and is sensitive to noises caused by CCD measurement errors and data interpolation errors. To compute a meaningful approximate solution, more *a priori* information should be incorporated to regularize the FMT problem [8]. Among different regularization methods, the Tikhonov regularization is a popular method that has been widely adopted in FMT problems [10,11]. Tikhonov method assumes that the "size" of the solution should not be very large and adds L2-norm constraint of the solution to the original problem, thus making the problem less sensitive to perturbations. The advantage of L2-norm regularization is that the problem is simple and can be solved efficiently by standard minimization algorithms, such as Newton method and conjugate gradient method. However, the solution is often over-smoothed with reduced intensities, and the localized features are lost during the reconstruction process [12]. In recent years, some researchers began to use

sparsity regularization for the optical tomography problems [8,9,12]. For FMT this is based on the fact that, the domains of the fluorescent sources are usually very small and sparse compared with the entire reconstruction domain [9]. This can be considered as valuable *a priori* information for FMT. A straightforward way to integrate the sparsity constraint is to replace the Tikhonov method with L0-norm regularization. However, the optimization problem becomes NP-hard in this case, and cannot be solved efficiently [13]. A tradeoff is to select a proper Lp-norm with $p \in [1,2)$. When p is within this range, large values in the solution are penalized less severely compared with Tikhonov regularization. Therefore, the Lp regularization ($p \in [1,2)$) has a higher tendency to promote the sparsity of the reconstructed result. This effect has been demonstrated in many published articles [8,13]. Another advantage of the sparsity regularization is that, it can still perform well when the measurement data is very limited. This has been well studied in the area of compressed sensing [14]. Besides, the Lp-norm with $p \in [1,2)$ is convex, so the FMT problem remains convex which is important for successful FMT reconstructions. However, this requires the design of new algorithms that can solve the optimization problems with general Lp-norm regularization. Another problem lies in the performance of the reconstruction algorithm. A long reconstruction time may become an obstacle for FMT to be transferred into practical use. Therefore, we believe that designing a fast reconstruction algorithm is always a hot spot for FMT.

In this paper, a fast reconstruction algorithm based on the iterated shrinkage method is proposed [15]. This algorithm decouples the original high dimensional optimization problem and converts it into a set of one dimensional optimization problems, which enables us to handle each one separately. More importantly, general Lp-norm regularization can be incorporated straightforwardly without increasing the complexity of the problem. Next, we provide two reconstruction strategies for different situations together with their complexity analysis. Last, we explain that due to the sparsity of the light sources, the computational burden can be greatly reduced, which leads to a fast reconstruction algorithm.

This paper is organized as follows. The iterated shrinkage based reconstruction algorithm is presented in section 2. In section 3, numerical simulations and physical experiments are conducted to evaluate the performance of the proposed method. Finally, we discuss the results and conclude this paper.

2. Method

2.1 Photon propagation model

For photon propagation in biological tissue within the near infrared spectral window, scattering is the dominant phenomenon over absorption. Therefore, the diffuse approximation to the radiative transfer equation has been extensively applied to model the photon transport [16–20]. For steady-state FMT with point excitation sources, the following coupled diffuse equations can be utilized to depict the forward problem:

$$\begin{cases} \nabla \cdot (D_x(r) \nabla \Phi_x(r)) - \mu_{ax}(r) \Phi_x(r) = -\Theta \delta(r - r_i) \\ \nabla \cdot (D_m(r) \nabla \Phi_m(r)) - \mu_{am}(r) \Phi_m(r) = -\Phi_x(r) \eta \mu_{af}(r) \end{cases} \quad (r \in \Omega) \quad (1)$$

where subscripts x and m denote the excitation and emission wavelengths, respectively; Ω denotes the domain of the problem; $\Phi_{x,m}$ is the photon flux density; $\mu_{ax,am}$ is the absorption coefficient and $D_{x,m} = 1/3(\mu_{ax,am} + (1-g)\mu_{sx,sm})$ is the diffusion coefficient, $\mu_{sx,sm}$ is the scattering coefficient, and g is the anisotropy parameter; $\eta \mu_{af}$ denotes the fluorescent yield which is to be reconstructed. Here, we assume that the absorption and scattering of the excitation light caused by the fluorescent probes will have little effect on the distribution of Φ_x , as the fluorescent probes often occupy a very small volume compared with the imaging

domain Ω . In this forward model, the excitation light is implemented as isotropic point sources located one mean free path of photon transport beneath the surface at different locations $r_l (l=1,2,\dots,L)$. Θ denotes the amplitude of the point sources. The coupled equations are complemented by Robin-type boundary conditions which depict the refractive index mismatch between the external medium and Ω :

$$\Phi_{x,m}(r) + 2AD_{x,m}(r)(\vec{v}(r)\cdot\Phi_{x,m}(r)) \quad (r \in \partial\Omega) \quad (2)$$

where \vec{v} denotes the outward normal to the surface. A is a constant depending on the optical reflective index mismatch at the boundary [21].

2.2 Linear relationship establishment

Instead of solving Eqs. (1) and (2) directly, they are posed in the weak solution forms. Discretizing the domain with small elements and employing the base functions as the test functions, the FMT problem can be linearized and the following matrix-form equations can be obtained.

$$[K_x]\{\Phi_x\} = \{S_x\} \quad (3)$$

$$[K_m]\{\Phi_m\} = [F]\{X\} \quad (4)$$

where $K_{x,m}$ is the system matrix. Matrix F is obtained by discretizing the unknown fluorescent yield distribution. Vector X denotes the fluorescent yield to be reconstructed. For each excitation point source at $r_l (l=1,2,\dots,L)$, Φ_x can be directly obtained by solving Eq. (3). Considering the *inverse crime* problem, Φ_x is calculated on a fine mesh using 2nd order Lagrange elements. Then it is projected onto a coarse mesh which will be used for the reconstruction of X with linear elements. As matrix K_m is symmetrical positive definite, Eq. (4) can be transformed into:

$$\{\Phi_{m,l}\} = [K_{m,l}^{-1}][F]\{X\} = [B_l]\{X\} \quad (5)$$

Removing the unmeasurable entries in $\Phi_{m,l}$ and corresponding rows in B_l , we have:

$$\{\Phi_{m,l}^{meas}\} = [A_l]\{X\} \quad (6)$$

Next, we assemble Eq. (6) for different excitation locations and obtain the following matrix-form equation:

$$\{\Phi\} = [A]\{X\} \quad (7)$$

The detailed descriptions can be found in [7].

2.3 Reconstruction based on iterated shrinkage method

As is mentioned, FMT is usually an ill-posed problem, which means that the dimension of the null space of A is not zero. Therefore, the solution is not unique in this case. Even if the FMT problem becomes less ill-posed when more fluorescence measurement data can be captured by the CCD camera, it can still remain ill-conditioned (with a large condition number). Therefore, errors in the FMT problem can be magnified, which will affect the reconstruction results. Errors are inevitable and can be introduced in several ways, e.g. the fluorescence measurement errors caused by CCD camera noise and the approximation errors caused by data interpolation. A standard way to regularize the problem is to incorporate additional constraints on the solution, which can be considered as a kind of *a priori* information:

$$\min_{X \geq 0} E(X) = \frac{1}{2} \|AX - \Phi\|_2^2 + \lambda R(X) \quad (8)$$

where $R(X)$ is the penalty function, and λ is a positive real number called the regularization parameter which is used to balance the two terms. In this paper, X is considered to be non-negative. When L2-norm penalty function is used, this becomes the popular Tikhonov regularization method. Here, we only consider the case when $R(X)$ is $\|X\|_p^p$ with $p \in [1, 2)$. In this case, the original energy function for the FMT problem with general Lp-norm regularization can be represented as follows:

$$E(X) = \frac{1}{2} \|AX - \Phi\|_2^2 + \lambda \|X\|_p^p = \frac{1}{2} \|AX - \Phi\|_2^2 + \lambda \sum_{i=1}^N x_i^p \quad (9)$$

where N is the dimension of X . Here, we define $X^{p-1} = [x_1^{p-1}, x_2^{p-1}, \dots, x_N^{p-1}]^T$. To minimize the energy function, we calculate the partial derivative of $E(X)$, and set the result to zero:

$$A^T AX - A^T \Phi + \lambda p X^{p-1} = 0 \quad (10)$$

When $p=1$ and $\exists x_i = 0$, Eq. (10) is not correct. This special case will be considered later. Equation (10) is non-linear except when $p=1$. If $A^T A$ is a diagonal matrix, then solving X is equivalent to solving each x_i individually, which is quite simple. However, this is generally not the case for FMT, and all x_i are closely coupled. Therefore, solving Eq. (10) is not trivial.

Next, we introduce a surrogate function with the following form [15]:

$$S(X; X_0) = \frac{c}{2} \|X - X_0\|_2^2 - \frac{1}{2} \|AX - AX_0\|_2^2 \quad (11)$$

where c is a constant scalar and X_0 is a constant vector. The parameter c should be chosen so that $S(X; X_0)$ is strictly convex. This implies that the Hessian matrix of $S(X; X_0)$ should be positive definite, $cI - A^T A \succ 0$. This can be satisfied by choosing parameter $c > \|A^T A\|_2 = \rho(A^T A)$. Adding $E(X)$ and $S(X; X_0)$ together, we get the new energy function as follows:

$$J(X; X_0) = \frac{1}{2} \|AX - \Phi\|_2^2 + \lambda \|X\|_p^p + \frac{c}{2} \|X - X_0\|_2^2 - \frac{1}{2} \|AX - AX_0\|_2^2 \quad (12)$$

Calculating the partial derivative of Eq. (12) with respect to X and setting the result to zero, we have:

$$\nabla J(X; X_0) = -[A^T (\Phi - AX_0) + cX_0] + \lambda p X^{p-1} + cX = 0 \quad (13)$$

Let $D = c^{-1} A^T (\Phi - AX_0) + X_0$ represent all the constant items, Eq. (13) can be further simplified as follows:

$$X - D + \frac{\lambda p}{c} X^{p-1} = 0 \quad (14)$$

In Eq. (14), all x_i are decoupled now, and solving Eq. (14) is equivalent to solving each x_i individually, which makes the use of general Lp-norm regularization possible. This is a distinctive advantage of the iterated shrinkage method compared with other methods, e.g. the

Lasso estimator that is based on L1 regularization [22]. Here, we take $p = 1$ for example. To solve a certain x_i , we firstly assumes that $x_i > 0$. Then, if $d_i > \lambda/c$, we have $x_i = d_i - \lambda/c$. If d_i does not satisfy the condition, we don't calculate x_i using Eq. (14) but simply set $x_i = 0$. Therefore, we can avoid the aforementioned special case. Now, we define a function named *Shrink*:

$$Shrink_{\lambda,c}(z; p = 1) = \begin{cases} z - \frac{\lambda}{c} & z > \frac{\lambda}{c} \\ 0 & z \leq \frac{\lambda}{c} \end{cases} \quad (15)$$

This function maps the values smaller than the threshold to zero; values larger than the threshold are “shrunk”, thus the name of the function. An important feature of the *Shrink* function (not limited to $p = 1$) is that it can be determined once and off-line. Using this *Shrink* function, the optimal solution to the optimization problem can be represented as:

$$X_{opt} = \arg \min_X J(X; X_0) = \begin{bmatrix} Shrink_{\lambda,c}(d_1; p = 1) \\ Shrink_{\lambda,c}(d_2; p = 1) \\ \vdots \\ Shrink_{\lambda,c}(d_N; p = 1) \end{bmatrix} \quad (16)$$

Next, we analyze the influence of the surrogate function on the optimization problem. Since $S(X; X_0)$ is strictly convex, it has a unique global minimum. Zeroing the derivative of $S(X; X_0)$ with respect to X leads to the equation:

$$A^T A(X - X_0) = c(X - X_0) \quad (17)$$

Since $X = X_0$ is obviously a solution to Eq. (17), it is the unique solution. Therefore, the global minimum of $S(X; X_0)$ is zero when $X = X_0$. From the above analysis, we know that the optimal solution to this optimization problem will have a bias towards X_0 . X_0 can be regarded as the *a priori* knowledge of X . However, this *a priori* information may not be obtainable in most cases. To resolve this problem, an iterative reconstruction scheme is adopted instead. For a new iteration $k + 1$, X_{k+1} is calculated by replacing X_0 with X_k :

$$x_{k+1,i} = Shrink_{\lambda,c}(d_{k,i}(X_k); p = 1) \quad (i = 1, 2, \dots, N) \quad (18)$$

Now, D becomes a function of X_k and can be updated iteratively. Here, we should point out that the convergence analysis in [15] cannot be directly used to analyze the proposed method. A major reason is that, the base functions in the finite element framework are not orthogonal, though they are linearly independent. However, it is shown in reference [23] that even with unorthogonal bases, the iterated shrinkage method is still practiced successfully. We share similar experience in the FMT reconstructions. This can be seen from the reconstruction results in section 3. Here, we provide a brief explanation to show that the proposed method still makes sense in this case. In fact, $J(X; X_k)$ is an upper bound of $E(X)$, which means $J(X; X_k) \geq E(X)$. Equality can be satisfied if and only if $X = X_k$. Then, for a certain iteration k , the following inequality exists:

$$E(X_{k+1}) \leq J(X_{k+1}; X_k) \leq J(X_k; X_k) = E(X_k) \quad (19)$$

Therefore, $E(X_k)$ is non-increasing as k grows.

Now, we consider the calculation of c satisfying $c > \rho(A^T A)$. A straightforward way is to calculate all the Eigen values of $A^T A$, and c can be set to the largest Eigen value plus ε , where ε is a small positive value. However, this could be rather time-consuming. Therefore, we provide two alternative ways to determine c . For the first one, we estimate the upper bound of the maximum Eigen value instead of actually calculating it, and c can be set to the upper bound plus ε . Several estimation algorithms can be adopted, e.g. the popular Minc method [24]. The second way is to calculate the maximum Eigen value directly using e.g. the Power method [25], and c can be set similarly. Sometimes, the estimated upper bound of maximum Eigen value may be too large compared with the true value. This will increase the influence of the surrogate function on the optimization problem and will slow down the convergence rate. Therefore, for the above two ways, we prefer the latter one, because it is more accurate.

Next, we provide two reconstruction strategies for different situations. Suppose we have obtained A and Φ . For a certain iteration k , D_k can be represented in the following two ways:

$$D_k = \frac{1}{c} A^T (\Phi - AX_k) + X_k \quad (20)$$

$$D_k = \frac{1}{c} A^T \Phi - \frac{1}{c} A^T AX_k + X_k \quad (21)$$

Equations (20) and (21) lead to the following two reconstruction strategies:

Reconstruction strategy 1

Require: $A_{m \times n}$, Φ and λ

- 1: Calculate the maximum Eigen value eig_{\max} of matrix $[0, A; A^T, 0]$
 - 2: $c = \widehat{eig}_{\max}^2 + \varepsilon$
 - 3: **for** $k = 0$ to k_{\max} **do**
 - 4: $V_k = \Phi - AX_k$
 - 5: $D_k = c^{-1} A^T V_k + X_k$
 - 6: **for** $i = 1$ to N **do**
 - 7: $x_{k+1,i} = Shrink(d_{k,i})$
 - 8: **end for**
 - 9: **end for**
-

Reconstruction strategy 2

Require: $A_{m \times n}$, Φ and λ

- 1: $M = A^T A$
 - 2: $B = A^T \Phi$
 - 3: Calculate the maximum Eigen value eig_{\max} of M
 - 4: $c = \widehat{eig}_{\max} + \varepsilon$
 - 5: **for** $k = 0$ to k_{\max} **do**
 - 6: $D_k = c^{-1} B - c^{-1} M X_k + X_k$
 - 7: **for** $i = 1$ to N **do**
 - 8: $x_{k+1,i} = Shrink(d_{k,i})$
 - 9: **end for**
 - 10: **end for**
-

For reconstruction strategy 1, we don't compute $A^T A$ since this is time-consuming. As we know, $\rho(A^T A)$ is equal to the square of the maximum singular value s_{\max} of A , and s_{\max} is equal to the maximum Eigen value eig_{\max} of the assembled matrix $[0, A; A^T, 0]$. So we can calculate eig_{\max} instead. For each iteration, two matrix-vector multiplications are needed. If we ignore the computations for the vector-vector operations (including the *Shrink* operations) and the calculation of eig_{\max} , which are relatively cheap in the total reconstruction process, the computational complexity of strategy 1 can be represented as $O(k_{\max} * 2mn)$.

For reconstruction strategy 2, we pre-calculate both $A^T A$ and $A^T \Phi$. The computational complexity of calculating $A^T A$ is $O(m * n^2)$. For each iteration, only one matrix-vector multiplication is needed. Then, the computational complexity of strategy 2 can be represented as $O((k_{\max} + m) * n^2)$.

From the above analysis, it is clear that when k_{\max} is not very large, strategy 1 is more efficient. For a large k_{\max} , especially when k_{\max} is much larger than the dimension of the problem, strategy 2 becomes a better choice.

When k_{\max} is very large, even strategy 2 will be quite time-consuming. Fortunately, if the sparsity characteristic of X is considered, the performance of both strategies can be greatly improved. Suppose X_k is sparse with n_k non-zero elements, and n_k is usually far less than n . When multiplying X_k by a matrix, the multiplications involving the zero elements in X_k can be simply ignored. Taking strategy 2 for example, the computational complexity of calculating MX_k is $O(n^2)$. If the sparsity of X_k is considered, only $n * n_k$ multiplications are needed, which is a small fraction of the original. For strategy 1, the computational complexity of calculating V_k can be reduced to $O(mn_k)$. However, V_k is not sparse any more. Therefore, computing D_k still needs $m * n$ multiplications.

3. Results

3.1 Simulation verifications

In this subsection, heterogeneous simulation experiments were conducted to demonstrate the performance of the proposed method. Figure 1(a) shows the heterogeneous cylindrical phantom we used, which was of 20mm in diameter and 20mm in height. Figure 1(b) is a slice image of the phantom in $z = 0$ plane. The phantom consisted of four kinds of materials to represent muscle (M), lung (L), heart (H) and bone (B), respectively. The optical parameters can be found in Table 1. The black dots in Fig. 1(b) represent the excitation light sources, which were modeled as isotropic point sources located one mean free path of photon transport beneath the surface in $z = 0$ plane.

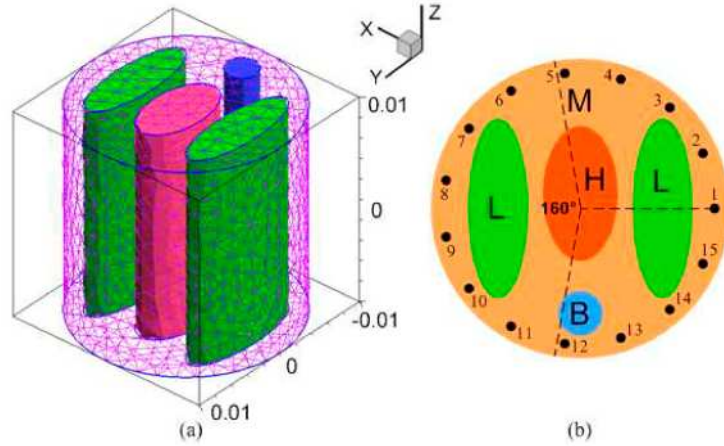


Fig. 1. Mouse-mimicking heterogeneous phantom with four kinds of materials to represent muscle (M), lung (L), heart (H) and bone (B), respectively. (a) is the 3D view of the phantom. (b) is the slice image of the phantom in $z = 0$ plane. The black dots in (b) represent the excitation point source locations. For each excitation location, fluorescence is measured from the opposite cylindrical side with 160° field of view.

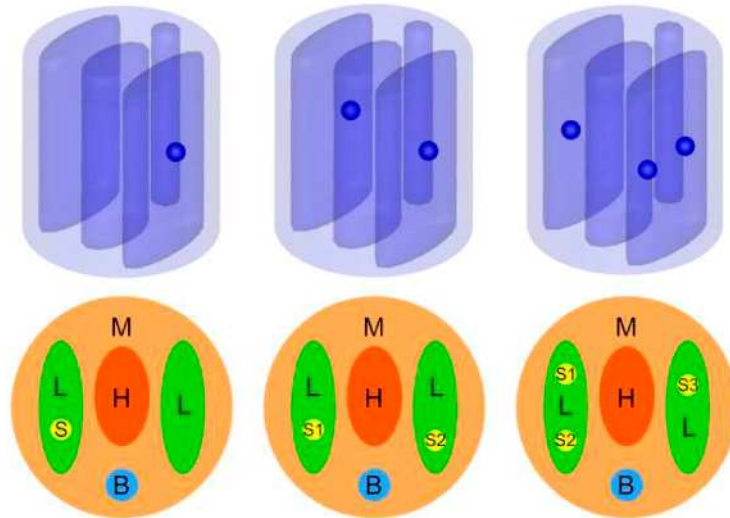


Fig. 2. Three different phantom setups for single fluorescent source (left column), double fluorescent sources (middle column) and three fluorescent sources (right column), respectively. The first row is the 3D views of the phantoms, and the second row is the slice images in $z = 0$ plane. All the fluorescent sources are spherical and are centered in $z = 0$ plane. The diameters of these spherical fluorescent sources are all set to be 2mm.

Table 1. Optical parameters of the heterogeneous phantom [26]

Material	$\mu_{ax}(mm^{-1})$	$\mu'_{sx}(mm^{-1})$	$\mu_{am}(mm^{-1})$	$\mu'_{sm}(mm^{-1})$
Muscle	0.0052	1.08	0.0068	1.03
Lung	0.0133	1.97	0.0203	1.95
Heart	0.0083	1.01	0.0104	0.99
Bone	0.0024	1.75	0.0035	1.61

As is mentioned, for FMT, the domains of the fluorescent sources are usually very small and sparse. Therefore, we used small spheres to represent the fluorescent sources in our experiments. Figure 2. shows the three different phantom setups for single source (left column), double sources (middle column) and three sources (right column), respectively. The first row of Fig. 2 is the 3D views of the phantom setups and the second row is the slice images in $z=0$ plane. All the fluorescent sources were of 2mm in diameter and were centered in $z=0$ plane. The fluorescent yields of these sources were set to be 8. These three phantom setups were discretized with tetrahedrons. For the forward problem, three fine meshes, each with about 23000 degrees of freedom (DOFs), were used for different phantom setups. Then the forward solutions were projected onto a single coarse mesh with 2710 DOFs, which was used for the inverse problem. Fluorescence measurement was implemented in transillumination mode. For each excitation source, measurement was taken from the opposite cylindrical side with 160° field of view (FOV), which is illustrated in Fig. 1(b). This means that all the nodes on the cylindrical surface within this FOV were considered to be measurable. When practical fluorescence measurements are taken using a CCD camera, the shot noise always exists, which obeys the Poisson distribution. If large numbers of photons are collected, the shot noise will approach a Gaussian distribution. Therefore, for FMT reconstructions with simulated data, Gaussian noise is often used to simulate the real case. In our simulation experiments, we added 5% Gaussian noise to the measurement data.

In this paper, L1-norm regularization was utilized. The maximum iteration number k_{\max} was set to be 30000, which was sufficiently large for these experiments. Since k_{\max} was larger than the dimension of the inverse problem, reconstruction strategy 2 was adopted. To better illustrate the performance of the proposed method, we compared it to the bound-constraint Newton method that minimized the original energy function $E(X)$ with Tikhonov regularization. Since the Newton method is a second-order optimization method, it generally converges faster than those first-order ones. Here, we set the maximum iteration number for the Newton method to be 300. Regularization parameter generally plays an important role in the FMT reconstructions. In this paper, the regularization parameters for the Newton method and the proposed method were manually optimized. Finding the optimal or near-optimal regularization parameters automatically will be our future work. Both reconstruction algorithms were implemented in C++, and all the reconstructions were carried out on a personal computer with 2.33GHz Intel Core2 duo CPU and 2GB RAM.

In the first set of experiments, fluorescence was excited by point sources from 15 different locations in sequence, which is illustrated in Fig. 1(b). Measurements were taken every 24° and a total of 15 data sets were acquired for the reconstruction of the fluorescent yield. Figures 3, 4 and 5 show the reconstruction results for three different phantom setups, which are presented in the form of slice images in $z=0$ plane and iso-surfaces for 30% of the maximum value. The small circles in the slice images denote the real positions of the fluorescent sources. From Figs. 3, 4 and 5 we can clearly see that, when only one fluorescent source exists, the results from both methods are quite satisfactory. However, when multiple sources exist, the over-smooth effect of Tikhonov regularization begins to appear. On the contrary, the proposed method can still preserve the sparsity of the sources very well in this case, and the reconstructed intensities are greater. To analyze these results quantitatively, we define the location error to be $\|L_r - L_0\|_2$, where L_0 is the real location of the source center and L_r is the location of the node with the maximum reconstructed value for that source. We also define the relative intensity error to be $|I_r - I_0|/I_0$, where I_0 is the true intensity of the source and I_r is the maximum reconstructed intensity. The quantitative comparisons between these results are presented in Table 2. From Table 2 we can see that, both reconstruction methods can obtain satisfactory source localizations. But the relative intensity errors for the proposed method are smaller.

Next, we compare the running time of the two methods. Since both methods need the pre-calculation of $A^T A$ and $A^T \Phi$, we set the starting point at the time when $A^T A$ and $A^T \Phi$ were just obtained. Then, the running time of the Newton method for three different phantom setups was 166.62s, 189.32s and 180.09s, respectively. And the running time of the proposed method was 1.16s, 2.11s and 4.46s, respectively. The proposed method is much more efficient.

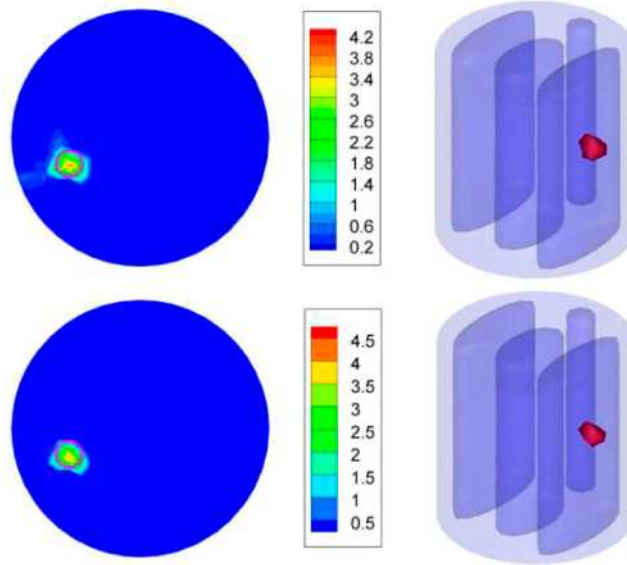


Fig. 3. Reconstruction results from the Newton method (first row) and the iterated shrinkage based method (second row) for 1 spherical fluorescent source and 15 measurement data sets. These results are presented in the form of slice images in $z = 0$ plane (left column) and isosurfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

Table 2. Quantitative comparisons between the results from the Newton method with Tikhonov regularization (Newton-L2) and the iterated shrinkage based method with L1 regularization (IS-L1) for 15 measurement sets

Phantom setup	Source No.	Location error (mm) (Newton-L2)	Location error (mm) (IS-L1)	Relative intensity error (Newton-L2)	Relative intensity error (IS-L1)
1 source	S	0.39	0.39	44.5%	38.2%
2 sources	S1	0.39	0.39	70.0%	45.3%
	S2	0.85	0.85	75.9%	66.4%
3 sources	S1	0.25	0.25	74.4%	33.0%
	S2	0.97	0.97	75.3%	63.4%
	S3	0.73	0.73	71.8%	63.8%

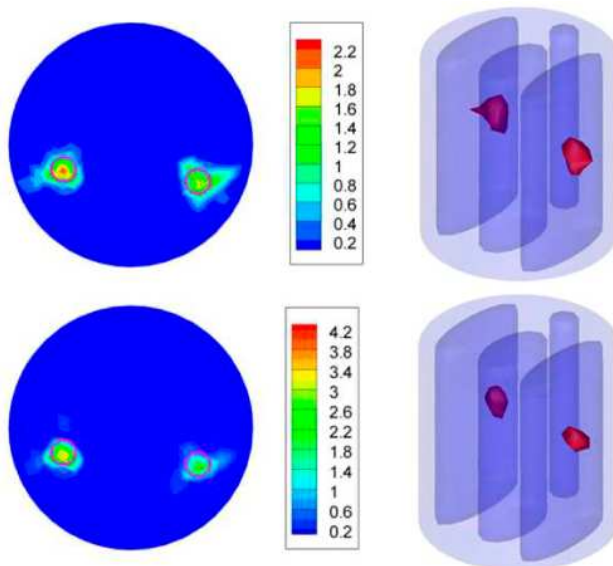


Fig. 4. Reconstruction results from the Newton method (first row) and the iterated shrinkage based method (second row) for 2 spherical fluorescent sources and 15 measurement data sets. These results are presented in the form of slice images in $z = 0$ plane (left column) and iso-surfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

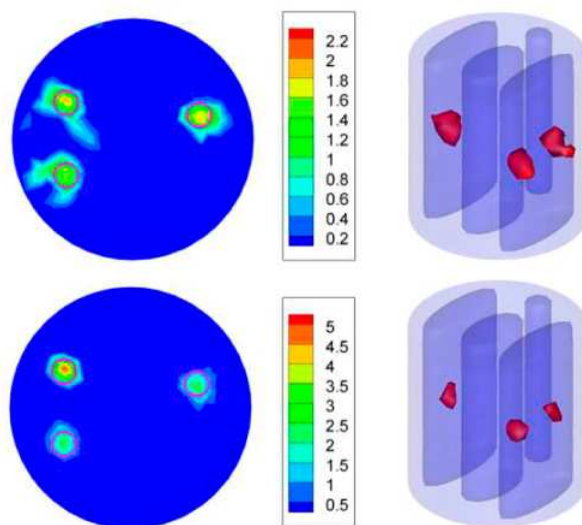


Fig. 5. Reconstruction results from the Newton method (first row) and the iterated shrinkage based method (second row) for 3 spherical fluorescent sources and 15 measurement data sets. These results are presented in the form of slice images in $z = 0$ plane (left column) and iso-surfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

In the second set of experiments, we reduced the amount of measurement data to simulate a much worse case. This is possible when long-time measurement is not appropriate or feasible. For instance, when imaging small animals like mice, the artifacts caused by movements must be taken into consideration. Besides, long-time measurement can cause the bleaching effect of the fluorescent probe, which may influence the reconstructed results. One way to resolve these problems is to reduce the number of fluorescence measurements. This

requires that we should be able to reconstruct the fluorescent sources from very limited data. It has been shown for bioluminescence tomography that, by using sparsity constraint, satisfactory results can still be achievable even with very limited imaging data [8]. Here, we only retained the measurement data sets generated by excitation point source 1, 6 and 11. Figures 6, 7 and 8 show the reconstruction results in this case. From these figures we can see that, when the measurement data is very limited and multiple fluorescent sources exist, the proposed method can obtain much better results compared to the Newton method with Tikhonov regularization. This demonstrates the applicability of the proposed method under more ill-posed conditions. Quantitative comparisons are presented in Table 3. For the Newton method, the reconstruction errors for source S1 and S2 in the third phantom setup are not presented, because the two sources cannot be separated in the result. The running time of the Newton method for three different phantom setups was 183.98s, 173.91s and 168.02s, respectively. And the running time of the proposed method was 1.84s, 2.97s and 2.42s, respectively.

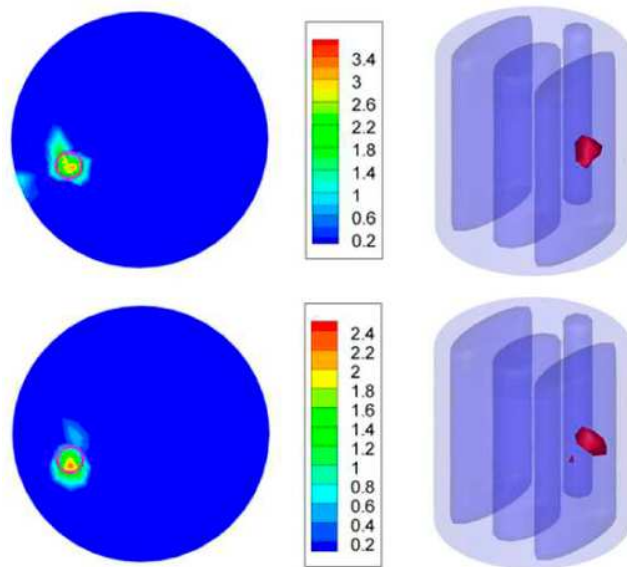


Fig. 6. Reconstruction results from the Newton method (first row) and the iterated shrinkage based method (second row) for 1 spherical fluorescent source and 3 measurement data sets. These results are presented in the form of slice images in $z = 0$ plane (left column) and isosurfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

Table 3. Quantitative comparisons between the results from the Newton method with Tikhonov regularization (Newton-L2) and the iterated shrinkage based method with L1 regularization (IS-L1) for 3 measurement sets

Phantom setup	Source No.	Location error (mm) (Newton-L2)	Location error (mm) (IS-L1)	Relative intensity error (Newton-L2)	Relative intensity error (IS-L1)
1 source	S	0.39	0.39	53.5%	68.9%
2 sources	S1	2.00	1.62	91.4%	71.4%
	S2	0.85	0.85	84.4%	64.0%
3 sources	S1	—	0.25	—	29.3%
	S2	—	0.97	—	56.3%
	S3	0.73	0.73	87.0%	60.2%

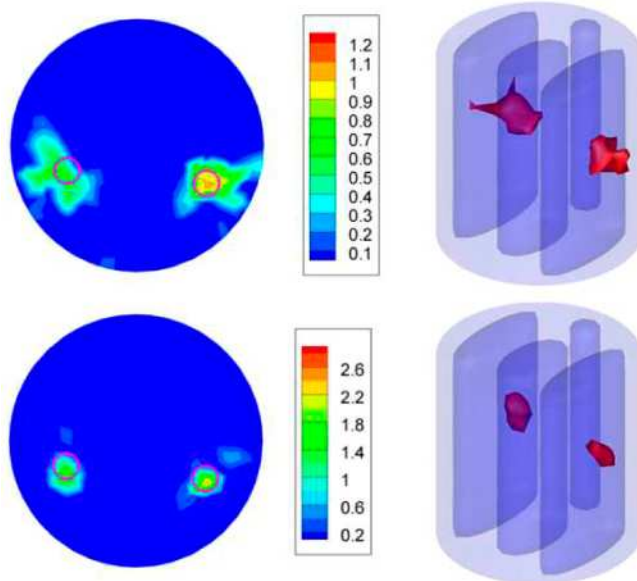


Fig. 7. Reconstruction results from the Newton method (first row) and the iterated shrinkage based method (second row) for 2 spherical fluorescent sources and 3 measurement data sets. These results are presented in the form of slice images in $z = 0$ plane (left column) and iso-surfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

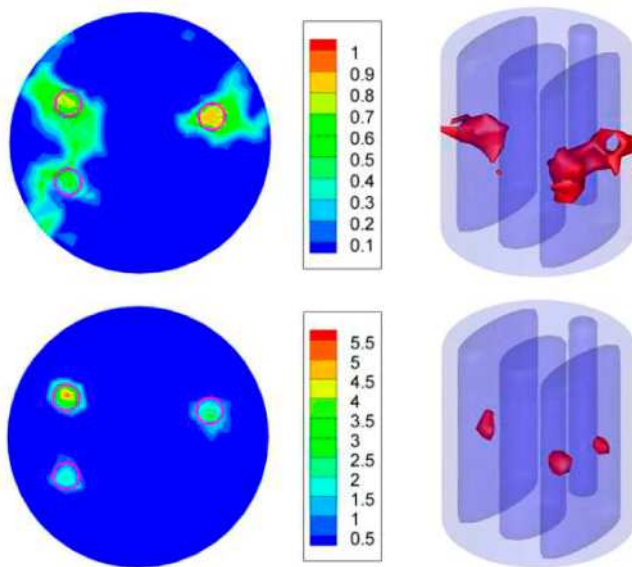


Fig. 8. Reconstruction results from the Newton method (first row) and the iterated shrinkage based method (second row) for 3 spherical fluorescent sources and 3 measurement data sets. These results are presented in the form of slice images in $z = 0$ plane (left column) and iso-surfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

3.2 Physical experiments

In this subsection, physical experiments were conducted to further test the proposed algorithm. Figure 9 illustrates the experimental setup. Excitation illumination was provided by a 671nm CW laser and the power was set to 20mW. The spot diameter of the laser beam was approximately 1mm. The fluorescence measurements were implemented in transillumination mode. A 10nm band-pass filter centered at 700nm was used to allow light transmission at the emission wave-length. The optical density of the filter at the excitation wave-length was larger than 5. Fluorescence was captured by a CCD camera (Princeton Instruments VersArray 1300B, Roper Scientific, Trenton, NJ) which was cooled to -110°C . To reconstruct the sources, the pixels of the measured fluorescence image should firstly be converted into the corresponding photon flux densities. Then, the calibrated image needs to be back-projected onto the surface of the phantom. Errors can be introduced during the mapping procedure. To avoid this problem and to test only the performance of the proposed reconstruction method, a cubic phantom was utilized in this experiment, and the back-projection was reduced to a point-to-point mapping. Figure 10(a) shows the phantom we used. The side length is 20mm. The phantom is made from polyoxymethylene, and the optical parameters for both excitation and emission wave-lengths, which are presented in Table 4, were determined by diffuse optical tomography. The sparsity-promoting characteristic of the proposed method can be better demonstrated when more than one source exists, which can be seen from our simulation results. Therefore, in this experiment, two small holes of 1.25mm in radius were drilled to allow two fluorescent sources. 2000nM Cy5.5 solution was used as the fluorescent source. The height of the two cylindrical sources was 2mm, and the centers were at $(-3.75\text{mm}, 3.75\text{mm}, 1\text{mm})$ and $(3.75\text{mm}, -3.75\text{mm}, 1\text{mm})$ respectively, which is illustrated in Fig. 10(b). To simulate a badly ill-posed situation, fluorescence was excited by point sources from only 4 different locations in $z = 0$ plane, which is illustrated in Fig. 10(c). The phantom was placed on a rotational stage, which was controlled by computer, to allow measurements from four sides.

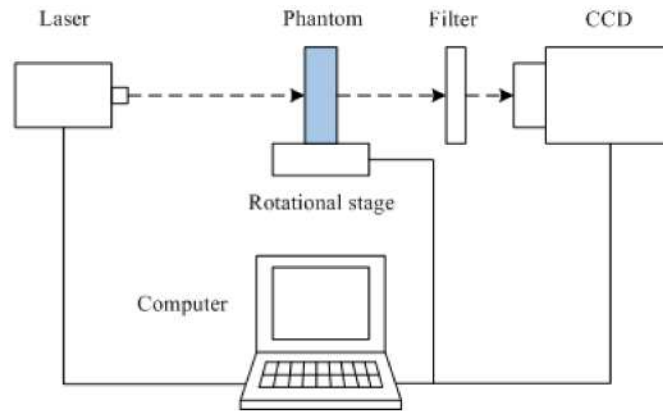


Fig. 9. The sketch of the experimental setup.

Table 4. Optical parameters of the cubic phantom

Wave-length (nm)	μ_a (mm^{-1})	μ_s' (mm^{-1})
671	0.00029	1.08
700	0.00051	1.11

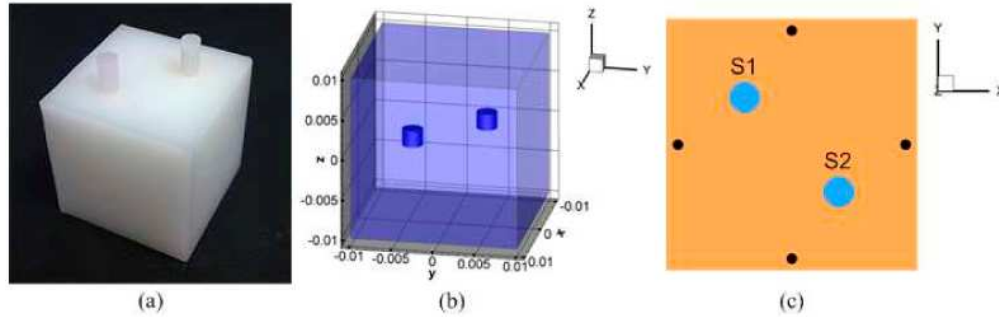


Fig. 10. The homogeneous cubic phantom with 2 cylindrical fluorescent sources. (a) is the photograph of the phantom. (b) is the 3D view of the phantom and the sources. (c) is the slice image of the phantom in $z = 0$ plane. The black dots in (c) represent the excitation point source locations.

In this experiment, the cubic phantom was discretized with tetrahedrons. For the forward problem, a fine mesh with about 21000 DOFs was adopted. And a coarse mesh with 2705 DOFs was used for the inverse problem. Figure 11 shows the reconstruction results which are presented in the form of slice images in $z = 1\text{mm}$ plane and iso-surfaces for 30% of the maximum value. From Fig. 11 we can clearly see that, due to the badly ill-posed situation and the over-smooth effect of Tikhonov regularization, the reconstructed fluorescent sources from the Newton method are spread within the entire domain, which makes result totally meaningless. The location errors are not presented for the Newton method, because the source locations cannot be identified from the result. On the contrary, the proposed method with L1 regularization can preserve the sparsity of the fluorescent sources very well, though some artifacts exist near the center. The location errors for the fluorescent sources S1 and S2 were 1.21mm and 0.82mm, respectively. The running time of the Newton method and the proposed method was 135.91 seconds and 2.56 seconds, respectively.

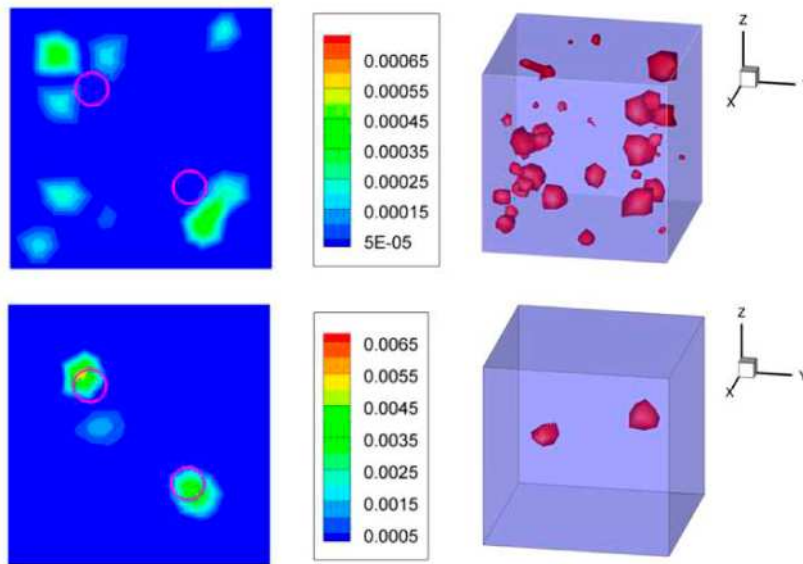


Fig. 11. Reconstruction results of the cubic phantom from the Newton method (first row) and the iterated shrinkage based method (second row) using 4 measurement data sets. These results are presented in the form of slice images in $z = 1\text{mm}$ plane (left column) and iso-surfaces for 30% of the maximum value (right column). The small circles in the slice images denote the real positions of the fluorescent sources.

4. Conclusion

In this paper, an iterated shrinkage based reconstruction algorithm is proposed. By integrating an additional surrogate function, the original high dimensional optimization problem can be decoupled into a set of one dimensional ones that can be solved easily. This also enables us to incorporate sparsity regularization in a graceful way. Two reconstruction strategies are provided for different situations together with their complexity analysis. Besides, we explain that due to the sparsity characteristic of the fluorescent sources, the efficiency of the algorithm can be greatly improved, which leads to a fast reconstruction method. Simulation verifications show that the proposed method outperforms the traditional bound-constrained Newton method with Tikhonov regularization in two ways. First, due to the sparsity constraint of our method, obvious improvements can be seen from these reconstruction results. Second, our method is much faster than the Newton method when the fluorescent sources are sparse. This is important if we want to transfer FMT into practical use. We also test our method under a more ill-posed condition and satisfactory results can still be achievable. Reconstruction of experimental data further demonstrates the performance of the proposed method.

For FMT reconstruction, the choice of the regularization parameter will have a significant impact on the results. A large parameter value can make the reconstructed solution deviate from the real distribution, while a small value will have little contribution to the regularization of the problem. Finding the optimal or near-optimal regularization parameter automatically still remains a challenging task. Generally speaking, two strategies can be used: determine the parameter in advance or update it heuristically. This will be our future research.

Another important issue lies in the accuracy of the photon propagation model itself. The diffuse equation has been extensively utilized to describe light transport in biological tissue, yet it is not applicable in certain regions, such as void or more absorptive regions. Several improved models, e.g. higher order approximation to radiative transfer equation, have been proposed to resolve the problem, though more computations are typically needed and the physical meanings are not such explicit. Since FMT reconstruction is a linear inverse problem in nature, the proposed reconstruction algorithm can potentially be utilized in these improved models.

In conclusion, we have developed a fast reconstruction algorithm with sparsity constraint for FMT. Numerical simulations and physical experiments show the merits of our method compared to the Newton method with Tikhonov regularization. *In vivo* mouse studies using the proposed method will be reported in the future.

Acknowledgments

This paper is supported by the Project for the National Basic Research Program of China (973) under Grant No.2006CB705700, Changjiang Scholars and Innovative Research Team in University (PCSIRT) under Grant No.IRT0645, CAS Hundred Talents Program, CAS scientific research equipment develop program under Grant No. YZ200766, the Knowledge Innovation Project of the Chinese Academy of Sciences under Grant No. KGCX2-YW-129, KSCX2-YW-R-262, the National Natural Science Foundation of China under Grant No. 30672690, 30600151, 60532050, 60621001, 30873462, 60910006, 30970769, 30970771, Beijing Natural Science Fund under Grant No.4071003, Science and Technology Key Project of Beijing Municipal Education Commission under Grant No.KZ200910005005.