



SODL-IR-FISTA: sparse online dictionary learning with iterative reduction FISTA for cone-beam X-ray luminescence computed tomography

XIN CAO,¹ WENLONG TANG,¹ HUIMIN GAO,¹ YIFAN WANG,¹ YI CHEN,^{2,4} CHENGYI GAO,³ FENGJUN ZHAO,¹ AND LINZHI SU^{1,5}

¹School of Information Science and Technology, Northwest University, Xi'an, Shaanxi 710127, China

²School of Electrical and Mechanical Engineering, The University of Adelaide, Adelaide SA 5005, Australia

³Department of Oncology, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi 710061, China

⁴yichen.cgz@gmail.com

⁵sulinzhi029@163.com

Abstract: Cone beam X-ray luminescence computed tomography (CB-XLCT) is an emerging imaging technique with potential for early 3D tumor detection. However, the reconstruction challenge due to low light absorption and high scattering in tissues makes it a difficult inverse problem. In this study, the online dictionary learning (ODL) method, combined with iterative reduction FISTA (IR-FISTA), has been utilized to achieve high-quality reconstruction. Our method integrates IR-FISTA for efficient and accurate sparse coding, followed by an online stochastic approximation for dictionary updates, effectively capturing the sparse features inherent to the problem. Additionally, a re-sparse step is introduced to enhance the sparsity of the solution, making it better suited for CB-XLCT reconstruction. Numerical simulations and in vivo experiments were conducted to assess the performance of the method. The SODL-IR-FISTA achieved the smallest location error of 0.325 mm in *in vivo* experiments, which is 58% and 45% of the IVTCG- L_1 (0.562 mm) and OMP- L_0 (0.721 mm), respectively. Additionally, it has the highest DICE similarity coefficient, which is 0.748. The results demonstrate that our approach outperforms traditional methods in terms of localization precision, shape restoration, robustness, and practicality in live subjects.

© 2024 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

With the continuous development of X-ray excitable nanophosphors, X-ray luminescence computed tomography (XLCT) has garnered significant attention due to its promising performance [1]. XLCT represents a hybrid imaging modality combining X-ray computed tomography (CT) with optical imaging [2]. High-resolution X-ray structural imaging with high sensitivity and high specificity optical molecular imaging can be naturally integrated in XLCT [3]. Compared to CT, XLCT is capable of acquiring molecular-level information, offering advantages in sensitivity and specificity. Compared with other optical molecular tomography (OMT) modalities, XLCT offers the following advantages. Firstly, in XLCT, the nanophosphors used can withstand multiple X-ray excitations and generate visible or near-infrared light (NIR) under X-ray irradiation. Its luminescence intensity can reach the single photon level, which is one order of magnitude higher than bioluminescence tomography (BLT) [4,5]. Secondly, compared with Cerenkov luminescence tomography (CLT) [6,7], XLCT avoids the potential impact of radioactive substances. Moreover, compared to fluorescence molecular tomography (FMT) [8,9], the energy attenuation of X-rays absorbed by biological tissues is minimal, which can avoid the interference of spontaneous fluorescence and background fluorescence, achieving high resolution and high signal-to-noise

ratio [10]. Therefore, XLCT has become a promising imaging technology that can be applied in basic research, drug development and clinical studies [11,12].

Since the introduction of the first XLCT imaging system by Xing et al., XLCT systems have been widely studied [13]. XLCT systems are categorized into pencil-beam/narrow-beam XLCT (PB/NB-XLCT) and cone-beam XLCT (CB-XLCT) based on the X-ray excitation and scanning methods [14]. PB/NB-XLCT emits X-rays in a strip shape and can achieve higher spatial resolution through translational scanning, but its data collection time is long and is not suitable for rapid biomedical applications [15]. CB-XLCT emits X-rays in a cone shape, eliminating the translational scanning process required for PB-XLCT and NB-XLCT, and significantly speeding up the imaging process at the expense of spatial resolution. Therefore, CB-XLCT is more suitable for fast biomedical applications, such as small animal whole-body imaging [16].

However, the CB-XLCT reconstruction is an ill-posed problem due to the complex excitation process and the high scattering of light during propagation in biological tissues [17]. To address these issues and enhance reconstruction quality, several methods have been proposed by researchers in recent years. For instance, some scholars have attempted to employ regularization methods to address this problem. Tzoumas et al. used L_1 -norm as convex relaxation of L_0 -norm to promote the sparsity of the solution [18]. Zhang et al. proposed a sparse nonconvex L_p ($0 < p < 1$) regularization method for efficient reconstruction [19]. Liu et al. proposed a method combining L_1 and total variation regularization (TV) to improve the imaging quality of CB-XLCT [15]. Gao et al. proposed a primordial dual Newton conjugate layer method based on restart strategy on the basis of L_1 norm to alleviate ill-posedness [14]. An elastic net- L_1L_2 method based on the combination of L_1 and L_2 regular terms was proposed for CB-XLCT reconstruction by Zhao et al. [17]. Beyond regularization, alternative approaches have been explored. Pu et al. demonstrated the use of compressed sensing (CS) techniques in CB-XLCT reconstruction to counteract the impact of sparse distribution of nanophosphors *in vivo* [20]. Guo et al. proposed the utilization of a mixed proton propagation model and Lasso-LSQR algorithm to enhance reconstruction accuracy [21]. Chen et al. proposed an algorithm based on dictionary learning and group structure strategy (DLGS) to effectively capture sparse features and enhance the interrelationship of elements in the same group [22]. In addition to the above methods, the reconstruction performance can also be improved by using optical properties, multi-spectral information and permission source regions [8,23,24].

In this paper, inspired by the application of dictionary learning in compressed sensing (CS), an online dictionary learning (ODL) method leveraging iterative reduction FISTA (IR-FISTA) is introduced to enhance the reconstruction quality of CB-XLCT. In the field of sparse representations, there are the two well-known computational approaches: Orthogonal Matching Pursuit (OMP) as a representative of greedy algorithms and FISTA as a representative of convex relaxations. Compared to OMP, which is used in the sparse coding stage of DLGS, the convex relaxation algorithms represented by FISTA are recommended for many applications due to their generality and reliability. Typically, they offer better accuracy, but they are relatively slower in terms of speed [25]. Therefore, we use IR-FISTA, which is much faster than FISTA, especially when the Lipschitz constant is large or the solution is highly sparse [26], aligning well with the demands of the CB-XLCT reconstruction problem. In the dictionary update stage, we adopt an online algorithm based on stochastic approximations, processing one sample at a time, updating dictionary by sequentially minimizing a quadratic local surrogate of the expected cost of the objective function, which makes the updating process more effective compared to global surrogate [27].

In order to evaluate the reconstruction performance of our method, a series of numerical simulations and simulation experiments were carried out. Our method was compared with L_0 -norm-based orthogonal matching pursuit (OMP- L_0) [28], and the incomplete variables truncated

conjugate gradient with L_1 -norm (IVTCG- L_1) [29] in terms of location accuracy, target shape, robustness, dual-source resolution and *in vivo* practicability.

The rest of this paper is structured as follows. Section 2 provides a detailed description of the CB-XLCT imaging model and the proposed method. In Section 3, the numerical simulation and *in vivo* experiments are designed, and the evaluation indexes are proposed. Section 4 summarizes the experimental results obtained using different reconstruction methods. Finally, Section 5 discusses and summarizes the main findings of this work.

2. Methodology

CB-XLCT imaging mainly includes two parts: forward problem and reverse reconstruction. The forward problem involves calculating the optical signal distribution. In this process, the light emitted by the target light source and its transmission to the surface of the organism are considered. Knowledge of the location of the target light source, the anatomical structure of the biological tissue, and the optical parameters is required. A photon propagation model is utilized to describe this process. Based on the forward problem, the optical information of the biological surface is obtained, and combined with specific reconstruction algorithms, the position information and other quantitative information of the target are reconstructed as reverse reconstruction process of CB-XLCT.

2.1. Photon propagation model

According to the customized optical/CT dual modality imaging system, X-rays interact with the target to produce photons, and the excitable nanophosphors will produce visible light or NIR light after being stimulated by X-ray, the source energy density $S(\mathbf{r})$ (W cm^{-3}) can be calculated as follows [30]:

$$S(\mathbf{r}) = \varepsilon X(\mathbf{r})\rho(\mathbf{r}) \quad (1)$$

where $X(\mathbf{r})$ is the X-ray intensity (W cm^{-3}) at position \mathbf{r} , and $\rho(\mathbf{r})$ is the density of the nanophosphor at position \mathbf{r} , ε is the light yield of nanoparticles (mg/mL). In CB-XLCT, a cone-beam incoherent X-ray source passing through biological tissues and exciting nanophosphors will produce photoelectric effect, which will make part of the X-ray electron flow absorbed or scattered. According to Beer-Lambert's law, when X-rays pass through tissues, the X-ray intensity distribution (energy attenuation) can be expressed as [31]:

$$X(\mathbf{r}) = X(\mathbf{r}_0) \exp \left\{ - \int_{\mathbf{r}_0}^{\mathbf{r}} \mu_t(\tau) d\tau \right\} \quad (2)$$

where $X(\mathbf{r}_0)$ is the intensity of the X-ray light source at the initial position \mathbf{r}_0 , and $\mu_t(\tau)$ is the X-ray attenuation coefficient at position τ , which can be obtained with attenuation-based CT technology. Then we analyze the transmission of emitted photons, the transmission of visible light or NIR light in biological tissue is often modeled by radiative transfer equation (RTE) [5]. However, due to the high computational cost of RTE, in general, the transmission of light in tissues with high scattering and low absorption can be simplified as the following steady-state diffusion equation (DE) model with Robin-type boundary condition, which is defined as follows [32,33]:

$$\begin{cases} -\nabla \cdot [D(\mathbf{r})\nabla\Phi(\mathbf{r})] + \mu_a(\mathbf{r})\Phi(\mathbf{r}) = S(\mathbf{r}) & (\mathbf{r} \in \Omega) \\ \Phi(\mathbf{r}) + 2k(\mathbf{r}, n, n')D(\mathbf{r})[v(\mathbf{r}) \cdot \nabla\Phi(\mathbf{r})] = 0 & (\mathbf{r} \in \partial\Omega) \end{cases} \quad (3)$$

where $D(\mathbf{r}) = [3(\mu_a(\mathbf{r}) + (1-g)\mu'_s(\mathbf{r}))]^{-1}$, which is the diffusion coefficient, and $\mu_a(\mathbf{r})$ and $\mu'_s(\mathbf{r})$ are the absorption and reduced scattering coefficients at the position \mathbf{r} , respectively; $\Phi(\mathbf{r})$ is the photon flux, and Ω is the domain of the imaged object and $\partial\Omega$ is the boundary of Ω ; $v(\mathbf{r})$ is

the outward unit normal vector on $\partial\Omega$, and $k(\mathbf{r}, n, n')$ is the boundary mismatch factor, which depends on the refractive index n in Ω and n' in the surrounding medium.

Based on the finite element method (FEM) [34], Eq. (3) can be transformed to obtain the following matrix equation:

$$\mathbf{M} \cdot \Phi = \mathbf{F} \cdot \boldsymbol{\varepsilon} \cdot X \cdot \rho \quad (4)$$

with

$$\begin{aligned} M_{i,j} &= \int_{\Omega} (D\nabla\psi_i \cdot \nabla\psi_j + \mu_a\psi_i\psi_j) dr + \frac{1}{2k} \int_{\partial\Omega} \psi_i\psi_j dr \\ F_{i,j} &= \int_{\Omega} \Phi(r)\psi_i\psi_j dr \\ X_{i,j} &= X_{i,j}(r) \end{aligned} \quad (5)$$

where ρ is the nanophosphors density to be reconstructed; $M_{i,j}$ and $F_{i,j}$ are the elements of matrix \mathbf{M} and \mathbf{F} , respectively; ψ_i and ψ_j denote the corresponding elements of the test function; $X_{i,j}(r)$ is the X-ray intensity for each vertex.

Since the matrix \mathbf{M} is positive definite, Eq. (4) can be rewritten as:

$$\Phi = \mathbf{G} \cdot \rho \quad (6)$$

where $\mathbf{G} = (\mathbf{M}^{-1}\mathbf{F}) \cdot \boldsymbol{\varepsilon} \cdot X$ is the weight matrix of size $m \times n$, and Φ is the photon measurements vector on the object surface, and its size is $m \times 1$. In the practical application of CB-XLCT, the noise of the XLCT imaging system needs to be considered. Equation (6) can be written as:

$$b = \Phi + \sigma = \mathbf{G} \cdot x + \sigma \quad (7)$$

where b represents the photon measurements vector with noise, σ is the system noise, and x is the unknown nanophosphors distribution in the imaging object.

2.2. Reconstruction based on the SODL-IR-FISTA method

The objective of CB-XLCT reconstruction is to estimate the distribution of nanophosphor x from the measured photon density Φ . However, due to the severe scattering of visible and NIR light in biological tissue, the system matrix \mathbf{G} becomes ill-conditioned, making it unrealistic to solve Eq. (6) directly. Considering the distribution of nanophosphors is small and sparse [35], to obtain a stable and robust solution, the L_0 regularization term can be used to find the solution by minimizing the following objective function [19]:

$$\underset{x \geq 0}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{G}x - b\|_2^2 \quad \text{s.t.} \quad x_0 \leq K \quad (8)$$

where K is the sparsity level of the reconstruction result. The CB-XLCT reconstruction problem is transformed into an optimization problem of x in Eq. (8).

The SODL-IR-FISTA method uses the dictionary (matrix \mathbf{G} can be regarded as a redundant dictionary) to continuously learn the sparse representation of the data, and by updating the dictionary, the updated dictionary can better represent the sparse data. Specifically, the solution is compartmentalized into three stages: sparse coding stage, dictionary update stage and re-sparse stage. The IR-FISTA is used in the sparse coding stage, which accelerates the solution under the premise of ensuring accuracy. In the dictionary update stage, an online algorithm based on stochastic approximation is used to further capture sparse features. Considering that the distribution of nanophosphor usually aggregates, the clustering algorithm in statistics is used for reference in the re-sparse stage to identify and eliminate outliers, as well as to control the sparsity of the solution, which can make the solution more accurate. The general flow of SODL-IR-FISTA is summarized in Algorithm 1.

2.2.1. Sparse coding stage

The sparse coding stage aims to address the objective function described in Eq. (8). However, due to its NP-hard nature, achieving a stable solution through conventional optimization methods is challenging. The convex relaxation algorithm represented by FISTA finds an approximation solution by transforming Eq. (8) into the following equation:

$$\underset{x \geq 0}{\operatorname{argmin}} \frac{1}{2} \|Dx - b\|_2^2 + \lambda \|x\|_1 \quad (9)$$

where D is the learned dictionary (initially $D = G$) and λ is the regularization parameter, the above equation is also known as the LASSO problem.

Algorithm 1. SODL-IR-FISTA for CB-XLCT reconstruction

Input: system matrix G , photon measurement Φ , regularization parameter λ , target number

Initialization: initial dictionary $D_0 = G$, maximum iteration number $\text{maxIter} = 500$, current iteration $k = 1$, the “past” information A_0 and B_0

While $k < \text{maxIter}$ **do**

 1): Sparse coding stage

 Compute x_k using D_{k-1} by Algorithm 2.

$$A_k \leftarrow A_{k-1} + x_k x_k^T$$

$$B_k \leftarrow B_{k-1} + b x_k^T$$

 2): Dictionary update stage

 Compute D_k using Algorithm 3.

$$x_{out} = x_{out} + x_k$$

 3): $k = k + 1$

End while

4): Re-sparse stage

Using outlier detection algorithm to get outlier in x_{out} .

If target > 1

 Using clustering algorithm, get grouping information from x_{out} .

End if

Update x_{out} , eliminate outliers and select appropriate groups based on grouping information.

Output: x_{out}

FISTA is celebrated for its simplicity and rapid convergence properties, with convergence speed decreasing quadratically with the number of the iterations. Although FISTA only needs to perform matrix-vector multiplications twice in each iteration, the computational cost of matrix-vector multiplications escalates significantly in large-scale problems involving an ill-conditioned matrix. This is because a large condition number of matrix would result in a large number of iterations. Consequently, FISTA may be inefficient for large-scale problems [36]. Note that in the iterations of FISTA, a significant portion of the components of x typically remain zero following each iteration, particularly as the iterates approach the solution. This means that updating these components is “wasteful”. Based on this observation, Wang et al. designed an iterative reduction algorithm [26], as shown below.

Using the current Karush-Kuhn-Tucker (KKT) information [37] to estimate the working set W^k of the current iteration.

Let

$$\begin{aligned} W_1(x^k) &= \{j | x_j^k \neq 0\} \\ W_2(x^k) &= \left\{ j | x_j^k = 0 \text{ and } \left| A_j^T(Ax^k - b) - \lambda \alpha \frac{x_j^k}{\|x\|} \right| > (1 - \tau)\lambda \right\} \\ W(x^k) &= W_1(x^k) \cup W_2(x^k) \end{aligned} \quad (10)$$

where τ is a small positive number. Moreover, let W_1^k, W_2^k, W^k represent $W_1(x^k), W_2(x^k), W(x^k)$, respectively.

Even though we perform a proximal iteration solely on the working set W^k , the computational expense of identifying this set equals that of performing a proximal iteration on a complete index set. Therefore, an effective approach involves executing multiple proximal iterations subsequent to identifying the working set. Due to this, after obtaining W^k , we solve the following simplified subproblem:

$$x_{W^k}^{k+1} = \underset{z \in R^{|W^k|}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|A_{W^k} z - b\|^2 + \lambda \|z\|_1 \right\} \quad (11)$$

and let $x_{(W^k)^c}^{k+1} = 0$, where $(W^k)^c = (W(x^k))^c := \{1, 2, \dots, n\} \setminus W(x^k)$. Hence, we have the procedures of iterative reduction FISTA as Algorithm 2.

Algorithm 2. Iterative reduction FISTA (IR-FISTA)

Input: matrix $A = D_{k-1}$, vector b , regularization parameter λ

Initialization: $\tau > 0, x^0 = 0$

For $k = 1, 2, 3, \dots$ **do**

$$W_1^k = \{j | x_j^k \neq 0\}$$

$$W_2^k = \left\{ j | x_j^k = 0 \text{ and } \left| A_j^T(Ax^k - b) - \lambda \alpha \frac{x_j^k}{\|x\|} \right| > (1 - \tau)\lambda \right\}$$

$$W^k = W_1^k \cup W_2^k$$

$$L_k = \|A_{W^k}^T A_{W^k}\|$$

Approximately solve the LASSO subproblems (11) with FISTA (see Algorithm 3).

Let

$$x_{W^k}^{k+1} = z^k, x_{(W^k)^c}^{k+1} = 0$$

if $\|g(f, x^k)\| < \varepsilon_{tol}$ **then**

Break;

End for

Output: $x^* = x^{k+1}$

Moreover, let

$$g_j(f, x) = \begin{cases} \max \left\{ \left| \frac{\partial f}{\partial x_j}(x) \right| - \lambda, \min \left\{ \left| \frac{\partial f}{\partial x_j}(x) + \lambda \operatorname{sign}(x_j) \right|, |x_j| \right\} \right\}, & x_j \neq 0, \\ \max \left\{ 0, \left| \frac{\partial f}{\partial x_j}(x) \right| - \lambda \right\}, & x_j = 0. \end{cases} \quad (12)$$

$g(f, x)$ is used to represent the KKT residual of the LASSO problem, while $g(f, x^*)$ can be used to measure the accuracy of the solution x^* .

For the FISTA algorithm, let

$$Q_L(x, y) = f(y) + \langle x - y, \nabla f(y) \rangle + \Omega(x) + \frac{L}{2} \|x - y\|^2 \quad (13)$$

where $\Omega(x) = \lambda x_1$. And let

$$\begin{aligned} \mathcal{P}_L(y^t) &= \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} Q(x, y^t) \\ &:= \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \frac{L}{2} \left\| x - \left(y^t - \frac{1}{L} A^T (Ay^t - b) \right) \right\|^2 + \lambda \|x\|_1 \end{aligned} \quad (14)$$

The specific process of FISTA algorithm can be found in Algorithm 3.

Algorithm 3. FISTA algorithm

Input: $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, \lambda > 0, x^0 \in \mathbb{R}^n, L_0 > 0, \eta > 1$

Initialization: $y^1 = x^0, \theta_1 = 1$

For $t = 1, 2, 3, \dots$ **do**

Find the smallest nonnegative integers i_t such that with $\bar{L} = \eta^{i_t} L_{t-1}$

$$\mathcal{P}_{\bar{L}}(y^t) \leq Q_{\bar{L}}(\mathcal{P}_{\bar{L}}(y^t), y^t)$$

Set $L_t = \bar{L}$ and compute
 $x^t = \mathcal{P}_{L_t}(y^t)$

$$\theta_{t+1} = \frac{1 + \sqrt{1 + 4\theta_t^2}}{2}$$

$$y^{t+1} = x^t + \frac{\theta_t - 1}{\theta_{t+1}} (x^t - x^{t-1})$$

End for

Output: $x^* = x^t$

2.2.2. Dictionary update stage

During the dictionary update stage, our method employs a block-coordinate descent approach coupled with warm restarts. One of its primary benefits is its lack of parameters, eliminating the need for any tuning of learning rates.

The new dictionary is obtained by the last updated dictionary and the obtained solution as warm restart, so that:

$$\begin{aligned} \mathbf{D}_k &\triangleq \underset{\mathbf{D} \in C}{\operatorname{argmin}} \frac{1}{k} \sum_{i=1}^k \left(\frac{1}{2} \|b - \mathbf{D}x_i\|_2^2 + \lambda \|x_i\|_1 \right) \\ &= \underset{\mathbf{D} \in C}{\operatorname{argmin}} \frac{1}{k} \left(\frac{1}{2} \operatorname{Tr}(\mathbf{D}^T \mathbf{D} A_k) - \operatorname{Tr}(\mathbf{D}^T \mathbf{B}_k) \right) \end{aligned} \quad (15)$$

Furthermore, the procedure necessitates storing only the matrices in $\mathbb{R}^{k \times k}$ and in $\mathbb{R}^{m \times k}$, not all the vectors x_k , where \mathbb{R} is the set of real numbers, \mathbb{R}^N is the vector space composed of all N -dimensional column vectors and $\mathbb{R}^{M \times N}$ is the vector space composed of all matrices of size $M \times N$. Concretely, in the dictionary update stage, this method sequentially updates each column of \mathbf{D} . In practice, the vectors x_k are sparse and the coefficients of the matrix A_k are often concentrated on the diagonal, which makes the block-coordinate descent more efficient [27]. Algorithm 4 describes the flow of the dictionary update algorithm.

2.2.3. Re-sparse stage

At this stage, the primary objective is to re-sparse the obtained solution and filter out the points that fail to meet the specified criteria. In most cases, the target data points should be closer to

each other in space, and the points that do not meet the requirements may be far away from other points. Inspired by data mining, a re-sparse method is designed.

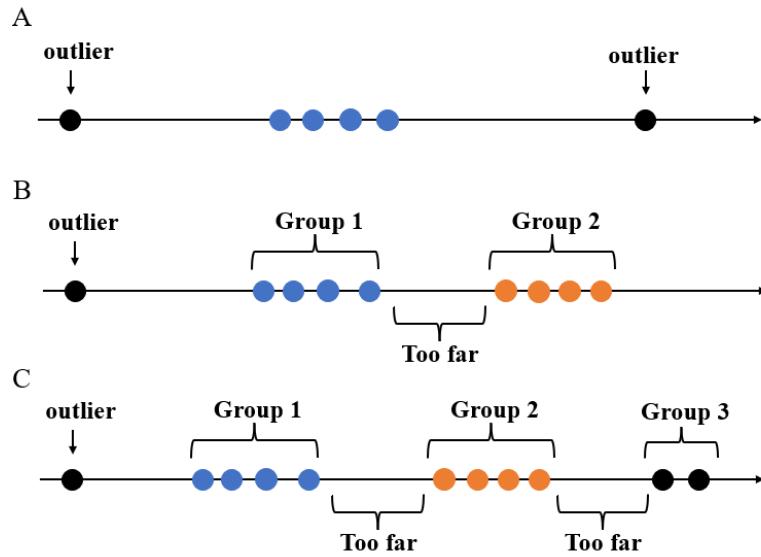


Fig. 1. Schematic diagram of re-sparse stage. The straight arrow represents a vector with many points on it, the circle indicates the point where the element value is greater than zero, and the black circle indicates represents outliers. (A) When target number is 1, only outlier detection is performed; (B) When the target number is greater than 1, perform clustering operation after outlier detection. The blue and orange circles form Group 1 and Group 2 respectively. (C) When the target number is greater than 1, select the appropriate group according to the target number. Here, takes the target number equal to 2 as an example, and classify Group 3 as outliers.

We use the distance-based outlier detection algorithm for reference. By calculating the average distance from each point to other points and determining the deviation points according to a threshold, we can identify potential outliers. When there are multiple light sources, the non-zero data points in the solution are divided into several specific groups by distance-based clustering algorithm. This clustering ensures that data points within the same group exhibit similar spatial attributes, while those in different groups exhibit different attributes. Specifically, after outlier detection, for the remaining non-zero elements, the distance between adjacent elements is compared to determine whether they belong to the same group. If the distance between two adjacent points is less than the set value, they are considered to belong to the same group, otherwise they are not considered to belong to the same group. After the clustering process, several groups are selected as the basis for multi-target reconstruction according to the target number, and the elements of other groups are treated as outliers for subsequent elimination. Three possible scenarios of re-sparse stage are shown in Fig. 1.

Algorithm 4. Dictionary update

Input: $D = [d_1, \dots, d_n] \in \mathbb{R}^{m \times n}$, $A = [a_1, \dots, a_n] \in \mathbb{R}^{n \times n}$, $B = [b_1, \dots, b_n] \in \mathbb{R}^{m \times n}$

Repeat

For $j = 1$ to n **do**

 Update the $j - th$ column to optimize for (15):

$$\begin{aligned} u_j &\leftarrow \frac{1}{A[j,j]}(b_j - Da_j) + d_j \\ d_j &\leftarrow \frac{1}{\max(\|u_j\|_{2,1})}u_j \end{aligned}$$

End for

 Until convergence

Output: D (updated dictionary)

3. Experiment design

In order to evaluate the effectiveness of the proposed method in CB-XLCT reconstruction, several numerical simulations and *in vivo* experiment were designed. Specifically, in terms of location accuracy, morphological recovery, robustness and *in vivo* practicability, two existing algorithms: OMP- L_0 and IVTCG- L_1 were compared. All experiments and procedures were conducted utilizing a laptop equipped with Intel Core i5-9300 H CPU (2.40 GHz) and 8GB RAM.

3.1. Numerical simulation setup

The heterogeneous phantom model was used for numerical simulation experiments. The model was a cylinder with a simulation height of 30 mm and a radius of 10 mm, including five organs and tissues: heart, liver, lung, muscle and bone. The 3D view of the model is depicted in Fig. 2(A). Table 1 displays the optical parameters of the tissues and organs at a wavelength of 650 nm within the model [22].

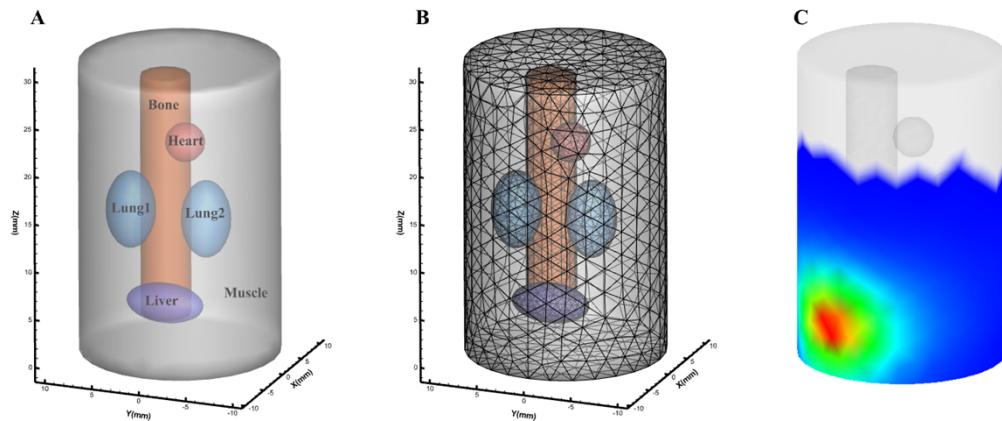


Fig. 2. Heterogeneous cylindrical phantom for numerical simulation studies: (A) The 3D view of the cylindrical phantom. (B) The 3D view of meshed cylindrical phantom. (C) The forward simulation result of single-target.

In the inverse reconstruction, COMSOL Multiphysics 5.6 (COMSOL, Inc., Burlington, Massachusetts) software [38] is used to discretize the model into a finite element mesh of 4626 nodes and 25840 tetrahedral elements, as shown in Fig. 2(B). Using molecular optical simulation

Table 1. Optical properties of the heterogeneous cylindrical phantom at 650 nm

Tissue	$\mu_a(r)[\text{mm}^{-1}]$	$\mu_s(r)[\text{mm}^{-1}]$	g
Muscle	0.010	4.000	0.900
Bone	0.002	20.000	0.900
Heart	0.200	16.000	0.850
Liver	0.035	6.000	0.900
Lungs	0.350	23.000	0.900

environment (MOSE, Version 2.3) software [39] implements the Monte Carlo (MC) method to obtain the forward simulation results of single-target and dual-target. The forward simulation result of single-target is shown in Fig. 2(C).

Three groups of experiments were set up in the model. In the single-target simulation experiment, a uniform sphere with a radius of 1 mm and a center of (-4.0, 5.0, 5.0) mm was used to simulate early tumors. In the dual-target experiment, two identical spheres with a radius of 1 mm were used at (-6.0, 3.0, 8.0) mm and (-6.0, 3.0, 16.0) mm to simulate tumors. In addition, a series of anti-noise simulation experiments were carried out to evaluate the robustness of the method. Gaussian noise at level of 5%, 10%, 15%, 20%, 25%, 30% was added in the single-source simulation experiment, and the reconstruction results under different Gaussian noise ratios were observed.

3.2. Implanted experiment

In order to evaluate the feasibility of SODL-IR-FISTA *in vivo* CB-XLCT, we collected data from female BALB/C nude mouse using an XLCT imaging system. All experimental procedures were approved by the Animal Ethics Committee of Northwest University, China. Specifically, we implanted a flexible plastic tube (with a radius of 1 mm and a height of 2 mm) containing $\text{Gd}_2\text{O}_2\text{S}: \text{Tb}$ nanophosphors (about 20 μl) into the mouse body and placed it at (15.2, 16.5, 20.0) mm to simulate an early tumor. The optical parameters of the tissues at a wavelength of 620 nm within our experiment are shown in Table 2. The imaging system included a controller, a CMOS X-ray detector panel, a cone beam X-ray source, an electron-multiplying CCD (EMCCD) camera and a rotation stage, as shown in Fig. 3. The cone beam X-ray source (Apogee, Oxford Instruments, USA) excited nanophosphors in biological tissues with a voltage of 50kVp and current of 1 mA. The EMCCD camera was placed perpendicular to the X-ray source, and the optical data with a wavelength of 620 nm were captured in a manner with an integration time and binning of 30 s and 2×2 , respectively. To obtain 3D volume data, the mouse was fixed on a rotating stage and rotated 360° at 1° intervals to obtain micro-CT images and X-ray projection images. And the CMOS X-ray detector panel was used to detect transmitted X-rays, and then obtained high-resolution CT imaging. Subsequently, the Feldkamp-David-Kress (FDK) method [40] was used to perform 3D reconstruction of the XCT projection, and the forward measurement data of the nanophosphor was mapped onto the surface. The mouse model was discretized into 8935 nodes and 49545 tetrahedral elements for reconstruction, as Fig. 4 illustrates.

3.3. Evaluation indexes

In order to quantify and analyze the positioning accuracy and shape recovery performance of different methods in CB-XLCT reconstruction, two commonly used evaluation indexes are adopted: location error (LE) [41] and Dice similarity coefficient (DICE) [42].

LE is the Euclidean distance between the reconstructed region center (x_r, y_r, z_r) and the actual region center (x_a, y_a, z_a). It is a quantitative measure to evaluate the positioning accuracy of the reconstruction result. The LE value is always greater than 0, and the smaller the LE value, the

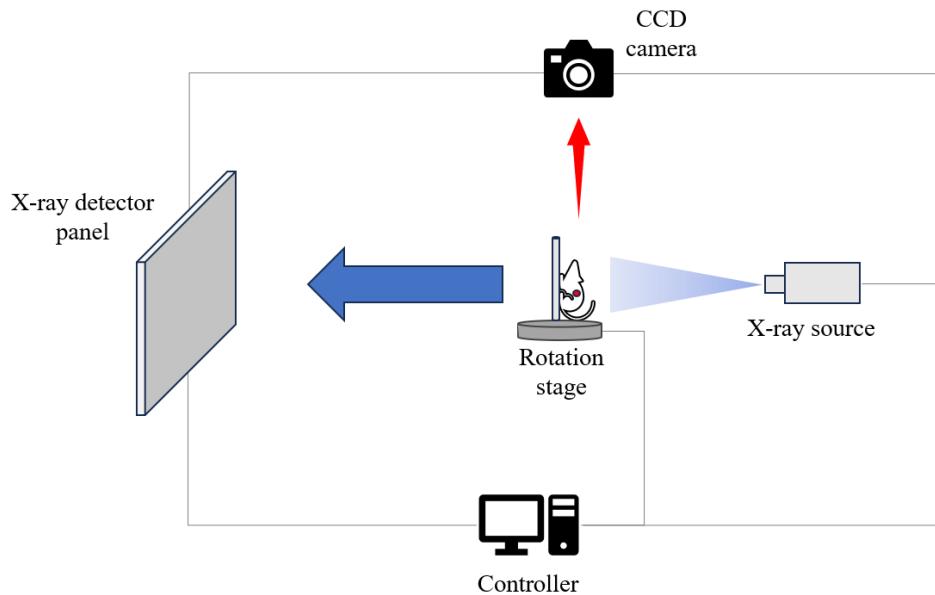


Fig. 3. The schematic diagram of the CB-XLCT system.

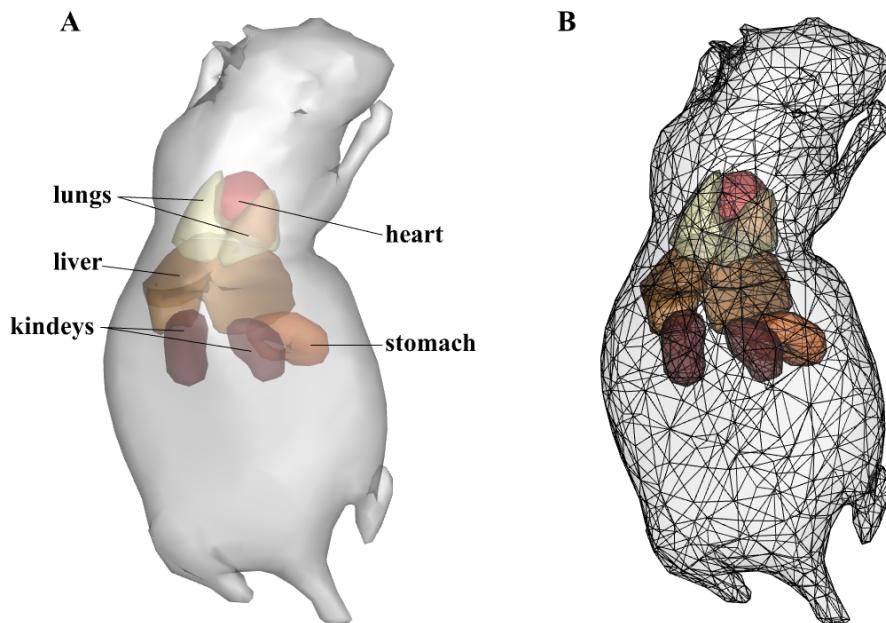


Fig. 4. (A) The 3D view of the mouse. (B) The 3D view of meshed biological tissues.

Table 2. Optical properties of the mice model at 620 nm

Tissue	$\mu_a(r)[mm^{-1}]$	$\mu_s(r)[mm^{-1}]$	g
Muscle	0.016	5.100	0.900
Kidneys	0.012	24.720	0.900
Heart	0.200	16.000	0.860
Liver	0.035	6.000	0.900
Lungs	0.350	23.000	0.900

higher the positioning accuracy of the reconstruction.

$$LE = \sqrt{(x_r - x_a)^2 + (y_r - y_a)^2 + (z_r - z_a)^2} \quad (16)$$

DICE measures the overlap between the reconstructed target region X_r and the actual target region X_a . The DICE value ranges from 0 to 1. The larger the value, the greater the similarity between the reconstructed target region and the actual target region, indicating that the higher the morphological recovery ability of the algorithm.

$$DICE = \frac{2|X_r \cap X_a|}{|X_r| + |X_a|} \quad (17)$$

4. Results

4.1. Numerical simulations results

4.1.1. Single-target experiment

The results of the single light source reconstruction experiment are shown in Fig. 5. The first row displayed the 3D view of the reconstruction results of each method, and the reconstructed source is visualized as a green area. In the section view, the yellow-green area represents a sectional view of the reconstruction results of the single-target at the $Z = 5$ mm plane, where the white circle represents the actual position and edge of the real target. Moreover, Table 3 displays the quantitative evaluation results of the three reconstruction methods. It can be seen that our method obtains a lower LE compared to other methods, which indicates that the reconstruction center of the SODL-IR-FISTA is closer to the true position. In addition, the DICE value of the SODL-IR-FISTA is higher, indicating that our method reconstructs the target region with the largest overlap with the true target region and has a better effect in morphological restoration.

Table 3. Quantitative results of different methods in single-target reconstruction experiment

Method	Reconstructed region center (mm)	LE (mm)	DICE
OMP- L_0	(-3.986, 4.531, 5.744)	0.880	0.573
IVTCG- L_1	(-3.776, 4.672, 5.099)	0.409	0.608
SODL-IR-FISTA	(-4.119, 5.166, 4.770)	0.308	0.717

4.1.2. Dual-target experiment

The reconstruction results of the dual-target experiment by the three methods are shown in Fig. 6. The representation of both the reconstructed and actual targets follows the same format as in the single-target experiment. It can be seen from 3D view and section view ($Y = 3$ mm) that the reconstruction results of SODL-IR-FISTA are the most accurate and the characterization of the actual target is better. The quantitative analysis of the reconstruction results of the three methods is shown in Table 4. The results indicate that, in comparison to other two methods, the LE of

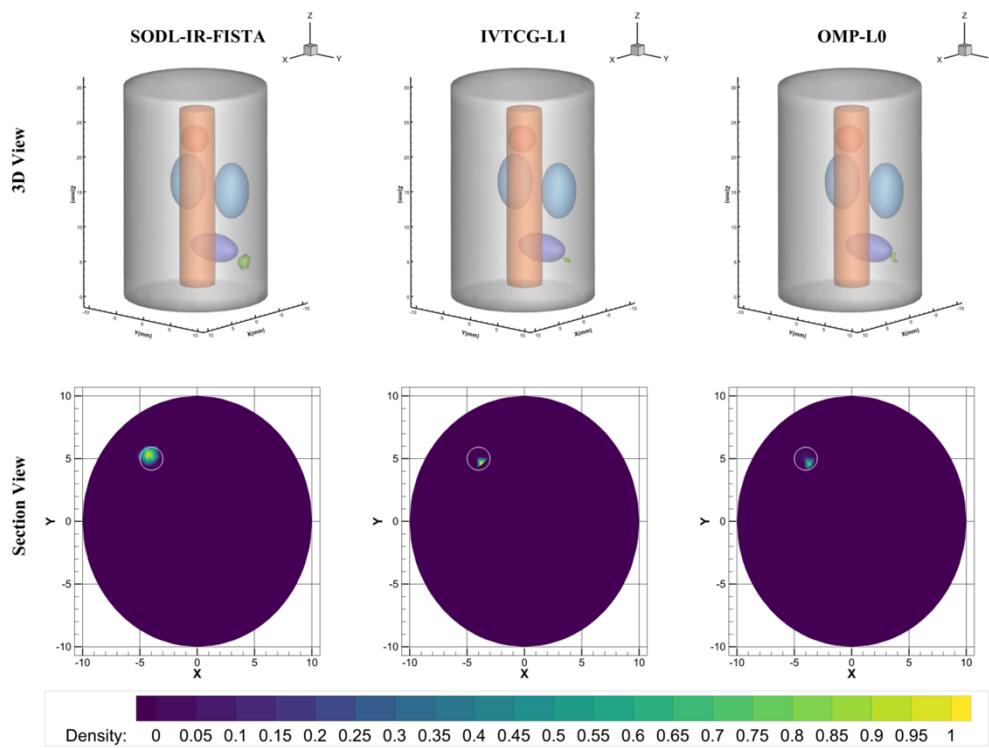


Fig. 5. Reconstruction results of three methods for a single spherical target.

each target result reconstructed by the SODL-IR-FISTA method is the smallest and the value of DICE is the largest. This shows that in the case of multiple targets, SODL-IR-FISTA still has good positioning and morphological reconstruction capabilities, and can accurately distinguish different targets.

Table 4. Quantitative results of different methods in dual-target reconstruction experiment

Method	Reconstructed region center (mm)	LE (mm)	Total LE (mm)	DICE	Average DICE
OMP- L_0	(-5.743, 2.953, 8.793)	0.835	1.355	0.509	0.549
	(-5.551, 2.737, 16.005)	0.520		0.589	
IVTCG- L_1	(-5.664, 2.782, 8.578)	0.703	1.124	0.571	0.574
	(-5.905, 2.792, 16.353)	0.421		0.576	
SODL-IR-FISTA	(-5.678, 2.861, 8.388)	0.523	0.825	0.672	0.729
	(-5.722, 2.922, 16.085)	0.302		0.786	

4.1.3. Anti-noise experiment

The results of the anti-noise experiment are shown in Fig. 7. Based on the single-target experiment, 5%, 10%, 15%, 20%, 25%, and 30% Gaussian noise are sequentially added to the measurement data to observe their influence on the reconstruction results. It can be observed from the figure that the fluctuations of LE and DICE are small, indicating that our method can maintain good reconstruction performance under the condition of significant noise, which proves that the method has strong robustness.

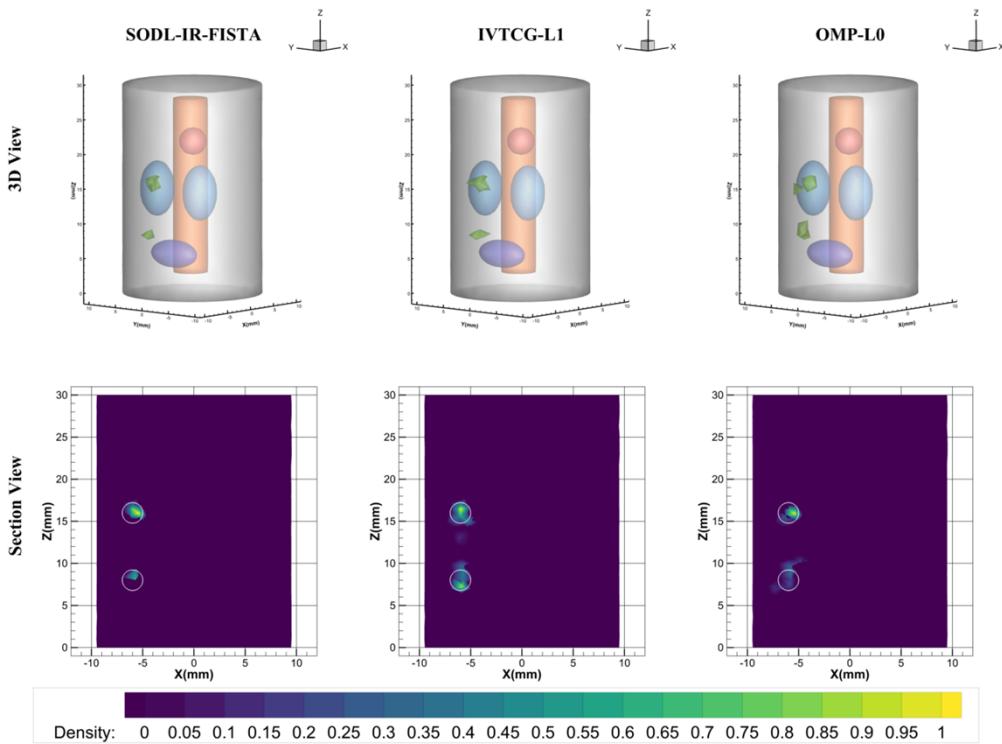


Fig. 6. Reconstruction results of four methods for dual-target.

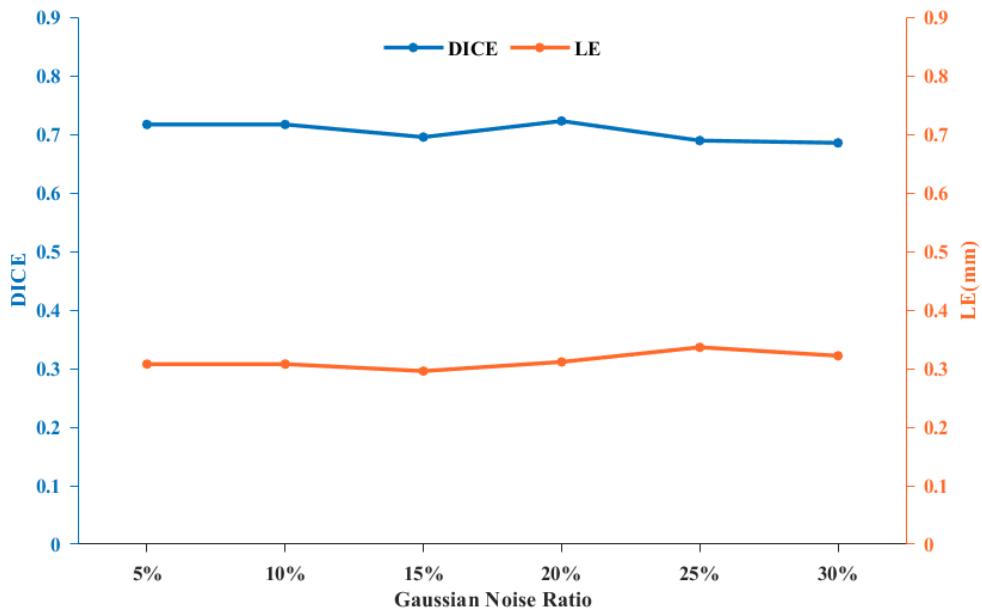


Fig. 7. Illustration of LE and DICE at different noise levels.

4.2. *In vivo* experiment results

The results of the *in vivo* experiment are shown in Fig. 8, the 3D view is expressed in the same way as the above experiments. In the section view, the red-yellow area is the reconstruction result, and the gray-white circular area represents the section of the real source. Observations from 3D view and section view ($Z = 20$ mm) indicate that the SODL-IR-FISTA method provides the optimal reconstruction effect. Concurrently, in order to quantitatively analyze the reconstruction results of different methods, LE and DICE were calculated and summarized in Table 5. From Table 5, the LE obtained by OMP- L_0 is 0.721 mm, which is more than twice that of SODL-IR-FISTA. The LE obtained by IVTCG- L_1 is 0.562 mm, while the SODL-IR-FISTA is only 0.325 mm. This indicates that the reconstruction result center of our method is closer to the actual position, which can also be confirmed by observing Fig. 8. And, compared with OMP- L_0 and IVTCG- L_1 , SODL-IR-FISTA has better performance in morphological recovery. The DICE of the reconstructed results is the highest among the three methods, with a value of 0.748, while the DICE of OMP- L_0 and IVTCG- L_1 are 0.392 and 0.380, respectively. These results indicate that, the spatial overlap between the reconstruction results of SODL-IR-FISTA and the real source is the highest, which means that its shape recovery is more accurate.

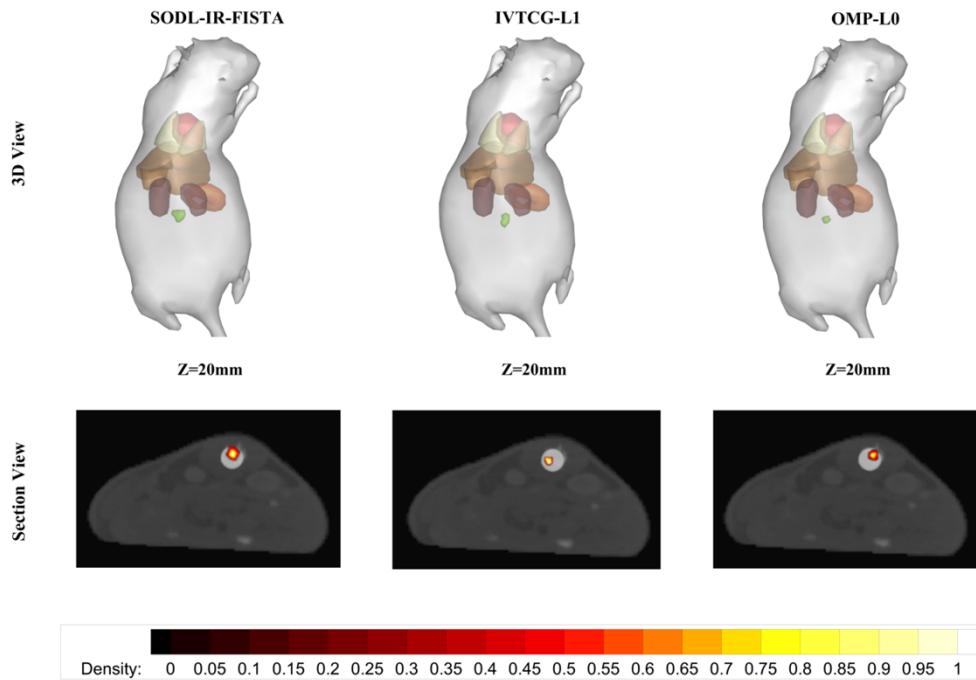


Fig. 8. Reconstruction results of the *in vivo* experiment.

Table 5. Quantitative results of *in vivo* experiment.

Method	Reconstructed region center (mm)	LE (mm)	DICE
OMP- L_0	(15.438, 16.821, 19.430)	0.721	0.392
IVTCG- L_1	(14.980, 16.537, 19.484)	0.562	0.380
SODL-IR-FISTA	(15.242, 16.821, 20.031)	0.325	0.748

5. Discussion and conclusion

CB-XLCT uses a cone-beam geometry to cover the entire imaging object, which can simplify the scanning process and shorten the imaging time. Therefore, it is more suitable for biological applications, especially fast biomedical imaging. Due to the diffusion propagation and high scattering of photons in biological tissues, the reconstruction process of CB-XLCT is ill-posed, resulting in inferior image quality. To solve this problem, in this paper, a reconstruction method called SODL-IR-FISTA is proposed. Since the distribution of nanophosphors in the body is small and sparse relative to the overall imaging object, this method utilizes the feature of dictionary learning in continuously learning sparse representations of data, effectively capturing sparse features of data and achieving sparse reconstruction of CB-XLCT. The SODL-IR-FISTA consists of three stages: sparse encoding, dictionary update, and re-sparse. The sparse encoding stage adopts the IR-FISTA algorithm, enhancing computational speed without compromising accuracy. The dictionary updating stage uses an online algorithm based on stochastic approximation to update the dictionary efficiently. In the re-sparse stage, an algorithm to remove outliers was designed to obtain more sparse and accurate solutions.

To verify the effectiveness of the proposed method, a series of numerical simulations and *in vivo* experiments were designed and carried out, and the results were compared with two common reconstruction methods, OMP- L_0 and IVTCG- L_1 . The experimental results are as follows. Firstly, the experimental results of single-target and dual-target show that SODL-IR-FISTA has good performance in reconstruction positioning accuracy and morphological restoration. Subsequently, the anti-noise experiments show that SODL-IR-FISTA has good robustness. Ultimately, the *in vivo* experiments and their results show the feasibility of SODL-IR-FISTA in biomedical research and the potential for CB-XLCT reconstruction in preclinical research.

However, despite the excellent performance of SODL-IR-FISTA in CB-XLCT reconstruction, there are still some challenges that need to be addressed. Firstly, during the dictionary update stage, the online update algorithm is not fast enough, which affects the overall efficiency of the method. Secondly, although most of the algorithm parameters do not require excessive adjustment, the target number and threshold values still need to be manually determined and cannot be automatically adjusted. Finally, more prior information can be used to obtain better results, such as group sparsity or smoothing prior.

In summary, this study proposes SODL-IR-FISTA method based on IR-FISTA and online dictionary learning to improve the reconstruction effect of CB-XLCT. The experimental results show that compared with other typical regularization-based methods, this method provides better reconstruction quality in terms of positioning accuracy and shape restoration. In addition, this method has strong robustness. Therefore, we believe that this method has great potential for preclinical application of CB-XLCT and can promote its *in vivo* application.

Funding. National Major Scientific Research Instrument Development Projects of China (82127805); Key Research and Development Program of Shaanxi Province (2019GY215, 2021ZDSF06-04, 2024SF-YBXM-681); National Natural Science Foundation of China (61701403, 61806164).

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

1. C. Carpenter, C. Sun, G. Pratx, *et al.*, “Hybrid x-ray/optical luminescence imaging: characterization of experimental conditions,” *Med. Phys.* **37**(8), 4011–4018 (2010).
2. M. Ahmad, G. Pratx, M. Bazalova, *et al.*, “X-ray luminescence and x-ray fluorescence computed tomography: new molecular imaging modalities,” *IEEE Access* **2**, 1051–1061 (2014).
3. H. Zhang, X. Huang, M. Zhou, *et al.*, “Adaptive shrinking reconstruction framework for cone-beam X-ray luminescence computed tomography,” *Biomed. Opt. Express* **11**(7), 3717–3732 (2020).
4. H. Guo, J. Yu, Z. Hu, *et al.*, “A hybrid clustering algorithm for multiple-source resolving in bioluminescence tomography,” *J. Biophotonics* **11**(4), e201700056 (2018).

5. Q. Zhang, H. Zhao, D. Chen, *et al.*, "Source sparsity based primal-dual interior-point method for three-dimensional bioluminescence tomography," *Opt. Commun.* **284**(24), 5871–5876 (2011).
6. Z. Hu, X. Chen, J. Liang, *et al.*, "Single photon emission computed tomography-guided Cerenkov luminescence tomography," *J. Appl. Phys.* **112**(2), 024703 (2012).
7. H. Guo, Z. Hu, X. He, *et al.*, "Non-convex sparse regularization approach framework for high multiple-source resolution in Cerenkov luminescence tomography," *Opt. Express* **25**(23), 28068–28085 (2017).
8. Z. Hu, C. Fang, B. Li, *et al.*, "First-in-human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-infrared-I/II windows," *Nat. Biomed. Eng.* **4**(3), 259–271 (2019).
9. J. Zhang, G. Zhang, Y. Chen, *et al.*, "Regularized reconstruction based on joint smoothly clipped absolute deviation regularization and graph manifold learning for fluorescence molecular tomography," *Phys. Med. Biol.* **68**(19), 195004 (2023).
10. H. Zhang, L. Hai, J. Kou, *et al.*, "OPK_SNCA: optimized prior knowledge via sparse non-convex approach for cone-beam X-ray luminescence computed tomography imaging," *Computer Methods and Programs in Biomedicine* **215**, 106645 (2022).
11. M. H. Oh, N. Lee, H. Kim, *et al.*, "Large-scale synthesis of bioinert tantalum oxide nanoparticles for X-ray computed tomography imaging and bimodal image-guided sentinel lymph node mapping," *J. Am. Chem. Soc.* **133**(14), 5508–5515 (2011).
12. J. M. Kinsella, R. E. Jimenez, P. P. Karmali, *et al.*, "X-ray computed tomography imaging of breast cancer by using targeted peptide-labeled bismuth sulfide nanoparticles," *Angew. Chem., Int. Ed.* **50**(51), 12308–12311 (2011).
13. G. Pratx, C. M. Carpenter, C. Sun, *et al.*, "X-ray luminescence computed tomography via selective excitation: a feasibility study," *IEEE Trans. Med. Imaging* **29**(12), 1992–1999 (2010).
14. P. Gao, K. Cheng, E. Schüller, *et al.*, "Restarted primal-dual Newton conjugate gradient method for enhanced spatial resolution of reconstructed cone-beam x-ray luminescence computed tomography images," *Phys. Med. Biol.* **65**(13), 135008 (2020).
15. T. Liu, J. Rong, P. Gao, *et al.*, "Regularized reconstruction based on joint L(1) and total variation for sparse-view cone-beam X-ray luminescence computed tomography," *Biomed. Opt. Express* **10**(1), 1–17 (2019).
16. P. Gao, J. Rong, H. Pu, *et al.*, "Sparse view cone beam x-ray luminescence tomography based on truncated singular value decomposition," *Opt. Express* **26**(18), 23233–23250 (2018).
17. J. Zhao, H. Guo, J. Yu, *et al.*, "A robust elastic net-l(1)l(2)reconstruction method for x-ray luminescence computed tomography," *Phys. Med. Biol.* **66**(19), 195005 (2021).
18. S. Tzoumas, D. Vernekohl, and L. Xing, "Coded-aperture compressed sensing x-ray luminescence tomography," *IEEE Trans. Biomed. Eng.* **65**(8), 1892–1895 (2018).
19. H. Zhang, G. Geng, S. Zhang, *et al.*, "Sparse non-convex lp regularization for cone-beam x-ray luminescence computed tomography," *J. Mod. Opt.* **65**(20), 2278–2289 (2018).
20. H. Pu, P. Gao, Y. Liu, *et al.*, "Principal component analysis based dynamic cone beam x-ray luminescence computed tomography: a feasibility study," *IEEE Trans. Med. Imaging* **38**(12), 2891–2902 (2019).
21. H. Guo, H. Zhao, J. Yu, *et al.*, "X-ray luminescence computed tomography using a hybrid proton propagation model and Lasso-LSQR algorithm," *J Biophotonics* **14**(11), e202100089 (2021).
22. Y. Chen, M. Du, G. Zhang, *et al.*, "Sparse reconstruction based on dictionary learning and group structure strategy for cone-beam X-ray luminescence computed tomography," *Opt. Express* **31**(15), 24845–24861 (2023).
23. Y. Tan and H. Jiang, "DOT guided fluorescence molecular tomography of arbitrarily shaped objects," *Med. Phys.* **35**(12), 5703–5707 (2008).
24. Z. Hu, M. Zhao, Y. Qu, *et al.*, "In vivo 3-dimensional radiopharmaceutical-excited fluorescence tomography," *J. Nucl. Med.* **58**(1), 169–174 (2017).
25. B. Dumitrescu and P. Irofti, *Dictionary Learning Algorithms and Applications* (Springer, 2018).
26. G. Wang, W. Yu, X. Liang, *et al.*, "An iterative reduction FISTA algorithm for large-scale LASSO," *SIAM J. Sci. Comput.* **44**(4), A1989–A2017 (2022).
27. J. Mairal, F. Bach, J. Ponce, *et al.*, "Online learning for matrix factorization and sparse coding," *Journal of Machine Learning Research* **11**, 19–60 (2010).
28. J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. Signal Process.* **60**(9), 4973–4976 (2012).
29. X. He, J. Liang, X. Wang, *et al.*, "Sparse reconstruction for quantitative bioluminescence tomography based on the incomplete variables truncated conjugate gradient method," *Opt. Express* **18**(24), 24825–24841 (2010).
30. H. Zhang, G. Geng, Y. Chen, *et al.*, "Combined multi-spectrum and orthogonal Laplacianfaces for fast CB-XLCT imaging with single-view data," *Opt. Rev.* **24**(6), 693–704 (2017).
31. D. Chen, S. Zhu, H. Yi, *et al.*, "Cone beam x-ray luminescence computed tomography: A feasibility study," *Med. Phys.* **40**(3), 031111 (2013).
32. A. D. Klose, V. Ntziachristos, and A. H. Hielscher, "The inverse source problem based on the radiative transfer equation in optical molecular imaging," *J. Comput. Phys.* **202**(1), 323–345 (2005).
33. L. V. Wang and H. I. Wu, *Biomedical Optics: Principles and Imaging* (John Wiley & Sons, 2007).
34. M. Schweiger, S. Arridge, M. Hiraoka, *et al.*, "The finite element method for the propagation of light in scattering media: boundary and source conditions," *Med. Phys.* **22**(11), 1779–1792 (1995).

35. X. Liu, X. Tang, Y. Shu, *et al.*, "Single-view cone-beam x-ray luminescence optical tomography based on Group_YALL1 method," *Phys. Med. Biol.* **64**(10), 105004 (2019).
36. Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course* (Springer Science & Business Media, 2013).
37. H.-C. Wu, "The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function," *European Journal of Operational Research* **176**(1), 46–59 (2007).
38. E. J. Dickinson, H. Ekström, and E. Fontes, "COMSOL Multiphysics®: Finite element software for electrochemical analysis. A mini-review," *Electrochem. Commun.* **40**, 71–74 (2014).
39. H. Li, J. Tian, F. Zhu, *et al.*, "A mouse optical simulation environment (MOSE) to investigate bioluminescent phenomena in the living mouse with the Monte Carlo method," *Academic Radiology* **11**(9), 1029–1038 (2004).
40. L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A* **1**(6), 612–619 (1984).
41. Y. Chen, W. Li, M. Du, *et al.*, "Elastic net-based non-negative iterative three-operator splitting strategy for Cerenkov luminescence tomography," *Opt. Express* **30**(20), 35282–35299 (2022).
42. Y. An, J. Liu, G. Zhang, *et al.*, "A novel region reconstruction method for fluorescence molecular tomography," *IEEE Trans. Biomed. Eng.* **62**(7), 1818–1826 (2015).