Efficient L1 regularization-based reconstruction for fluorescent molecular tomography using restarted nonlinear conjugate gradient

Junwei Shi,[1] Bin Zhang,[1] Fei Liu,[1] Jianwen Luo,[1,2] and Jing Bai[1,*]

[1]*Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China*

[2]*Center for Biomedical Imaging Research, School of Medicine, Tsinghua University, Beijing 100084, China*

**Corresponding author: deabj@tsinghua.edu.cn*

Received June 17, 2013; revised August 12, 2013; accepted August 22, 2013;
posted August 23, 2013 (Doc. ID 192378); published September 13, 2013

For the ill-posed fluorescent molecular tomography (FMT) inverse problem, the L1 regularization can protect the high-frequency information like edges while effectively reduce the image noise. However, the state-of-the-art L1 regularization-based algorithms for FMT reconstruction are expensive in memory, especially for large-scale problems. An efficient L1 regularization-based reconstruction algorithm based on nonlinear conjugate gradient with restarted strategy is proposed to increase the computational speed with low memory consumption. The reconstruction results from phantom experiments demonstrate that the proposed algorithm can obtain high spatial resolution and high signal-to-noise ratio, as well as high localization accuracy for fluorescence targets.

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

http://dx.doi.org/10.1364/OL.38.003696

For the fluorescent molecular tomography inverse problem (FMT-IP), the reconstruction is highly susceptible to noise and numerical errors. To compute a meaningful approximate solution, L2 regularization is usually incorporated into the solution, and the Tikhonov algorithm is a widely used L2-regularized method. However, the fast spatial changes in the solution are often over-smoothed, and the localized features are lost during the reconstruction process [1]. The usage of L1 regularization to reconstruct the fluorescence distribution and improve the reconstructed image quality has been an active area of research. Unlike L2 regularization (such as Tikhonov), there is no analytic formula for the optimal solution to the L1-regularized FMT-IP whose objective function is convex but not differentiable. Thus, solving the L1-regularized FMT-IP is more of a computational challenge than the L2-regularized FMT-IP. Moreover, although L1 regularization has been widely used in magnetic resonance imaging (MRI) and computed tomography (CT) [2,3], specific algorithms should be developed for FMT-IP owing to its severe ill-condition. Recently, several researchers have proposed reconstruction methods and variants for solving L1-regularized FMT-IP, typically including L1-Ls based on the primal-dual interior-point approach [1], truncated Newton (TN) method based on interior-point approach [4], iterative shrinkage-thresholding algorithm (IST) [5], separable paraboloidal surrogates algorithm based on majorization-minimization (SPS-MM) framework [6], and the greedy algorithm based on match pursuit (MP) framework [7]. However, in these algorithms, the inverse Hessian matrix needs to be calculated or preconditioned, which consumes large amounts of runtime and memory space, especially for the large-scale FMT-IP. Although the greedy algorithm performs well for FMT-IP, it demands that the columns in the weight matrix of FMT-IP are nearly incoherent. Therefore, the matrix preprocessing is time-consuming.

In this Letter, a method based on an L1-regularized nonlinear conjugate gradient (L1-NCG) descent algorithm with backtracking line search [3] is used to solve the L1-regularized FMT-IP. Unlike the L1-Ls, TN and SPS-MM algorithms, L1-NCG as a conjugate gradient algorithm does not involve the inversion of Hessian matrix. In contrast to the IST algorithm, L1-NCG does not involve the calculation the eigenvalues of Hessian matrix. Compared with the greedy algorithm described in [7], L1-NCG does not need the singular value decomposition of the weight matrix. However, the convergence speed of L1-NCG decreases after some iterations due to the ill-condition of FMT-IP. In order to increase the convergence speed, we propose a restarted strategy, where the searching descent direction and initial value are reset, and the permission region is set meanwhile, after a fixed iteration of L1-NCG. With the help of permission region, the computational speed can be increased with low memory consumption. Moreover, the restarted strategy profits from the fact that the convergence of L1-NCG depends on the iteratively updated initial value.

The forward model used to predict photon propagation in highly scattering media, such as biological tissues, is based on the coupled diffusion equations (DE) with Robin-type boundary condition [8]. Through the normalized Born approximation to the DE, the nonlinear FMT-IP can be linearized and all the position-dependent gain factors in forward model can be canceled out [9]. The normalized Born approximation can be formulated as follows [9]:

$$\frac{\phi_{\mathrm{em}}(r_d,r_s)}{\phi_{\mathrm{ex}}(r_d,r_s)}=\frac{S_0\upsilon}{G(r_s,r_d,\lambda_1)D^{\prime}}\int G(r_d,r,\lambda_2)x(r)G(r_s,r,\lambda_1)d^3r, \tag{1}$$

where $\phi_{\mathrm{em}}(r_d,r_s)$ denotes the average intensities at emission wavelength λ_2 and $\phi_{\mathrm{ex}}(r_d,r_s)$ denotes the average

intensities at excitation wavelength λ_1. $G(r_s, r, \lambda_1)$ denotes the analytically calculated photon field at excitation wavelength λ_1 induced at position r by a source at position r_s in the tissue. $G(r_d, r, \lambda_2)$ is the Green's function which describes photon propagation from point r to the detector point r_d at emission wavelength λ_2. D^f is the diffusion coefficient of the tissue at emission wavelength λ_2, υ is the speed of light in the tissue, and S_0 is a calibration factor accounting for various system gains and attenuation factors. After the image domain is discretized, the FMT problem can be formulated as the following linear matrix equation:

$$\Phi_f = WX, \tag{2}$$

where W denotes the weight matrix mapping the reconstructed vectorized fluorochrome concentrations X into the fluorescent measurement vector Φ_f.

To guarantee the uniqueness and stability of FMT-IP and to preserve high spatial resolution, L1 regularization is adopted. Then, the corresponding optimization function is formulated as follows:

$$\arg\min_{X\geq 0} \|WX - \Phi_f\|_2^2 + \lambda\|X\|_1, \tag{3}$$

where λ is the regularization parameter balancing the data fitting and L1 penalty. Through normalizing the columns of matrix W, defined in Eq. (3), the high attenuation of emission light coming from deep fluorochromes in the tissue can be compensated. The normalization diagonal matrix W_{nor} is defined as follows:

$$W_{\text{nor}}(i,j) = \|W_i\|_2^{-1} \qquad i = j \quad \text{or} \quad 0\ i \neq j, \tag{4}$$

where W_i is the ith column of matrix W. In order to unify the magnitude of the measurement data for different FMT problems, the vector Φ_f is normalized by its maximum. Then, Eq. (4) is rewritten as follows:

$$X_{L1} = \arg\min_{X\geq 0}(f(X) = \|W_{\text{new}}X - \Phi_{f\text{new}}\|_2^2 + \lambda\|X\|_1), \tag{5}$$

$$W_{\text{new}} = WW_{\text{nor}}; \qquad \Phi_{f\text{new}} = \Phi_f / \max(\Phi_f). \tag{6}$$

Equation (5) is solved through the NCG descent algorithm with backtracking line search [3], where the gradient of $f(X)$ formulated as follows:

$$\nabla f(X) = 2W_{\text{new}}^T(W_{\text{new}}X - \Phi_{f\text{new}}) + \lambda\nabla\|X\|_1, \tag{7}$$

where the L1 norm $\|X\|_1$ is the sum of absolute values. Equation (7) is not well defined for all values of X, as the absolute value function is not a smooth function. Instead, the L1 norm $\|X\|_1$ is approximated with a smooth function by using the relation $|X| \approx \sqrt{X * X + \mu}$, where μ is a positive smooth parameter. The basic NCG iteration for Eq. (5) is

$$X_{L1}^{k+1} = X_{L1}^k + t_k d_k, \tag{8}$$

where d_k denotes the search direction and t_k denotes the step size. With $d_0 = -\nabla f(X_{L1}^0)$, the search direction is updated as follows:

$$d_k = -\nabla f(X_{L1}^k) + \gamma_k d_{k-1}, \tag{9}$$

where γ_k is given based on the Fletcher–Reeves (FR) method. Then, for Eq. (8), the backtracking line search is used to find an appropriate step size t_k. The L1-regularized NCG (L1-NCG) is summarized as follows:

1. Set stopping criteria by number of iterations, $k_{\max} = 100$. Set the parameters for backtracking line search, $\alpha = 0.05$, $\beta = 0.6$. Set original value, $X_{L1}^0 = 0$. Set original search direction, $d_0 = -\nabla f(X_{L1}^0)$. $k = 0$.
2. $t_k = 1$, while $f(X_{L1}^k + t_k d_k) > f(X_{L1}^k) + \alpha t_k d_k \nabla f(X_{L1}^k)$, then $t_k = \beta t_k$(compute the step size).
3. $X_{L1}^{k+1} = X_{L1}^k + t_k d_k$(update X_{L1}^{k+1} with step size t_k and search direction d_k), and then compute the parameter γ_k used to update search direction d_k as $\gamma_k = \|\nabla f(X_{L1}^{k+1})\|_2^2 / \|\nabla f(X_{L1}^k)\|_2^2$.
4. $d_{k+1} = -\nabla f(X_{L1}^{k+1}) + \gamma_k d_k$(update the search direction with newly computed X_{L1}^{k+1} and γ_k).
5. $k = k + 1$, if $k > k_{\max}$, stop, otherwise go to step 2.

Actually, in the L1-NCG method, the search directions are not strictly conjugate between each other and the step sizes are approximated suboptimally, unlike the linear conjugate gradient method, and the severe ill-condition of the FMT-IP makes it difficult to achieve satisfactory reconstruction results. In this Letter, restarted strategy is adopted to remit the aforementioned issue of the L1-NCG algorithm. When the iterations k in L1-NCG equals to $k_{\max}$, the search direction d_{k+1} is reset as $d_{k+1} = -\nabla f(X_{L1}^{k+1})$, rather than $d_{k+1} = -\nabla f(X_{L1}^{k+1}) + \gamma_k d_k$. With this restarted search direction, the convergence speed can be increased and the results are closer to the true value. After the first $k_{\max}$ iterations, the negative value in $X_{L1}^{k_{\max}}$ is set to zero and then used as the initial value for the next iteration. The convergence speed of L1-NCG can be increased through using the restarted strategy, as L1-NCG profits from a good initial value. As permission region, in matrix W_{new}, the columns corresponding to zeros in $X_{L1}^{k_{\max}}$ can be removed. Then, the dimension of Eq. (5) scales down significantly, which increases the computational speed and reduces the memory consumption. Moreover, by means of the permission region, the ill-posedness can be alleviated. The restarted L1-NCG (re-L1-NCG) algorithm is summarized as follows:

1. Set the restart iterations, $k_{\text{restart}} = 10$, $k = 1$.
2. Compute the solution $X_{L1-\text{re}}^1$ of Eq. (5) using aforementioned L1-NCG with $k_{\max}$ iterations. Set the negative value in $X_{L1-\text{re}}^1$ to zero.
3. By removing columns corresponding to zero values in $X_{L1-\text{re}}^k$, W_{new} is compressed into $W_{\text{new}}^{\text{re}}$ (permission-region strategy). Then, the optimization function turns into $\arg\min_{X\geq 0}(f(X) = \|W_{\text{new}}^{\text{re}}X - \Phi_{f\text{new}}\|_2^2 + \lambda\|X\|_1)$.
4. $X_{L1-re}^0 = \max\{X_{L1-re}^k, 0\}$(reset initial value).
5. Solve $\arg\min_{X\geq 0}(f(X) = \|W_{\text{new}}^{\text{re}}X - \Phi_{f\text{new}}\|_2^2 + \lambda\|X\|_1)$ for $X_{L1-\text{re}}^{k+1}$ using aforementioned L1-NCG (with

restarted search direction) with $k_{\max}$ iterations and the initial value for L1-NCG is set as X^0_{L1-re}.

6. $k = k + 1$, if $k > k_{\text{restart}}$, stop, otherwise go to step 3.

In addition, the degree of ill-posedness varies with different FMT problems at specific experimental parameters. Thus, choosing the regularization parameter of Eq. (3) is a difficult task, as the regularization parameter depends on the degree of ill-posedness. However, the reconstruction results with re-L1-NCG are not sensitive to the regularization parameter in the proper range due to the normalization of weight matrix W and measurement vector Φ_f. In this Letter, λ is manually optimized as 15.

In order to evaluate the performance of the proposed algorithm, phantom experiments were conducted on the FMT system [10]. The detector field-of-view corresponding to each point excitation source was about 180°, and 36 projections were adopted. A transparent glass cylinder with a diameter of 30 mm and height of 60 mm filled with 1% intralipid was employed as the phantom. The geometry configuration is shown in Figs. 1(a1) and 1(a2). Two cylinders (with a diameter of 4 mm) filled with 10 μL Indocyanine Green with a concentration of 1.3 μM were used as the fluorescent targets. The edge-to-edge distance between the two targets was approximately 8 mm, and the positions were at (2.2 mm, −5.5 mm, 26.5 mm) and (−2.1 mm, 5.8 mm, 26.5 mm), respectively. The transverse CT image of the two cylinder fluorophores is shown in Fig. 1(f1). The Kirchhoff approximation (KA) method, as an analytical method, was used to solve the FMT forward model, in order to reduce the computation burden of numerical methods [11]. The size of the weight matrix W was 8571×4995, where the detector spatial sampling was set as 2 mm and the discretized mesh resolution of the phantom was set as 1.5 mm. The normalized Born measurements in Eq. (1) were calculated from the ratio between the light intensities measured at the emission and excitation wavelengths.

In order to demonstrate the advantage of the proposed algorithm, reconstruction results obtained with different algorithms (L2-regularized FMT-IP with Tikhonov algorithm, L1-regularized FMT-IP with re-L1-NCG and L1-NCG algorithms) are shown in Fig. 1. To emphasize the importance of the permission-region strategy used in the re-L1-NCG algorithm, the reconstruction result using L1-NCG with only 100 iterations (L1-NCG-100) are shown in Figs. 1(b1) and 1(b2). It can be seen that the reconstructed fluorophore distribution is highly spread out and there is little noise in the result. However, by comparison, the true fluorophores in Fig. 1(a2) are included in the result in Fig. 1(b2). Therefore, the reconstruction result obtained from L1-NCG-100 can serve as the permission region.

Like the Tikhonov algorithm, the main memory consumption of the re-L1-NCG algorithm depends on the size of matrix W^TW [see Eq. (7)]. Thus, in this Letter, the memory consumption of different algorithms is compared based on the size of matrix W^TW. Corresponding to step 3 in the re-L1-NCG algorithm, the size of the weight matrix W was greatly reduced from 8571×4995 to 8571×1051, thus reducing the memory consumption needed in Eq. (7) and improving the computational speed of the re-L1-NCG algorithm. Although there may be some wrong permission regions owing to the noise, the permission region can work well as the wrong anatomical priors in multimode FMT reconstruction [12]. Above all, L1-NCG as a representative gradient method relies heavily on the initial value, and the results obtained by L1-NCG-100 can be used as the initial value for rapid convergence in the re-L1-NCG algorithm.

Figures 1(c1) and 1(c2) show the reconstruction result using the proposed re-L1-NCG algorithm, where L1-NCG-100 was restarted for 10 times with the restarted strategy. As shown in Figs. 1(c1) and 1(c2), the reconstruction result using the re-L1-NCG has higher spatial resolution and signal-to-noise (SNR). In order to compare the reconstructed results with the same iterations as re-L1-NCG, the result using L1-NCG with 1000 iterations (L1-NCG-1000) are shown in Figs. 1(d1) and 1(d2). The reconstruction result obtained from L1-NCG-1000 has a lower SNR and the reconstructed fluorescence targets

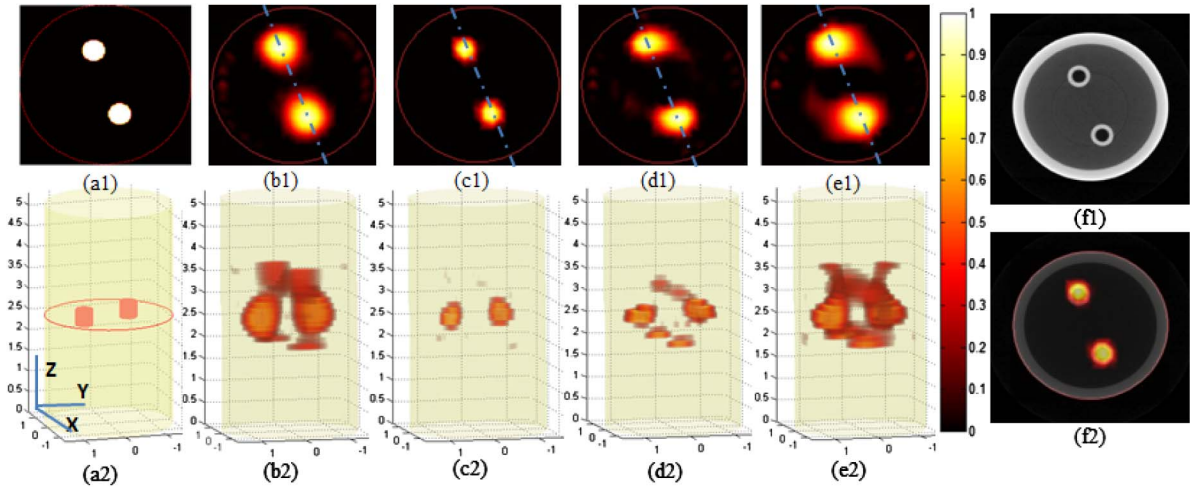

Fig. 1. Reconstruction results. (a1) True cross section in the $z = 26$ mm plane, (a2) 3D rendering of true fluorescent targets. Slice (b1) and stereo (b2) are the reconstructed results obtained from L1-NCG-100. Slice (c1) and stereo (c2) correspond to re-L1-NCG algorithm. Slice (d1) and stereo (d2) correspond to L1-NCG-1000. Slice (e1) and stereo (e2) correspond to Tikhonov algorithm. (f1) The transverse section of CT image in $z = 26$ mm plane. (f2) The transverse overlay of (c1) and (f1).

Table 1. Runtime and Memory Consumption for re-L1-NCG, L1-NCG-1000 and Tikhonov

Method	Runtime(s)	Size of W^TW
re-L1-NCG	44.5	1051×1051
L1-NCG-1000	214.1	4995×4995
Tikhonov	45.5	4995×4995

tend to become superficial. On one side, the severe ill-condition of the FMT-IP leads to the slow L1-NCG convergence. On the other side, the noise in the first few iterations leads to the wrong convergence, as the convergence of L1-NCG benefits from a good initial value.

The reconstruction results using Tikhonov algorithm correspond to Figs. 1(e1) and 1(e2). Compared with re-L1-NCG, the reconstruction results obtained from Tikhonov have lower SNR and spatial resolution. In order to quantitatively evaluate the performance of the proposed algorithm, the runtime and memory consumption are shown in Table 1. Benefitting from the permission region, re-L1-NCG is more efficient than L1-NCG and Tikhonov.

To illustrate the localization accuracy, the cross section [Fig. 1(c1)] is overlaid onto the corresponding CT result [Fig. 1(f1)] in Fig. 1(f2). It can be observed that the proposed method detects the two fluorescence targets with fairly good accuracy. The profiles along the dash dotted lines [in Figs. 1(b1), 1(c1), 1(d1), and 1(e1)] are shown in Fig. 2 in order to better clarify the performance of the proposed re-L1-NCG algorithm. As shown, the locations of the peaks of the result obtained from re-L1-NCG have a better match with the true locations, while the result obtained from L1-NCG-1000 tends toward boundary pixels, and the full width at half-maximum (FWHM) corresponding to L1-NCG-1000 and Tikhonov is larger than that of re-L1-NCG.

In conclusion, we have proposed a method using re-L1-NCG to achieve high-quality L1-regularized FMT reconstruction (even with a large scale), where the restarted strategy is adopted to increase the convergence speed with low memory consumption (see Table 1). Results from phantom experiments demonstrate that the proposed method is capable of localizing fluorescence targets at high efficiency with high spatial resolution and high SNR. In the future, we will verify the proposed re-L1-NCG algorithm for high-resolution and limited-projection FMT reconstruction.

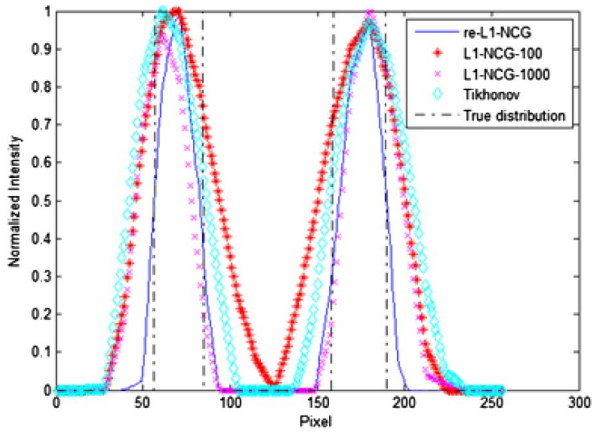

Fig. 2. Intensity profiles along the dash–dotted lines in Figs. 1(b1), 1(c1), 1(d1), and 1(e1).

This work is supported by the National Basic Research Program of China (973) under Grant No. 2011CB707701; the National Natural Science Foundation of China under Grant No. 81227901, 81071191, 81271617; the National Major Scientific Instrument and Equipment Development Project under Grant No. 2011YQ030114; and the National Science and Technology Support Program under Grant No. 2012BAI23B00.

References

1. H. Gao and H. Zhao, Opt. Express **18**, 1854 (2010).
2. J. C. Ramirez-Giraldo, J. Trzasko, S. Leng, L. Yu, A. Manduca, and C. H. McCollough, Med. Phys. **38**, 2157 (2011).
3. M. Lustig, D. L. Donoho, and J. M. Pauly, Magn. Reson. Med. **58**, 1182 (2007).
4. X. He, Y. Hou, D. Chen, Y. Jiang, M. Shen, J. Liu, Q. Zhang, and J. Tian, J. Biomed. Imaging **2011**, 1 (2011).
5. D. Han, J. Tian, S. Zhu, J. Feng, C. Qin, B. Zhang, and X. Yang, Opt. Express **18**, 8630 (2010).
6. J. Dutta, S. Ahn, C. Li, S. R. Cherry, and R. M. Leahy, Phys. Med. Biol. **57**, 1459 (2012).
7. J. Shi, X. Cao, F. Liu, B. Zhang, J. Luo, and J. Bai, J. Opt. Soc. Am. **30**, 437 (2013).
8. S. R. Arridge, Inverse Probl. **15**, R41 (1999).
9. Soubret, Antoine, J. Ripoll, and V. Ntziachristos, IEEE Trans. Med. Imaging. **24**, 1377 (2005).
10. F. Liu, X. Liu, D. Wang, B. Zhang, and J. Bai, Ann. Biomed. Eng. **38**, 3440 (2010).
11. J. Ripoll, M. N. Vesperinas, R. Weissleder, and V. Ntziachristos, Opt. Lett. **27**, 527 (2002).
12. B. Zhang, S. Liu, X. Cao, F. Liu, X. Wang, J. Luo, B. Shan, and J. Bai, IEEE Trans. Multimedia **99**, 1 (2013).