Generalized conditional gradient method with adaptive regularization parameters for fluorescence molecular tomography

YI CHEN,1,2 MENGFEI DU,1,2 JUN ZHANG,1,2 GEGE ZHANG,1,2 LINZHI SU,1,2,4 KANG LI,1,2 FENGFUN ZHAO,1 HUANGJIAN YI,1 LIN WANG,3 AND XIN CAO1,2,5

1School of Information Science and Technology, Northwest University, Xi’an, Shaanxi 710127, China
2National and Local Joint Engineering Research Center for Cultural Heritage Digitization, Xi’an, Shaanxi 710127, China
3School of Computer Science and Engineering, Xi’an University of Technology, Xi’an, Shaanxi 710048, China
4sulinzhi029@163.com
5xin_cao@163.com

Abstract: Fluorescence molecular tomography (FMT) is an optical imaging technology with the ability of visualizing the three-dimensional distribution of fluorescently labelled probes in vivo. However, due to the light scattering effect and ill-posed inverse problems, obtaining satisfactory FMT reconstruction is still a challenging problem. In this work, to improve the performance of FMT reconstruction, we proposed a generalized conditional gradient method with adaptive regularization parameters (GCGM-ARP). In order to make a tradeoff between the sparsity and shape preservation of the reconstruction source, and to maintain its robustness, elastic-net (EN) regularization is introduced. EN regularization combines the advantages of L1-norm and L2-norm, and overcomes the shortcomings of traditional Lp-norm regularization, such as over-sparsity, over-smoothness, and non-robustness. Thus, the equivalent optimization formulation of the original problem can be obtained. To further improve the performance of the reconstruction, the L-curve is adopted to adaptively adjust the regularization parameters. Then, the generalized conditional gradient method (GCGM) is used to split the minimization problem based on EN regularization into two simpler sub-problems, which are determining the direction of the gradient and the step size. These sub-problems are addressed efficiently to obtain more sparse solutions. To assess the performance of our proposed method, a series of numerical simulation experiments and in vivo experiments were implemented. The experimental results show that, compared with other mathematical reconstruction methods, GCGM-ARP method has the minimum location error (LE) and relative intensity error (RIE), and the maximum dice coefficient (Dice) in the case of different sources number or shape, or Gaussian noise of 5%–25%. This indicates that GCGM-ARP has superior reconstruction performance in source localization, dual-source resolution, morphology recovery, and robustness. In conclusion, the proposed GCGM-ARP is an effective and robust strategy for FMT reconstruction in biomedical application.

© 2023 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

1. Introduction

Fluorescence molecular imaging (FMI) is a non-invasive imaging technique that detects the surface fluorescence distribution emitted from fluorescent sources within biological tissue by using a high-sensitivity detectors such as charge-coupled device (CCD) or scientific complementary metal–oxide–semiconductor (sCMOS) camera, photomultiplier tubes (PMT), and other photosensitive technologies [1–3]. However, due to the absorption and scattering of light
transmission [4,5], FMI can only obtain the photon distribution information on the surface of the object, but three-dimensional (3D) spatial information of the fluorescent target cannot be obtained [6]. Therefore, based on FMI, fluorescence molecular tomography (FMT) has been developed as an imaging modality to obtain the spatial distribution and quantitative analysis of interior fluorescent probes via the reconstruction method, which overcomes the difficulty of quantifying FMI [7,8]. Because of its high specificity, strong sensitivity, and low cost, FMT has been widely applied in preclinical study and diagnosis based on various small animal models [9–11].

Due to the severe scattering of photons propagation through heterogeneous tissues, the complexity of the photon propagation model, and the highly ill-posed inverse problem, FMT reconstruction is still a challenging problem [12–14]. To solve these critical problems, researchers have proposed many different model-based optimization methods, which need appropriate priors or penalties to promote reconstruction and restrict the search space to a specific solution set. One of the effective strategies is to utilize the anatomical information of different biological tissues, which is obtained by computed tomography (CT) or magnetic resonance imaging (MRI), and is used as the prior information of the photon propagation model to construct a more accurate forward model [15,16], so the spatial resolution can be promoted. Furthermore, the $L_p$-norm regularization ($p \in [0, 2]$) of the unknown fluorescent source is also used to constrain the FMT reconstruction [17–19]. Mathematically, $L_0$-norm is a sparsest constraint, but the reconstruction algorithm based on $L_0$-norm is a combinatorial optimization problem and an NP-hard problem [20,21]. When $p = 2$, the over-smoothness of the $L_p$-norm will cause reconstruction artifacts and exacerbates the noise effect [22,23]. When the value of $p$ is small ($p \in (0, 1)$), the $L_p$-norm tend to produce an over-sparse and incomplete boundary of the reconstruction target [24]. To overcome the drawbacks of the classical sparsity regularization, Zou et al. proposed an elastic-net (EN) regularization [25], which combines $L_1$ and $L_2$ norms with different weights. It not only ensures the sparsity, but also improves the smoothness of the reconstruction source [26]. Liu et al. applied EN regularization to FMT reconstruction for the first time and achieved good reconstruction performance [27]. Later, Wang et al. proposed an adaptive parameter search elastic net (APSEN) method for FMT reconstruction [26], which addresses the issue of parameter selection in reconstruction. However, the coordinate descent method is adopted by APSEN to optimize the objective function, and the optimal solution is solved linearly along the coordinate axis, which is easy to fall into local optimization when solving large-scale sparse models.

It should be noted that the reconstruction performance of the regularization method will be affected by the regularization parameters [28]. The larger the regularization parameters, the greater the influence of the regularization term on the reconstruction results. However, the optimal selection of regularization parameters is usually unknown, which depends on the specific reconstruction problem and the properties of the sparse vector [29]. Recently, some methods for selecting the optimal regularization parameters have been proposed, such as L-curve, U-curve, and cross-validation [30–32], which are of great significance to improve the reconstruction performance of FMT.

In addition to numerical methods, deep learning has also been introduced into FMT reconstruction. Guo et al. proposed an end-to-end 3D depth encoder network [33], which can significantly improve image quality and reduce reconstruction time. Meng et al. proposed a local connection network based on the K-nearest neighbor to improve the morphological reconstruction performance of FMT [34]. Zhang et al. proposed a three-dimensional fusion double-sampling convolution neural network to achieve FMT ultra-high spatial resolution reconstruction [35]. The deep learning method reconstructs the fluorescent source by directly establishing an end-to-end mapping model on a large dataset, which can greatly eliminate the modeling error [36]. Nevertheless, the deep learning method has two common shortcomings. On the one hand, the trained neural network can only be used for specific imaging objects, which means its generalization
ability is weak [37–39]. On the other hand, compared with traditional mathematical methods, the interpretability of neural network is poor.

In this work, a generalized conditional gradient method for adaptive regularization parameters (GCGM-ARP) is proposed. The ill-conditioned inverse problem is interpreted as a non-differentiable object function with EN regularization, in which the EN regularization is a combination of $L_1$ and $L_2$-norms to balance the sparsity and shape recovery of the reconstructed fluorescent source. GCGM-ARP adopts L-curve to effectively generate regularization parameters suitable for different fluorescence distributions. The objection function is solved using the generalized conditional gradient method (GCGM). The per-iteration cost of the GCGM is particularly cheap and globally convergent, making it widely applied in fast sparse approximation. The method consists of two minimization sub-problems: determining the descent direction and determining the step size. The minimization process of determining the descent direction is simplified to a shrinkage process by the iterative soft thresholding algorithm (ISTA) [40], and the step size sub-problem can be efficiently solved by the alternating direction method of multipliers (ADMM) [41].

To assess the performance of the GCGM-ARP method, a series of numerical simulations and in vivo experiments were carried out. Iterative shrinkage with $L_1$-norm (IS-$L_1$) [13], incomplete variables truncated conjugate gradient with $L_1$-norm (IVTCG-$L_1$) [42], Nesterov’s method with EN regularization (N-EN) [27], and APSEN method were used for comparison. The results of the numerical experiments showed that GCGM-ARP performed the lowest location error (LE) and relative intensity error (RIE), as well as the highest Dice coefficient (Dice), as compared to the other methods. These findings indicated the superiority of the GCGM-ARP methods in terms of localization, shape recovery, and dual-source resolution. Moreover, the in vivo experiments further verified the practical applicability of GCGM-ARP method.

The novelty of this paper lies in the proposal of a GCGM-ARP method based on the EN regularization model. Firstly, in comparison to other methods based on the EN regularization model applied to FMT, the principle of the L-curve in the GCGM-ARP method is easier to understand and use in parameter selection. Secondly, the GCGM algorithm avoids falling into local optimization by selecting the gradient descent direction.

The remainder of this paper is structured as follows: the Section 2 introduces the FMT reconstruction model and GCGM-ARP reconstruction algorithm. The Section 3 introduces the evaluation index, the design of experiments and the result of experiments. Finally, the Section 4 gives discussion and conclusion of our work.

2. Methodology

2.1. Photon propagation model

In highly scattering media, such as biological tissues, the photon propagation model can be approximated by the diffusion equation (DE). In steady-state FMT with point excitation sources, the photon propagation can be described by a coupled diffusion equation with Robin boundary condition, which is defined as [43,44]:

\[
\begin{align*}
\nabla \cdot [D_r(r) \nabla \Phi_r(r)] - \mu_{at}(r) \Phi_r(r) & = -\Theta \delta(r - r_l), \quad (r \in \Omega) \\
\nabla \cdot [D_m(r) \nabla \Phi_m(r)] - \mu_{am}(r) \Phi_m(r) & = -\Phi_r(r) \eta \mu_{af}(r), \quad (r \in \Omega) \\
2D_x,m(r) \nabla \Phi_{x,m}(r) + q \Phi_{x,m}(r) & = 0, \quad (r \in \partial \Omega)
\end{align*}
\]
where \( \nabla \) is the gradient operator, \( \Omega \) denotes the entire domain of the problem, and \( r \) denotes the position of the nodes in \( \Omega \); \( r_l \) is the position of a point excitation source with an amplitude of \( \Theta \) of the mean free path of photon transmission located below the surface of \( \Omega \). \( \eta \) is the quantum efficiency, and \( q \) is the optical reflectivity. The subscripts \( x \) and \( m \) denote the excitation wavelength and emission wavelength, respectively. \( \Phi_x(r) \) and \( \Phi_m(r) \) denote the excitation and emission photon flux density at position \( r \), respectively. \( D_{x,m} = \frac{1}{3} [\mu_{ax,m} + (1-g)\mu_{sx,m}] \) denotes the diffusion coefficient, where \( g \) denotes the anisotropy parameter, \( \mu_{ax,m} \) and \( \mu_{sx,m} \) denote the absorption coefficient and scattering coefficient respectively. \( \eta \mu_{af}(r) \) is the unknown fluorophore distribution to be reconstruct, where \( \mu_{af} \) is absorption coefficient of the fluorescence agent.

By using the finite element method (FEM) to discretize the photon propagation model, the partial differential Eq. (1) is linearized into the following linear equations:

\[
\begin{align*}
M_x \Phi_x &= S_x \\
M_m \Phi_m &= G X
\end{align*}
\]

where \( M_x \) and \( M_m \) is forward system matrices of the excitation and emission propagation process, respectively. The vector \( S_x \) represents the distribution of discrete excitation point source. The matrix \( G \) is obtained according to the discretization of fluorescent source, which can be calculated as follows:

\[
G_x(i,j) = \int_{\Omega} \Phi_x(r) B_i(r) B_j(r) \, dr
\]

where \( B_i(r) \) and \( B_j(r) \) denote the base functions of the node \( i \) and node \( j \), respectively. The vector \( X \) represents the reconstructed fluorescent source. \( M_m \) is symmetric positive matrix \([45] \), so the relationship between \( \Phi_m \) and \( X \) in Eq. (2) can be expressed as a Matrix-form equation:

\[
\Phi_m = M_m^{-1} G X = A_m X
\]

Because only partial nodes can be detected, the immeasurable nodes are removed from \( \Phi_m \) and the corresponding rows are removed from \( A_m \). The final matrix equation can be expressed as:

\[
\Phi = AX
\]

where \( \Phi \in \mathbb{R}^{M \times 1} \) represents the measured photon flux at the boundary of the biological tissue, \( A \in \mathbb{R}^{M \times N} \) denotes system matrix, and \( X \in \mathbb{R}^{N \times 1} \) denotes the distribution of internal fluorescent source.

In general, the inverse problem of Eq. (5) is ill-posed, so it is impractical to solve \( X \) directly. To alleviate the ill-posedness of the inverse problem, the EN regularization was adopted in FMT reconstruction, which is defined as:

\[
\min_X E(X) = \frac{1}{2} \| AX - \Phi \|_F^2 + \alpha \| X \|_1 + \frac{\beta}{2} \| X \|_2^2
\]

where \( E(X) \) represents the objective function. \( \| \cdot \|_F \) represents the Frobenius norm; \( \alpha > 0 \) and \( \beta > 0 \) represent the regularization parameters of \( L_1 \)-norm and \( L_2 \)-norm, respectively. The objective
The function in Eq. (6) can be equivalently expressed as:

$$\min_X E(X) = \frac{1}{2} \|AX - \Phi\|_F^2 + R(X)$$  \hspace{1cm} (7)

where $$R(X) = \alpha \|X\|_1 + \frac{\beta}{2} \|X\|_2^2$$ is the EN regularization. Compared with the classical $$L_1$$-norm sparse regularization, there is an additional $$L_2$$-norm in $$R(X)$$. And EN regularization has the advantages of both $$L_1$$-norm regularization and $$L_2$$-norm regularization [46]. Moreover, statistically, EN regularization is more stable than the classical $$L_1$$-norm regularization [47]. So, it is more suitable for ill-conditioned inverse problem.

Obviously, EN regularization is non-differentiable, so the simple gradient descent minimization method cannot be used to solve the problem. For this reason, here, the GCGM-ARP algorithm was utilized to address the inverse problem.

### 2.2. Reconstruction based on GCGM-ARP scheme

The GCGM algorithm is proposed by Bredies K et al. to address the following form of minimization problem [48]:

$$\min_X F(X) + G(X)$$  \hspace{1cm} (8)

where $$F$$ is assumed to be smooth, but $$G$$ is assumed to be proper, convex, lower semi-continuous and coercive in Hilbert space. More detailed conditions about the functional $$G$$ that need to be satisfied can be found in [48].

In this work, GCGM is considered for FMT reconstruction. The objective function in Eq. (6) is rewritten as follows:

$$E(X) = F(X) + G(X)$$  \hspace{1cm} (9)

where:

$$F(X) = \frac{1}{2} \|AX - \Phi\|_F^2 - \Psi(X)$$

$$G(X) = R(X) + \Psi(X)$$  \hspace{1cm} (10)

$$\Psi(X) = \frac{\lambda}{2} \|X\|_2^2 - \frac{\beta}{2} \|X\|_2^2, \ \lambda > 0$$

There are two reasons for $$\Psi(X) = \frac{\lambda}{2} \|X\|_2^2 - \frac{\beta}{2} \|X\|_2^2$$. Firstly, in this case, $$\Psi(X)$$ has possess of desirable properties, such as proper, convex, lower semi-continuous and coercive in Hilbert space. Secondly, it can be found that ISTA can be applied to the minimization of the objective function $$E(X)$$, as the both sub-problems of GCGM involve $$L_1$$-norm regularization.

Thus, Eq. (8) can be expressed as:

$$\min_X \{X = F(X) + G(X)\}$$  \hspace{1cm} (11)

Next, the GCGM from the literature [48] is showed in Algorithm 1.

It is worth noting that throughout this paper, the operator symbol $$\langle \cdot, \cdot \rangle$$ represents the inner product.

In Algorithm 1, there are two sub-problem of minimization that need to be solved, which are the determining direction sub-problem and the determining step size sub-problem. To explain our proposed method more intuitively, these two sub-problems will be discussed respectively.
Algorithm 1. GCGM

Input: System matrix $A$, measured surface photon distribution $\Phi$, regularization parameters $\alpha$ and $\beta$.

Initialization: fluorescent source distribution $X^0$, descent direction $Y^0$, step size $Z^0$, iteration number index $k$, maximum iteration number $K$, threshold error $err = 1e^{-6}$

While $\|X^{k+1} - X^k\|_2 > err$ or $k < K$ do
1: Determine a descent direction $Y^k$ as a solution of
\[
\min_Y \langle A^T (AX - \Phi) - (\lambda - \beta)X, Y \rangle + \alpha \|Y\|_1 + \frac{\lambda}{2} \|Y\|_2^2
\]
2: Determine a step size $Z^k$ as a solution of
\[
\min_Z F(Y^k + Z \cdot (Y^k - X^k)) + G(X^k + Z \cdot (Y^k - X^k))
\]
3: $X^{k+1} = X^k + Z^k \cdot (Y^k - X^k)$
4: $k = k + 1$

Output: $X = X^K$

2.2.1. Sub-problem of determining the descent direction

Since the Fréchet derivative of $F$ is used in Algorithm 1, its mathematical formula is given:

\[
F'(X) = A^T (AX - \Phi) - (\lambda - \beta)X.
\] (12)

Thus, the descent direction $Y^k$ in the Algorithm 1 can be rewritten as follows:

\[
\min_Y \langle A^T (AX^k - \Phi) - (\lambda - \beta)X^k, Y \rangle + \alpha \|Y\|_1 + \frac{\lambda}{2} \|Y\|_2^2
\] (13)

The Eq. (13) can be solved by explicit calculation. Then, the explicit componentwise expression can be obtained:

\[
Y^k_i + \frac{\alpha}{\lambda} \text{sign}(Y^k_i) = \left( \frac{1}{\lambda} A^T (\Phi - AX^k) + \left( 1 - \frac{\beta}{\lambda} \right) X^k \right)_i
\] (14)

where the subscript $i$ represents the $i$-th component. The solution of the Eq. (14) can be expressed by soft threshold function $S_\gamma$, which is defined as:

\[
S_\gamma(t) = \begin{cases} 
    t - \frac{\alpha}{\lambda}, & t \geq \frac{\alpha}{\lambda} \\
    0, & -\frac{\alpha}{\lambda} \leq t \leq \frac{\alpha}{\lambda} \\
    t + \frac{\alpha}{\lambda}, & t \leq -\frac{\alpha}{\lambda}
\end{cases}
\] (15)

where $\gamma = \frac{\alpha}{\lambda}$ and $t \in \mathbb{R}$.

It is obvious that determining the descent direction sub-problem can be addressed by ISTA, which can be formulated as:

\[
Y^k = S_\gamma \left( \frac{1}{\lambda} A^T (\Phi - AX^k) + \left( 1 - \frac{\beta}{\lambda} \right) X^k \right)
\] (16)

2.2.2. Sub-problem of determining step size

In this sub-problem, our objective function can be simplified to the following form:

\[
\min_Z F(X^k + WZ) + G(X^k + WZ)
\] (17)
where \(X^k\) is a constant, and to avoid confusion, we let \(X^k = X\). The symbol \(W \in \mathbb{R}^{N \times N}\) is a diagonal matrix, which can be defined as follows:

\[
W = \begin{pmatrix}
Y_1^k - X_1^k \\
\vdots \\
Y_N^k - X_N^k
\end{pmatrix}
\]  

Combining Eq. (6) and Eq. (10), then \(G(X)\) can be formulated as:

\[
G(X) = \alpha \|X\|_1 + \frac{\lambda}{2} \|X\|_2^2
\]  

(19)

ADMM is a method for solving convex optimization problems. It breaks the objective function of the original problem into smaller sub-problems, then solves each sub-problem in parallel, and finally balances the solution of each sub-problem to obtain the global solution of the original problem \([49]\). It can be noticed that the objective function of the Eq. (17) is non-differentiable but convex, so ADMM is applied to address the minimization problem in the Eq. (16).

An auxiliary variable is defined as \(S = X + WZ\). According to ADMM, Eq. (17) can be reformulated as an optimization problem with equality constraints:

\[
\frac{1}{2} \|A(X + WZ) - \Phi\|_F^2 + \frac{\lambda}{2} \|X + WZ\|_2^2 - \frac{\rho}{2} \|X + WZ\|_2^2 + \alpha \|S\|_1 + \frac{\beta}{2} \|S\|_2^2,
\]  

s.t \(S = X + WZ\)

(20)

The augmentation Lagrange function of the Eq. (20) can be expressed as:

\[
\mathcal{L}_\rho(S, Z, V) = \frac{1}{2} \|A(X + WZ) - \Phi\|_F^2 + \frac{\lambda}{2} \|X + WZ\|_2^2 - \frac{\rho}{2} \|X + WZ\|_2^2 + \alpha \|S\|_1 \\
+ \frac{\beta}{2} \|S\|_2^2 + V^T(S - X - WZ) + \frac{\rho}{2} \|S - X - WZ\|_2^2
\]

(21)

where \(\rho>0\) is a penalty parameter, \(V\) is a Lagrange multiplier. The Eq. (21) is solved by the following iterative scheme:

\[
\begin{align*}
S^{k+1} &= \arg\min_S L\rho(S, Z^k, V^k) \\
&= \arg\min_S \left\{ \alpha \|S\|_1 + \frac{\beta}{2} \|S\|_2^2 + V^T(S - X - WZ) \right\} \\
&\quad + \frac{\rho}{2} \|S - X - WZ\|_2^2 \\
S^{k+1} &= \arg\min_Z L\rho(S^k, Z, V^k) \\
&= \arg\min_Z \left\{ \frac{1}{2} \|A(X + WZ) - \Phi\|_F^2 + \frac{\lambda}{2} \|X + WZ\|_2^2 \right\} \\
&\quad + \frac{\rho}{2} \|S - X - WZ\|_2^2 \\
V^{k+1} &= V^k - \rho(S^{k+1} - X - WZ^{k+1})
\end{align*}
\]  

(22)

where \(Z^{k+1}\) can be solved directly by derivation, but \(S^{k+1}\) can not. However, the soft threshold function in Eq. (15) can be adopted to address this optimization problem. First, the explicit
calculation result of the $S^{k+1}$ minimization problem is as follows:

$$S^{k+1} + \frac{\alpha}{\beta + \rho} \text{sign}(S^{k+1}) = \frac{1}{\beta + \rho} (\rho X + WZ - V)$$  \hspace{1cm} (23)$$

By comparing Eq. (13), the solution of Eq. (22) can be obtained, as follows:

$$S^{k+1} = \frac{1}{\beta + \rho} (\rho X + WZ - V).$$  \hspace{1cm} (24)$$

So, the sub-problem of determining the step size is solved.

**Algorithm 2. GCGM-ARP**

**Input:** System matrix $A$, measured surface photon distribution $\Phi$, regularization parameters $\alpha$ and $\beta$.

**Initialization:** fluorescent source distribution $X^0$, descent direction $Y^0$, step size $Z^0$, iteration number index $k$, maximum iteration number $K$, error threshold $err = 1e^{-6}$, penalty parameter $\rho$, optimal regularization parameter list $O_\alpha = [\alpha_1, \alpha_2; \cdots, \alpha_p]$, regularization parameter $\beta$.

**Step1:** Reconstructed fluorescent source obtained by using different regularization parameter $\alpha$.

For $j = 1$ to $p$ do
1: optimal regularization parameter $\alpha = O_\alpha[j]$
2: While $\|X^k - X^{k-1}\|_2 > err$ or $k < K$ do
   1: The descent direction $Y^k$ is determined via Eq. (16)
   2: The step size $Z^k$ is determined by iterative Eq. (22)
   3: $X^{k+1} = X^k + Z^k : (Y^k - X^k)$
   4: $k = k + 1$
End while
3: Save the reconstruction source $X^*_j$ obtained each time
End for

**Step2:** The optimal parameter $\alpha^*$ and its corresponding reconstructed fluorescent source $X^*$ are obtained by L-curve.

**Output:** $X = X^*$.

### 2.2.3. Adaptive regularization parameters

As described above, it is difficult to select the optimal regularization parameters in FMT reconstruction. If $\beta$ in Eq. (6) is forced to equal 0, the adaptive EN regularization will reduce into the adaptive lasso. According to the opinion of Zou and Hastie [25], it can be easily proved that, no matter what the value of $\beta$ is, the adaptive EN regularization will be reduced to an adaptive lasso [50]. Thus, regularization parameter $\beta$ will be set empirically, and only regularization parameter $\alpha$ needs to be adjusted adaptively.

For the optimization problem based on Eq. (6), the L-curve will be applied to adaptively adjust the regularization parameter $\alpha$. It can show the tradeoff between the size of the regularization solution and its fitness to the given data, when the regularization parameters change. In general, the L-curve is a log-log plot, which is composed of the solution norm $X_2$ and the corresponding residual norm $\|AX - \Phi\|_2$.

Based on above statement, the complete procedure of GCGM-ARP strategy is summarized in Algorithm 2.
3. Experiments and results

In this section, numerical simulations and in vivo experiment were performed to validate the reconstruction performance of the GCGM-ARP strategy. Furthermore, IS-\(L_1\), IVTCG-\(L_1\), N-EN, APSEN methods were selected for comparison in terms of accuracy, shape recovery, and in vivo practicability. All the programs for the reconstruction algorithm were implemented using MATLAB (2019b) on desktop computer with Intel Core i3-10100 CPU (3.60 GHz) and 8GB RAM.

3.1. Regularization parameters and evaluation index

Suppose \(\eta_1\) and \(\eta_2\) are the regularization parameter of \(L_1\) and \(L_2\) norms used in the algorithm for the above comparison. In order to ensure the convergence of all algorithms, the maximum number of iterations was set to 1000 and the error thresholds were set to 1e-6 through experience. Moreover, the optimal regularization parameters of GCGM-ARP method were determined by L-curve, and the best regularization parameters of IS-\(L_1\), IVTCG-\(L_1\), and N-EN methods were determined according to related work and experience [13,27,42]. APSEN method can adaptively adjust regularization [26]. The optimal regularization parameters of different algorithms in different experiments have been shown in Table 1.

<table>
<thead>
<tr>
<th>Methods</th>
<th>IS-(L_1)</th>
<th>IVTCG-(L_1)</th>
<th>N-EN</th>
<th>APSEN</th>
<th>GCGM-ARP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-source</td>
<td>(\eta_1 : 2e-10)</td>
<td>(\eta_1 : 1e0)</td>
<td>(\eta_1 : 1e-5)</td>
<td>(\eta_1 : 0.5)</td>
<td>(\alpha : 8e-6)</td>
</tr>
<tr>
<td></td>
<td>(\eta_2 : 1e-6)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\beta : 1e-5)</td>
<td></td>
</tr>
<tr>
<td>Dual-source</td>
<td>(\eta_1 : 2e-10)</td>
<td>(\eta_1 : 1e0)</td>
<td>(\eta_1 : 1e-5)</td>
<td>(\eta_1 : 0.9)</td>
<td>(\alpha : 1e-5)</td>
</tr>
<tr>
<td></td>
<td>(\eta_2 : 1e-6)</td>
<td>(\eta_2 : 1.1e-9)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\beta : 1e-5)</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>(\eta_1 : 2e-10)</td>
<td>(\eta_1 : 1e0)</td>
<td>(\eta_1 : 1e-5)</td>
<td>(\eta_1 : 0.5)</td>
<td>(\alpha : 2e-6)</td>
</tr>
<tr>
<td></td>
<td>(\eta_2 : 1e-6)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\beta : 1e-5)</td>
<td></td>
</tr>
<tr>
<td>Robustness</td>
<td>(\eta_1 : 2e-10)</td>
<td>(\eta_1 : 1e0)</td>
<td>(\eta_1 : 1e-5)</td>
<td>(\eta_1 : 0.6)</td>
<td>(\alpha : 1e-5)</td>
</tr>
<tr>
<td></td>
<td>(\eta_2 : 1e-6)</td>
<td>(\eta_2 : 1.67e-9)</td>
<td>(\eta_2 : 1.67e-9)</td>
<td>(\beta : 1e-5)</td>
<td></td>
</tr>
<tr>
<td>15%</td>
<td>(\eta_1 : 2e-10)</td>
<td>(\eta_1 : 1e0)</td>
<td>(\eta_1 : 1e-5)</td>
<td>(\eta_1 : 0.8)</td>
<td>(\alpha : 1e-4)</td>
</tr>
<tr>
<td></td>
<td>(\eta_2 : 1e-6)</td>
<td>(\eta_2 : 1.25e-9)</td>
<td>(\eta_2 : 1.25e-9)</td>
<td>(\beta : 1e-5)</td>
<td></td>
</tr>
<tr>
<td>25%</td>
<td>(\eta_1 : 2e-10)</td>
<td>(\eta_1 : 1e0)</td>
<td>(\eta_1 : 1e-5)</td>
<td>(\eta_1 : 0.5)</td>
<td>(\alpha : 1e-5)</td>
</tr>
<tr>
<td></td>
<td>(\eta_2 : 1e-6)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\eta_2 : 2e-9)</td>
<td>(\beta : 1e-5)</td>
<td></td>
</tr>
</tbody>
</table>

To quantitatively assess the accuracy of FMT reconstruction using different methods, LE, Dice, and RIE were adopted as quantitative indexes in this work.

LE measures the Euclidean distance between the real source center and the reconstructed source center, which is defined as:

\[
LE = \|L_{real} - L_{recon}\|_2
\]

(25)

where \(L_{real}\) and \(L_{recon}\) denote the center coordinates of the real fluorescent source and the reconstructed source region, respectively. The lower the LE, the more accurate the reconstruction location.

Dice is introduced to assess the spatial overlap performance of the real source region and the reconstructed source region, which is defined as:

\[
Dice = \frac{2|R_{real} \cap R_{recon}|}{|R_{real}| + |R_{recon}|}
\]

(26)

where \(R_{real}\) and \(R_{recon}\) represent the reconstructed region and the real fluorescent region, respectively. The higher the Dice index, the better the morphological reconstruction. Specifically, the reconstructed source region is determined by the non-zero tetrahedron region based on \(X\).
RIE is used to evaluate the intensity deviation between the real source intensity and the reconstructed source intensity, which is defined as:

$$RIE = \frac{|I_{\text{real}} - I_{\text{rcon}}|}{I_{\text{real}}}$$  \hspace{1cm} (27)

where $I_{\text{real}}$ and $I_{\text{rcon}}$ is the intensity of real fluorescent source and the reconstructed source, respectively. The smaller the RIE, the better the fluorescent intensity recovery of the reconstructed source.

3.2. Experimental setting

3.2.1. Numerical simulation

The heterogeneous model based on a cylinder with radius of 10 mm and height of 30 mm was adopted in the numerical simulations. It consisted of five biological tissues: muscle, heart, lung, liver, and bone, as shown in Fig. 1(a). In this work, all the experiments were carried out under the excitation wavelength of 680 nm and the emission wavelength of 750 nm. The optical parameters of each biological tissue have been listed in detail in Table 2 [51].

![Figure 1](image-url)

**Fig. 1.** (a) is the 3D view of the cylinder model with single-source. (b) is the mesh for inverse problem. (c) shows the position of 4 excitation sources. (d) is the photon distribution simulated on the surface by the single fluorescence source.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>$\mu_a (\text{mm}^{-1})$</th>
<th>$\mu'_a (\text{mm}^{-1})$</th>
<th>$\mu_d (\text{mm}^{-1})$</th>
<th>$\mu'_d (\text{mm}^{-1})$</th>
<th>$g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscle</td>
<td>0.0745</td>
<td>0.4115</td>
<td>0.0474</td>
<td>0.3122</td>
<td>0.97</td>
</tr>
<tr>
<td>Bone</td>
<td>0.0521</td>
<td>2.4415</td>
<td>0.0326</td>
<td>2.114</td>
<td>0.93</td>
</tr>
<tr>
<td>Liver</td>
<td>0.3016</td>
<td>0.6676</td>
<td>0.1921</td>
<td>0.6023</td>
<td>0.93</td>
</tr>
<tr>
<td>Lung</td>
<td>0.1681</td>
<td>2.1569</td>
<td>0.1045</td>
<td>2.0477</td>
<td>0.93</td>
</tr>
<tr>
<td>Heart</td>
<td>0.0504</td>
<td>0.9437</td>
<td>0.0331</td>
<td>0.8203</td>
<td>0.90</td>
</tr>
</tbody>
</table>

In the forward process, a mesh consisting of 29415 nodes and 168046 tetrahedral elements was used to simulate the photon propagation. The intensity of each source is 1 nw/mm³, and the fluorescence distribution of the surface was simulated using the molecular optical simulation environment (MOSE) [51] based on Monte Carlo method, as shown in Fig. 1(d). In the inverse process, the cylinder model was segmented into a mesh using Comsol Multiphysics software [52]. The mesh included 4626 nodes and 25840 tetrahedral elements, as shown in Fig. 1(b). The
excitation sources were located at four positions along the $Z = 15\text{mm}$ plane, as illustrated in Fig. 1(c).

We designed single-source and dual-source simulation experiment to evaluate the performance of the GCGM-ARP reconstruction method. In the single-source simulation experiment, one spherical fluorescence source with a radius of 1 mm was placed at (-1, -1, 15) mm. In the dual-source simulation experiment, two cylindrical sources with a radius of 1 mm and a height of 2 mm, which were placed at the center coordinates of S1(2, 1, 10) mm and S2(2, 1, 18) mm, respectively.

As noise is inevitable in FMT, the anti-noise experiment is also designed to evaluate the robustness of our method. Gaussian noise of 5%, 15%, 25% was added to the measurement data based on single-source numerical simulation, and then reconstructed using different methods to observe the robustness and accuracy of these methods under different Gaussian noise ratios.

3.2.2. In vivo experiment

With the guidelines of the Animal Ethics Committee of the Northwestern University of China, in vivo experiment was carried out to further investigate the practical performance of GCGM-ARP. In vivo experimental data set was collected from an adult BALB/c mouse by a dual-modality FMT/CT system, and the chief components of the dual-modality FMT/CT system was exhibited in Fig. 2(a). To minimize the suffering of mouse, all animal experiments were performed under isoflurane gas anesthesia (3% isoflurane-air mixture). The detailed collection process was introduced as follows.

First, a spherical fluorescent bead with a radius of 1 mm containing Cy5.5 solution was implanted into the abdominal cavity of mouse as a fluorescence target. The fluorescent bead is wrapped in a plastic material, which can be easily detected by micro-CT to locate the real fluorescent region. Six hours later, a 680 nm continuous wave semiconductor laser was used to provide excitation illumination, and the surface fluorescence image with a 120° field of view was collected by a thermoelectric cooled electron multiplying charge coupled device (EMCCD) camera (-80°C, iXonEM+ 888) with an exposure time of 1 s. The emission light was captured and restrained the noise by a 750 ± 10 nm bandpass filter. After the acquisition of fluorescence image, the structure information of mouse needs to be collected by Micro-CT system (tube
voltage of 60kVp, x-ray power of 40W). The CT image was processed into 3D volume data, and Amria 5.2 (Amria, Visage Imaging, Australia) was used to segment the main organs, including muscle, lung, heart, stomach, liver, and kidney, and then integrate into the xenogeneic mouse model. The true central position of the fluorescent target is (17.5, 21.5, 9.0) mm, and the optical parameter for different organs were from the literature [53]. Thus, the structure of the in vivo experimental process can be summarized, as shown in Fig. 2(b).

### 3.3. Experimental results

#### 3.3.1. Single-source simulation reconstruction

In the spherical single-source simulation experiment, the 3D and slice views of the reconstructed results with different methods were shown in Fig. 3. In the 3D view, different organs were depicted in different colors, and the reconstruction source was depicted in red. In the slice view, the white circle represents the shape and region of the real source, while the red region represents the source of the reconstruction. By observing the slice view, it can be found that although all the methods can approximately reconstruct the fluorescence source region, the overlapping between GCGM-ARP and real source is more, so a better morphological reconstruction is obtained. The quantitative analysis of the reconstruction results of the four methods was listed in Table 3. From these results, it is obvious that our method achieves the smallest LE and RIE, and the largest Dice coefficient. This indicated that GCGM-ARP possessed of superior positioning ability and shape recovery ability.

<table>
<thead>
<tr>
<th>Method</th>
<th>True center (mm)</th>
<th>Reconstructed center (mm)</th>
<th>LE (mm)</th>
<th>Dice</th>
<th>RIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS-L₁</td>
<td>(-1, -1, 15)</td>
<td>(-0.198, -0.721, 15.310)</td>
<td>0.904</td>
<td>0.496</td>
<td>0.302</td>
</tr>
<tr>
<td>IVTCG-L₁</td>
<td>(-1, -1, 15)</td>
<td>(-0.829, -0.643, 14.311)</td>
<td>0.794</td>
<td>0.625</td>
<td>0.824</td>
</tr>
<tr>
<td>N-EN</td>
<td>(-1, -1, 15)</td>
<td>(-0.983, -1.193, 14.541)</td>
<td>0.498</td>
<td>0.693</td>
<td>0.553</td>
</tr>
<tr>
<td>APSEN</td>
<td>(-1, -1, 15)</td>
<td>(-8.145, -1.077, 14.737)</td>
<td>0.331</td>
<td>0.795</td>
<td>0.337</td>
</tr>
<tr>
<td>GCGM-ARP</td>
<td>(-1, -1, 15)</td>
<td>(-0.835, -0.884, 14.820)</td>
<td>0.270</td>
<td>0.952</td>
<td>0.181</td>
</tr>
</tbody>
</table>

#### 3.3.2. Dual-source simulation reconstruction

In Fig. 4, the 3D view of the reconstruction result in the first row, and the other rows correspond to transverse view, sagittal view, and coronal view, respectively. The enlarged view next to the slice view provides a close view of the distribution of the reconstructed area. The reconstruction sources obtained by IS-L₁ and IVTCG-L₁ were over-sparse, although N-EN and APSEN showed improved performance, the reconstruction result obtained by GCGM-ARP reconstruction was closest to the real sources. Obviously, the dual-source reconstruction performance of GCGM-ARP is better than other methods. The detailed reconstruction results of the four algorithms were shown in Table 4, which further confirms our observations. The LE of S1 and S2 obtained by GCGM-ARP method is much lower than that of other methods, which was 0.432 mm and 0.314 mm, and the Dice is up to 0.839 and 0.834, which demonstrated that GCGM-ARP achieved more accurate dual-source localization and morphological recovery.

#### 3.3.3. Robustness experiment

The influence of different intensity of Gaussian noise on the reconstruction results of the four methods is shown in the Fig. 5. Due to the influence of background fluorescent signal, the reconstruction results of the four methods are different from the single-source numerical simulation experiments without noise. However, under different ratios of Gaussian noise, compared with other methods, GCGM-ARP shows excellent reconstruction performance, with
Fig. 3. (a), (c), (e), (g) and (i) show the 3D reconstruction results of the IS-$L_1$, IVTCG-$L_1$, N-EN, APSEN, and GCGM-ARP, respectively. (b), (d), (f), (h), and (j) show the transverse slice views in the $Z = 15$ mm of the corresponding four methods.

Table 4. Quantitative results in the dual-source numerical simulation.

<table>
<thead>
<tr>
<th>Method</th>
<th>True center (mm)</th>
<th>Reconstructed center (mm)</th>
<th>LE (mm)</th>
<th>Dice</th>
<th>RIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS-$L_1$</td>
<td>(2, 1, 10)</td>
<td>(1.010, 1.075, 10.620)</td>
<td>1.170</td>
<td>0.462</td>
<td>0.833</td>
</tr>
<tr>
<td></td>
<td>(2, 1, 18)</td>
<td>(1.313, 0.349, 17.485)</td>
<td>1.077</td>
<td>0.446</td>
<td></td>
</tr>
<tr>
<td>IVTCG-$L_1$</td>
<td>(2, 1, 10)</td>
<td>(1.654, 0.732, 9.262)</td>
<td>0.858</td>
<td>0.548</td>
<td>0.749</td>
</tr>
<tr>
<td></td>
<td>(2, 1, 18)</td>
<td>(2.973, 1.407, 17.718)</td>
<td>1.092</td>
<td>0.298</td>
<td></td>
</tr>
<tr>
<td>N-EN</td>
<td>(2, 1, 10)</td>
<td>(2.524, 0.648, 9.925)</td>
<td>0.636</td>
<td>0.545</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2, 1, 18)</td>
<td>(1.851, 1.265, 17.510)</td>
<td>0.576</td>
<td>0.538</td>
<td>0.731</td>
</tr>
<tr>
<td>APSEN</td>
<td>(2, 1, 10)</td>
<td>(1.766, 0.671, 9.975)</td>
<td>0.489</td>
<td>0.783</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2, 1, 18)</td>
<td>(1.850, 1.324, 17.517)</td>
<td>0.571</td>
<td>0.698</td>
<td>0.674</td>
</tr>
<tr>
<td>GCGM-ARP</td>
<td>(2, 1, 10)</td>
<td>(1.848, 1.359, 10.220)</td>
<td>0.432</td>
<td>0.839</td>
<td>0.502</td>
</tr>
<tr>
<td></td>
<td>(2, 1, 18)</td>
<td>(1.832, 0.802, 18.177)</td>
<td>0.314</td>
<td>0.834</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 4. Reconstruction results of different method for dual-source simulation.
the lowest LE and RIE, and the largest Dice. Consequently, anti-noise experiments showed that GCGM-ARP is the most robust among these methods.

![Fig. 5](image)

**Fig. 5.** The quantitative analysis of different methods under different Gaussian noise levels.

### 3.3.4. In vivo experiment reconstruction

The 3D view and slice view were displayed in Fig. 6, which prove the practicality of our methods in vivo. According to Fig. 6, it was showed that there was a deviation between the reconstruction position of IS-$L_1$ and IVTCG-$L_1$, and there were many reconstruction artifacts produced by N-EN. While the performance of APSEN improved, it still cannot match the accuracy achieved by GCGM-ARP, which achieved more accurate reconstruction, including the most accurate spatial location, and introduced less reconstruction artifacts. Here, it should be noted that, because of the unknown real source intensity, all the evaluation indicators related to the source intensity (RIE) were unavailable in vivo experiments. As shown in Table 5, the GCGM-ARP method indicated the best accuracy with the least LE and largest Dice similarity. These quantitative results further indicated the superior performance of GCGM-ARP method in obtaining the morphology and localization of fluorescence probe distribution in mouse.

<table>
<thead>
<tr>
<th>Method</th>
<th>True center (mm)</th>
<th>Reconstructed center (mm)</th>
<th>LE (mm)</th>
<th>Dice</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS-$L_1$</td>
<td>(17.5, 21.5, 9.0)</td>
<td>(17.998, 20.641, 8.616)</td>
<td>1.065</td>
<td>0.254</td>
</tr>
<tr>
<td>IVTCG-$L_1$</td>
<td>(17.5, 21.5, 9.0)</td>
<td>(16.652, 21.356, 8.296)</td>
<td>1.112</td>
<td>0.254</td>
</tr>
<tr>
<td>N-EN</td>
<td>(17.5, 21.5, 9.0)</td>
<td>(17.625, 20.998, 8.456)</td>
<td>0.760</td>
<td>0.325</td>
</tr>
<tr>
<td>APSEN</td>
<td>(17.5, 21.5, 9.0)</td>
<td>(17.518, 20.846, 9.178)</td>
<td>0.678</td>
<td>0.625</td>
</tr>
<tr>
<td>GCGM-ARP</td>
<td>(17.5, 21.5, 9.0)</td>
<td>(17.893, 21.477, 8.818)</td>
<td>0.434</td>
<td>0.659</td>
</tr>
</tbody>
</table>

**Table 5.** Quantitative results of the in vivo experiment.
Fig. 6. The 3D view and slice view at the $Z = 9.0$ mm plane of reconstruction results obtained by different method for the *in vivo* experiment.
4. Discussion and conclusion

In this work, the EN regularization with adaptive parameters via GCGM was proposed to improve the recovery the 3D distribution of the fluorescent source. The regularization parameter was determined by the L-curve to select the optimal regularization parameter according to different data. Compared with the classical $L_1$-norm and $L_2$-norm, EN regularization offered greater flexibility to balance the sparsity and smoothness of the reconstruction source. Nevertheless, it also increased the computational complexity, so the GCGM-ARP with iteration was employed to enhance the accuracy and reduce the computational complexity. Specifically, GCGM-ARP splits the inverse problem into two sub-problems, namely gradient descent direction and step size, which were addressed by ISTA and ADMM. In addition, GCGM-ARP is simple and easy to implement and has convergence, which can guarantee the stability solution for FMT.

The effectiveness of GCGM-ARP was demonstrated through three groups of numerical simulation experiments and one *in vivo* experiment. In the numerical simulation experiments, the reconstruction results and quantitative analysis indicated that GCGM-ARP outperformed IS-$L_1$, IVTCG-$L_1$, N-EN, and APSEN method in terms of positioning accuracy, shape recovery, and dual-source positioning ability. Moreover, in the robustness test, GCGM-ARP showed high reconstruction accuracy and morphological recovery ability despite the increased ill-posedness of FMT reconstruction. The *in vivo* experiment further proved the superiority and practicability of the method. It was worth noting that because of the noise in the *in vivo* experiment and the error caused by organ segmentation, the results of reconstruction were worse than that of using the same method in numerical simulation, but GCGM-ARP still obtained satisfactory results. Overall, these experiments demonstrated the effectiveness of GCGM-ARP in improving reconstruction accuracy, spatial resolution, and dual-source resolution.

Although GCGM-ARP had achieved satisfactory reconstruction results, there are still some limitations. Firstly, the speed of ISTA algorithm used in GCGM-ARP can be further optimized by incorporating acceleration methods such as gradient projection to accelerate convergence. Secondly, the accuracy of optical parameters estimation can significantly affect the reconstruction results. Currently, the optical parameters used in this study were estimated based on the relevant tissues in the relevant literature [53], but using near infrared imaging method to measure the optical properties of tissue and background optical properties can improve the accuracy of reconstruction results. Additionally, in the *in vivo* experiment, the specific quantification of the effect of low-fluorescence in the non-target region on the captured fluorescence distribution has not been conducted. Further analysis of the effect of low-fluorescence can better simulate the complexity of the environment *in vivo* and enhance the reliability of *in vivo* experiments. Finally, The L-curve composed of a series of discrete points limits the accuracy of the optimal regularization parameters, thus affecting the reconstruction results. Therefore, developing a method for selecting more accurate optimal parameters under time constraint will be the focus of our work in the future.

In summary, among the four methods used in the experiments, the reconstruction results of IS-$L_1$ and IVTCG-$L_1$ are poor in location and morphology of the reconstructed source. The reconstruction results of N-EN method are show higher accuracy in location and shape recovery compared to the former two methods. The reconstruction performance of APSEN is similar to that of GCGM-ARP, and better than that of N-EN. However, GCGM-ARP achieves lower LE and RIE, as well as a higher Dice, which proves its better performance in localization, morphology and fluorescent intensity reconstruction. This work has the potential to promote the application of FMT in preclinical or clinical biology.

Funding. National Natural Science Foundation of China (61701403, 61806164); Key Research and Development Program of Shaanxi Province (2019GY215, 2021ZDLFS06-04); China Postdoctoral Science Foundation (2018M643719); Graduate Innovation Program of Northwest University (CX2023185).

Disclosures. The authors declare no conflicts of interest.
Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon request.

References


