Total variation regularization for 3D reconstruction in fluorescence tomography: experimental phantom studies

Ali Behrooz,[1] Hao-Min Zhou,[2] Ali A. Eftekhar,[1] and Ali Adibi[1,*]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, 777 Atlantic Dr., Atlanta, Georgia 30332, USA

[2]School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, Georgia 30332, USA

*Corresponding author: adibi@ece.gatech.edu

Received 25 May 2012; revised 17 August 2012; accepted 16 October 2012;
posted 31 October 2012 (Doc. ID 169359); published 30 November 2012

Fluorescence tomography (FT) is depth-resolved three-dimensional (3D) localization and quantification of fluorescence distribution in biological tissue and entails a highly ill-conditioned problem as depth information must be extracted from boundary measurements. Conventionally, L_2 regularization schemes that penalize the Euclidean norm of the solution and possess smoothing effects are used for FT reconstruction. Oversmooth, continuous reconstructions lack high-frequency edge-type features of the original distribution and yield poor resolution. We propose an alternative regularization method for FT that penalizes the total variation (TV) norm of the solution to preserve sharp transitions in the reconstructed fluorescence map while overcoming ill-posedness. We have developed two iterative methods for fast 3D reconstruction in FT based on TV regularization inspired by Rudin–Osher–Fatemi and split Bregman algorithms. The performance of the proposed method is studied in a phantom-based experiment using a noncontact constant-wave trans-illumination FT system. It is observed that the proposed method performs better in resolving fluorescence inclusions at different depths.

OCIS codes: 100.3190, 100.6950, 170.3010, 170.5280.

1. Introduction

Fluorescence tomography (FT) is an emerging three-dimensional (3D) optical imaging modality used for *in vivo* noninvasive depth-resolved localization and quantification of fluorescent-tagged inclusions, e.g., cancer lesions and test drugs, buried a few centimeters deep in biological tissue. FT is extensively employed in early cancer detection as well as drug monitoring and discovery [1,2]. In this imaging technique, the tissue is illuminated with red visible or near infrared light at different boundary locations at the excitation wavelength of exogenously administered fluorophores that label target inclusions. The fluorescent signal emitted by fluorophore probes is collected at several locations on the skin. These surface measurements are then used in an inversion algorithm to reconstruct the 3D distribution of fluorophores in the tissue [1].

Reconstructing the 3D fluorescent distribution from boundary measurements is a highly ill-conditioned problem, as depth information should be extracted from data collected on the surface [1]. Therefore, inverse solvers resort to regularization techniques [3] to overcome ill-posedness and minimize artifacts and errors arising from the ill-posed nature of the problem. Conventionally, Laplacian [4] and Tikhonov [5] regularization methods that penalize L_2 norm of the solution or a linear transformation of the solution, and therefore compromise accuracy for stability [3], have been used for 3D reconstruction in FT. These L_2 norm regularization algorithms entail smoothing effects and result in continuous and spatially overspread reconstructions while damping out noise-induced artifacts. They filter out the high-frequency

1559-128X/12/348216-12$15.00/0

and edge-type features of the fluorescence distribution and impair the resolution offered by FT reconstruction while removing artifacts caused by noise and modeling errors [3]. Results reported for 3D FT reconstructions using Laplacian and Tikhonov regularization methods [4,5] possess negligible spreading and are relatively well-resolved in shallow depths. However, as the depth of the fluorescent inclusions increases, the need for stronger regularization rises, which in turn results in high spatial spreading and poor resolution. As a result, the resolution of L_2 regularization algorithms degrades as the depth of the inclusions increases. Structural and anatomical priors have been used to improve the performance of L_2 norm regularization methods and adapt the regularization parameters to the geometry and a priori information of the FT problem. This has been shown to greatly enhance the accuracy of 3D reconstructions in a few studies [6–9]. However, the performance of adaptive methods relies on availability of priors. Thus, in problems with limited or error-contaminated priors, adaptive methods cannot perform optimally.

While L_2 regularization is used extensively for FT reconstruction, regularization of the FT inverse problem using other L_p norms has been a subject of study in recent years [10,11]. Among them, L_1 regularization has been shown to perform optimally for reconstruction of sparse and localized fluorescent distributions in scenarios such as early-stage cancer detection [11]. Various implementations of L_1 regularization have been applied to FT and shown to improve 3D reconstruction of sparse fluorophore distributions [11–14]. Optimal performance of L_1 regularization methods is limited to scenarios with sparsity priors and cannot be applied to general cases with nonlocalized extended fluorescent distributions. Row-action iterative inversion algorithms are also extensively used for FT reconstruction, among which algebraic reconstruction techniques (ARTs) [15], notable for their memory efficiency and speedy convergence, have been shown to yield fast and stable reconstructions in FT [16,17]. Nevertheless, the performance of ART deteriorates with increase in modeling errors and data noise level. ARTs do not possess strong regularizing power and are not suitable for performing reconstructions on noisy data and error contaminated models.

In this paper, we report the use of total variation (TV) regularization for FT reconstruction where the TV seminorm of the solution is penalized to stabilize the reconstruction against artifacts and errors in data and modeling. While application of TV regularization to nonlinear FT and use of TV seminorm jointly with L_1 norm as penalty terms for regularization have been the subject of very recent studies [18,19], the use of TV regularization has not been studied for linear FT to our knowledge. We apply the proposed TV regularization technique to two-dimensional (2D) simulated FT data with different noise levels and compare its performance with conventional regularization and iterative reconstruction techniques. Moreover, we investigate the performance of TV regularization in 3D reconstruction of fluorescent inclusions in a phantom-based experiment where two fluorophore-filled tubes are placed inside a liquid tissue phantom that is excited by near infrared laser radiation and imaged by a noncontact cooled charged-coupled device (CCD) camera. We compare the performance of the proposed method with conventional regularization techniques in reconstructing the fluorescent tubes at different depths.

2. Theory and System Modeling

Propagation of near infrared light in diffusive media like biological tissue can be mathematically modeled by the diffusion approximation to radiative transfer equation [20], which yields a first order partial differential equation that describes the behavior of diffusive photons as follows:

$$\nabla \cdot D(r)\nabla\Phi(r) - \mu_a(r)\Phi(r) = -q(r), \tag{1}$$

where $\Phi(r)$ represents the average light intensity, $\mu_a(r)$ is the absorption coefficient, $D(r)$ is the diffusion coefficient, and $q(r)$ is the source strength at location r [1,20]. By applying the finite element method to Eq. (1) and discretizing the tissue volume using a tetrahedral mesh, we can relate the fluorophore concentration at each voxel to the surface measurements of emitted fluorescent light intensity through a linear system of equations [21]. Surface measurements consist of fluorescent signal intensity readings $\Phi(r)$ on the boundary of the turbid medium at detector locations. Fluorophore concentration at each voxel multiplied by the quantum efficiency of the fluorophore and the intensity of excitation light yields $q(r)$. Hence, using prior knowledge of $\mu_a(r)$ and $D(r)$ values in the medium, boundary measurements of $\Phi(r)$ can be related to $q(r)$ and in turn, to fluorophore concentration at each voxel through a linear operator. More details about the linear model of FT can be found in Appendix A and [21].

When a tetrahedral finite element (FE) mesh is used to discretize the turbid medium, all the quantities are transformed into discrete vector quantities. For a mesh with K voxels, the fluorophore concentration at each voxel is stacked in the $K \times 1$ vector x. For N_s source positions and N_d detector positions, there are $N_s \times N_d$ surface measurements that are usable as data for FT reconstruction and are represented by the $N_sN_d \times 1$ vector y, and the noise and modeling errors present in the measurements are represented by the $N_sN_d \times 1$ vector n. Therefore, the linear model relating fluorophore concentrations to surface measurements using the discretized version of Eq. (1) can be expressed in the form of an $N_sN_d \times K$ system matrix M, which depends on the optical properties of the tissue and the measurement geometry [21] as follows:

$$y = Mx + n. \tag{2}$$

FT reconstruction methods aim at estimating the fluorophore distribution x using data vector y, system matrix M, and statistical priors on n. Since M is a highly ill-conditioned matrix with a relatively large condition number [1,3,4], unbiased estimation techniques, such as nonregularized weighted least squares [22], result in unstable and artifact-contaminated reconstructions. Therefore, inverse solvers use regularized least-squares techniques that penalize the norm of a linear transformation of the solution along with the data fidelity term to avoid unstable solutions as formulated below:

$$\min_x \|y - Mx\|^2 + \lambda^2 \|Lx\|^2, \tag{3}$$

where is λ the regularization parameter, and L is the regularization weight matrix.

The norm used for the data fidelity term $\|y - Mx\|$ is usually chosen to be L_2 to provide a least-squares fit to the data. The norm for the penalty term $\lambda^2\|Lx\|$, with a linear operator L acting on the solution vector x is conventionally selected to be L_2 [3–5] while recently other L_p $(p \neq 2)$ norms have been used for cases with specific priors like sparsity [10–14] as discussed previously in Section 1. The major limitation of L_2-norm regularization stems from its oversmoothing property where, similar to low-pass filtering for de-noising of one-dimensional (1D) signals, high-frequency and edge-type features of the reconstructed map are removed while filtering out artifacts and noise [3]. Therefore artifacts are removed and solution is stabilized to the detriment of sharp transitions and well-resolved inclusions being oversmoothed. In L_2 regularization, the spatial resolution of the reconstruction algorithm is impaired at the expense of providing stability. To shed more light on this matter and provide a better insight into the resolution of image reconstruction, we study point spread functions (PSFs), a concept used in resolution studies to compare the spatial resolving power of image reconstruction methods applied to linear problems [23]. For a given medium and data acquisition geometry, PSF is the reconstruction of an image with a single nonzero pixel (or nonzero voxel in the case of FT) in perfect theoretical settings with no noise or errors present. For FT, an image reconstructed from noiseless data generated from a fluorophore distribution that is zero everywhere except at one voxel represents the PSF of that voxel. PSFs reveal certain qualities of image reconstruction techniques irrespective of the level of noise and modeling errors. They reveal the distortions and artifacts generated in an image when projected to the data domain and back to the image domain again through a reconstruction algorithm. For L_2 norm regularization mathematically formulated as

$$\min_x \|y - Mx\|_2^2 + \lambda^2 \|Lx\|_2^2, \tag{4}$$

in perfect noiseless theoretical settings suitable for exploring PSFs, the reconstructed image x_{rec} can be related to the true image x_{orig} through the linear system [22,23]

$$(M * M + \lambda^2 L * L)x_{\text{rec}} = M * Mx_{\text{orig}}. \tag{5}$$

Therefore, to calculate the PSFs of L_2 norm regularization we need to substitute x_{orig} in Eq. (5) by a vector with a single nonzero entry and solve for x_{rec}. In Section 3, we will compute and plot the PSFs to compare the resolution offered by L_2 regularization and the proposed TV regularization.

3. Proposed Approach

In this paper we propose TV regularization for 3D reconstruction in FT and explore its advantages over conventional L_2 regularization. TV image reconstruction is a widely used method in image processing and has been shown to enhance contrast and resolution of image reconstruction in medical imaging modalities such as bioluminescence tomography and emission tomography [24,25]. We investigate applying two fast implementations of TV regularization to FT. In this regularization technique, the penalty term added to the least-squares data fidelity is the TV norm of the solution as follows:

$$\min_x \|y - Mx\|_2^2 x + \lambda^2 \|x\|_{\text{TV}}, \tag{6}$$

where the discretized 3D fluorophore distribution is stacked in the 1D vector x. If u denotes the 3D continuous scalar fluorophore distribution function, the TV norm can be expressed as

$$\|x\|_{\text{TV}} = \int |\nabla u| \mathrm{d}\Omega. \tag{7}$$

The integral in Eq. (7) is taken over the space where x is defined, $d\Omega$ represents the differential element for volume, and ∇u is the gradient of u. TV norm can be viewed as the L_1 norm of the gradient of the 3D fluorescent distribution. Hence in TV regularization, instead of penalizing the size of the solution, the size of the variations in the solution is penalized. Unlike L_2 regularization, the edges and sharp transitions are not smoothed and only highly oscillatory and variational components are filtered out in TV regularization [26]. Therefore, TV regularization is expected to yield a better resolution over conventional regularization techniques. To compare the resolving power of TV regularization with L_2 regularization, we compare their PSFs in a 2D numerical study. In L_2 regularization, as discussed previously, columns of the resolution matrix $R = (M * M + \lambda^2 L * L)^{-1} M * M$ are the PSFs of voxels (entries) in x. As λ increases, i.e., the strength of regularization increases, the term $\lambda^2 L * L$ becomes larger, which results in more off-diagonal nonzero entries in the resolution matrix. Hence, the resolution diminishes with increase in regularization strength [22].

In TV regularization the reconstructed fluorescent map x_{rec} is related to the true fluorescent map x_{orig} by

$$\left(M * M + \lambda^2 \nabla . \frac{\nabla}{|\nabla u|}\right) x_{\text{rec}} = M * M x_{\text{orig}}, \qquad (8)$$

which can be derived by differentiating Eq. (6) and setting it to zero. Also in arriving at Eq. (8), we assume that $|\nabla u_{\text{rec}}| = |\nabla u_{\text{orig}}|$, considering the edge-preserving behavior of TV regularization. To obtain the PSFs for TV regularization, we solve for x_{rec} in Eq. (8) when x_{orig} is substituted by a vector with only one nonzero entry. Figure 1 depicts a comparison of PSFs for TV regularization and L_2 regularization of moderate strengths in a 2D numerical study. For an off-edge voxel, both regularization techniques create spreading around the voxel. However, for on-edge voxel, TV regularization has no spreading across the edge and hence does not diffuse or smooth it while L_2 regularization creates spreading in all directions just like it did for off-edge voxels.

We observe in Fig. 1 that the PSF of TV regularization has almost no smoothing or regularizing effect for on-edge voxels whereas it has moderate level of spreading or regularization for off-edge voxels. This edge-preserving adaptive nature of TV regularization can be explained by the fact that [as expressed in Eq. (8)] the regularizing term or the operator in TV regularization, $\lambda^2 \nabla \cdot (\nabla / |\nabla u|)$, depends on the solution or more precisely the gradient of the solution. The denominator $|\nabla u|$ plays an important role in the performance of TV regularization. It assumes large values on edges and small values on smooth or flat regions of u, making $\lambda^2 \nabla \cdot (\nabla / |\nabla u|)$ to be small for voxels sitting on the edge and large for those away from edges or sharp transitions. So TV regularization can be viewed as an adaptive form of regularization. For the voxels away from edges in smooth regions there is strong regularization that damps out oscillations and high frequency components; for voxels sitting on an edge, there is weak or almost no regularization and hence the high frequency components (that make up the edge) are not filtered out. The PSFs do not possess smoothing around voxels sitting on edges, and therefore they are preserved when fed through the regularization algorithm. This edge-preserving property of TV regularization allows for higher reconstruction resolution compared to L_2 regularization.

To implement TV regularization for FT, we need to numerically solve the minimization in Eq. (6), which is a nonlinear optimization problem. We use two different approaches to solve for x in Eq. (6). We use a variational method inspired by Rudin–Osher–Fatemi (ROF) TV de-noising [27] and an iterative method inspired by linearized split Bregman iteration [28,29] to solve the optimization problem in Eq. (6).

In the first approach, an ROF-based TV regularization algorithm is employed to solve the minimization problem [27]. Let T denote the continuous to discrete plus reshape operator mapping 3D fluorescent distribution function u to discretized 1D fluorescent distribution vector x. Then, Eq. (6) can be written in terms u of as

$$\min_u \|y - MT(u)\|_2^2 + \lambda^2 \int |\nabla u| \mathrm{d}\Omega, \qquad (9)$$

where λ is a scalar variable that controls the strength of regularization. Equation (9) is a nonlinear minimization problem in u and we will use the gradient descent method [27] to solve it iteratively. Hence, we set the gradient of the objective equal to the difference between successive iterations divided by the step size Δt:

$$\frac{u^{(k+1)} - u^{(k)}}{t} = 2\,T * M * (y - MT(u)) + \lambda^2 \nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right). \qquad (10)$$

The finite difference method is applied to Eq. (10) to discretize u and solve for it numerically. The

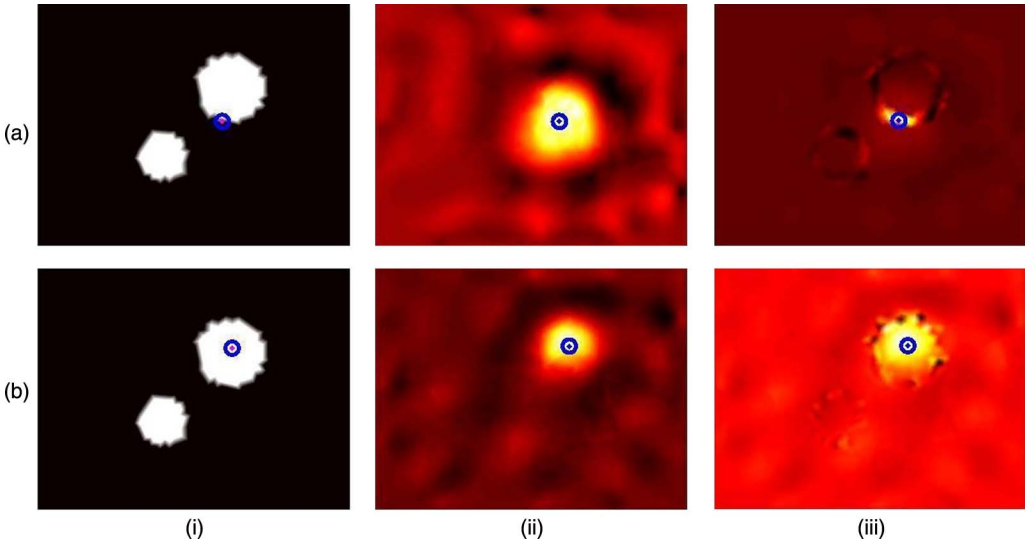

Fig. 1. (Color online) Comparison of PSF of L_2 regularization versus TV regularization for (a) on-edge and (b) off-edge voxels: (i) original fluorescence distribution, (ii) PSF for L_2 regularization of a moderate level, and (iii) PSF for TV regularization of a moderate level.

volume over which u is defined is discretized by a uniform rectangular mesh, u and its spatial derivatives are approximated by finite difference equivalents [27]. To ensure fast convergence, the initial guess to this iterative scheme is set to be the solution of L_2 regularization. Further details regarding the numerical implementation of this algorithm are provided in Appendix B. The choice of λ and Δt is fundamental to the convergence of the algorithm. $\Delta t(t)$ denotes the time step in the numerical solver and should be chosen to obtain the highest possible rate of convergence to the steady-state solution while meeting certain constraints required for maintaining numerical stability, e.g., Δt must be small enough to satisfy the Courant–Friedrichs–Lewy condition [30]. λ scales the trade-off between the data fidelity and regularization. To ensure best performance, we choose these parameters empirically; we perform a search in the interval ranging from the lowest to the highest singular values of the system matrix M [3], and similar to the L-curve method [31], select the regularization parameter that offers the best trade-off between the data fidelity and the solution norm [31]. Optimal selection of the regularization parameter is crucial to the successful performance of the proposed TV regularization method. If λ is too small, the problem will be under-regularized and the reconstructions will be artifact-contaminated. If λ is too large, the problem will be over-regularized and the accuracy of the reconstructions will be excessively jeopardized.

In the second approach, we use the recently explored split Bregman iteration [28,29] to solve the minimization problem in Eq. (8). Detailed derivation of this heuristic algorithm for TV minimization can be found in [32]. Unlike L_2 regularization, the optimization problem describing TV regularization may possess several local minima, only some of which yield the desirable solution. Therefore, under certain conditions, an iterative solver may converge to the undesirable local minima. In the ROF model, we avoided this by using the solution to L_2 regularization as an initial guess. For Bregman iteration we cannot use an initial guess; therefore we need to modify the algorithm to bias it away from undesirable local minima and toward the desired solution. We do this by adding an extra term to the objective of the minimization that biases it toward a good initial guess: the solution to L_2 regularization, u_{L2}. The modified biased minimization problem can be written as

$$\min_u \int |\nabla u| \mathrm{d}\Omega + \frac{\mu}{2}\|MT(u) - y\|_2^2 + \frac{\alpha}{2}\|u - u_{L2}\|_2^2, \tag{11}$$

where, for the purposes of simplicity in the final form of the algorithm, the regularization parameter λ is removed and instead a positive weight $\mu/2$ has been adopted for data fidelity. The relaxation parameter α controls the strength of the bias term that guides the algorithm to the vicinity of the least-squares solution to avoid convergence to unwanted local minima of the TV regularization. To solve the optimization in Eq. (11), split Bregman algorithm uses the concept of subgradient space [28,29,32] to determine the direction and magnitude of descent and repeats this iteratively to converge to the solution. Mathematically, the minimization in Eq. (11) is relaxed to the following iterative updating scheme by introducing auxiliary variables b^k and d^k, which lie in the subgradient [28,29],

$$u^{k+1} = \min_u \frac{\mu}{2}\|MT(u) - y\|_2^2 + \frac{\alpha}{2}\|u - u_{L2}\|_2^2 + \frac{\beta}{2}\|d^k - \Delta u - b^k\|_2^2, \tag{12}$$

where β is the relaxation weight of the subgradient term and determines the direction of the descent associated with each iteration. A large β would stipulate a descent in a direction very close to the gradient at the expense of jeopardizing the data fidelity. α, β, and μ are selected empirically by performing an exhaustive search (similar to the approach described for the regularization parameter λ). By setting the derivative of Eq. (12) to zero, we obtain

$$(\mu T * M * MT + \alpha I - \beta)u^{k+1} = \mu T * M * y + \alpha u_{L2} + \beta \nabla * (d^k - b^k), \tag{13}$$

where Δ denotes the Laplacian operator. Equation (13) along with updates for the auxiliary variables [32] describe each iteration of our proposed biased split Bregman algorithm for TV regularization applied to FT reconstruction. The system of equations in Eq. (13) is symmetric positive definite; therefore we solve it using the preconditioned conjugate gradient method [33]. Further details regarding numerical implementation of this algorithm are provided in Appendix B.

4. Numerical Studies

To compare the performance of our proposed method with L_2 regularization and row-action iterative reconstruction methods, we apply them to 2D simulated FT data and compare the results. Data are contaminated with different levels of noise. Figure 2 depicts the 2D simulated FT configuration. The two white blobs represent fluorescent inclusions with unit quantum efficiency and absorption of $\mu_{\mathrm{fl}} = 0.1\ \mathrm{mm}^{-1}$ in an 8 cm by 6 cm rectangular turbid medium with absorption of $\mu_a = 0.01\ \mathrm{mm}^{-1}$ and scattering of $\mu_s' = 1\ \mathrm{mm}^{-1}$, which mimic the optical properties of biological tissue. The blue squares and red circles represent sources and detectors. We have eight sources and eight detectors so the measurement vector y has 64 entries. We use a forward solver that simulates the propagation of diffuse light in turbid medium using finite element formulation of the diffusion equation expressed in Eq. (1) to obtain y. We contaminate y by three different typical

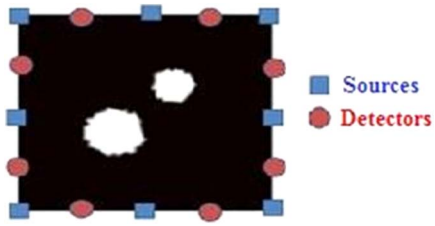

Fig. 2. (Color online) Simulated 2D fluorescence tomography configuration: two fluorescent blobs in a 2D turbid slab with eight sources and eight detectors around it used for excitation and data acquisition.

levels of additive white Gaussian noise to get 50, 40, and 30 dB signal-to-noise ratios (SNRs) in the data. The vector y is then used as the input for reconstruction using different approaches. Figure 3 depicts the 3D reconstructions for L_2 regularization, ART, and the two implementations of the proposed TV regularization. For the three studies, the relaxation parameters in the split Bregman iteration were optimized to be $\alpha = 520$, and $\beta = 2.1$. The regularization parameters in the two implementations of the TV regularization were selected based on the data SNR. The three sets of the regularization parameters used in the simulation studies were $\lambda = 4.8 \times 10^{-6}$, 8.9×10^{-5}, 9.1×10^{-4}, and $\mu = 1.4 \times 10^{3}$, 1.0×10^{3}, 7.5×102, for 50, 40, and 30 dB SNRs, respectively. It should be noted that the optimal relaxation parameters (α and β) are robust and do not depend on the data SNR, whereas the optimal regularization parameters (λ and μ) depend on the data SNR and must be optimized using an L-curve-type search [31] for each study.

The numerical studies presented in Fig. 3 reveal the effect of noise on the performance of different reconstruction algorithms. ART is an unbiased and nonregularized reconstruction technique, and thus, it does not converge to the solution for low SNR data. L_2 regularization retrieves the two blobs in all scenarios as depicted in Fig. 3(i), but the harm of its spreading and oversmoothing effects on the resolution of the reconstructed image is evident particularly for low SNR. As shown in Fig. 3(iii), ROF-based TV regularization reconstructs the blobs with high resolution in high SNR scenarios, while for low SNR scenarios, artifacts show up in the form of edge distortion and shifting. Split Bregman-based TV regularization also performs well for high SNR data and demonstrates a high resolving power, but for low SNR data, as shown in Fig. 3(iv), artifacts appear in the reconstructions and distort the shapes of the reconstructed blobs.

The least squares relative estimation error as defined in Eq. (14) is calculated for each reconstruction algorithm and plotted in Fig. 4:

$$\varepsilon = \frac{\|x - \hat{x}\|_2}{\|x\|_2}, \tag{14}$$

where $\hat{x}$ represents the reconstructed solution. We maintain that this error does not reflect the edge-preserving advantage of reconstruction algorithms as effectively as visual inspection of the reconstructed inclusions. Also, we computed and plotted the Michelson contrast [34] for each reconstructed image, as formulated below, to provide a quantitative analysis on the resolution improvements offered by the TV regularization:

$$C = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \tag{15}$$

where $I_{\max}$ represents the mean of the peak fluorophore concentrations in the two reconstructed blobs and $I_{\min}$ is the minimum fluorophore concentration

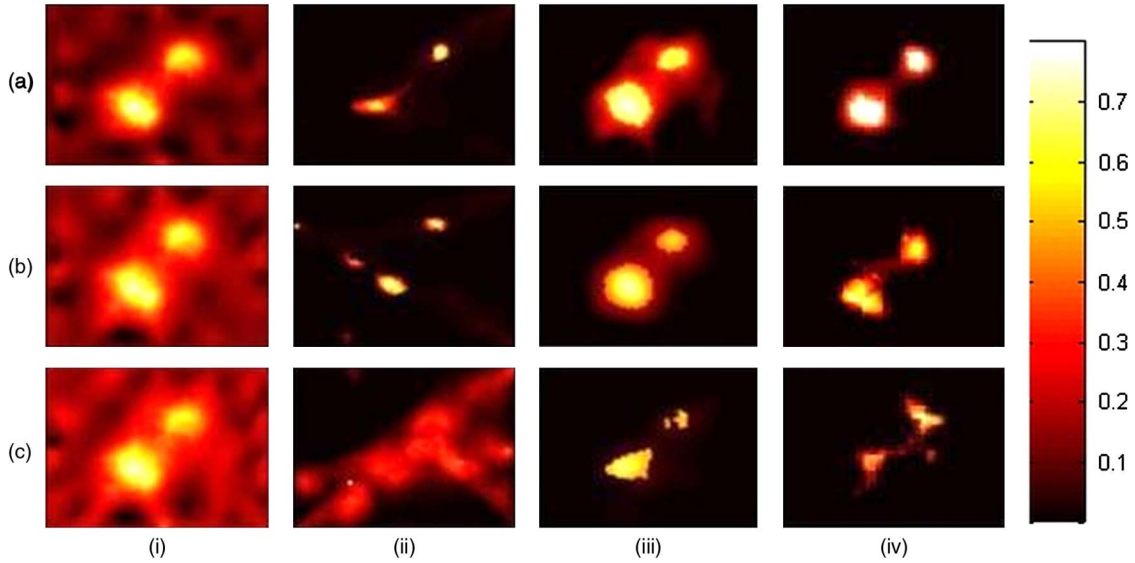

Fig. 3. (Color online) Reconstructed fluorophore distributions for 2D simulated data with (a) SNR = 50 dB, (b) SNR = 40 dB, and (c) SNR = 30 dB, by (i) L_2 regularization, (ii) algebraic reconstruction technique (ART), (iii) time marching ROF TV regularization, and (iv) iterative Bregman TV regularization.

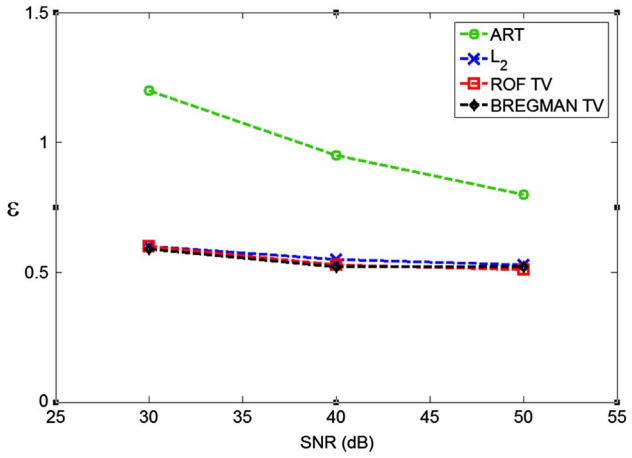

Fig. 4. (Color online) Relative estimation errors for reconstructed fluorescent distributions corresponding to L_2 regularization, ART, and the proposed ROF and iterative Bregman TV regularization for data SNR = 30, 40, and 50 dB.

on the line segment connecting the location of the peaks. Figure 5 depicts a plot of contrast values for each reconstructed image. As plotted in Fig. 4, the accuracy of the TV regularization reconstructions is better than ART and no less than the conventional L_2 regularization. Figure 5 demonstrates that the resolution offered by the TV regularization is higher compared to L_2 regularization due to its edge-preserving property. ART also provides a relatively high resolution for high-SNR data because of its weakly regularizing nature, but its performance degrades as the SNR is decreased, and for low-SNR data, TV regularization has higher resolution and accuracy compared to ART. Overall, Figs. 4 and 5 demonstrate that TV regularization is advantageous to the conventional ART and L_2 regularization algorithms and performs better in terms of accuracy and resolution compared to ART and L_2 regularization, respectively.

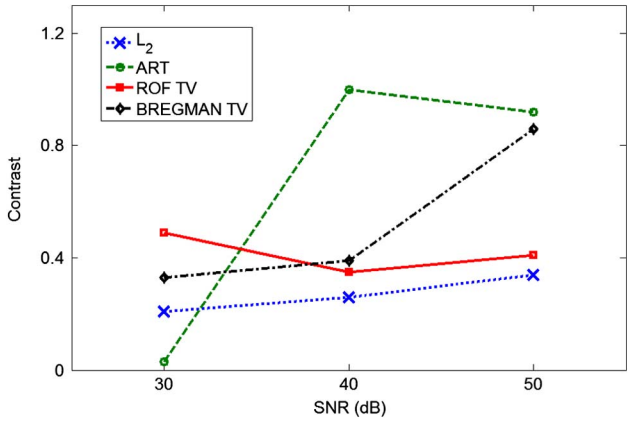

Fig. 5. (Color online) Michelson contrast [defined in Eq. (15)] is computed and plotted for the reconstructed fluorescent distributions corresponding to L_2 regularization, ART, and the proposed ROF and iterative Bregman TV regularization for data SNR = 30, 40, and 50 dB.

5. Experimental Results

We have used a noncontact constant-wave (CW) transillumination phantom-based FT system for validation of our proposed method. Figure 6 depicts a schematic and a picture of the experimental setup used in this study. A He-Ne 20 mW laser produces CW light at 632 nm whose power is adjusted and measured using neutral density filters and a power detector. The laser light is coupled into a multimode fiber whose tip lies on the tissue phantom to direct the light to a point on the tissue phantom surface. Translation stages are used to change the position of the tip of the fiber on the surface of the tissue phantom in order to illuminate it at different locations. Colored glass filters mounted on a motorized filter wheel are used to allow for separate sequential imaging of the transillumination signal at excitation wavelength and the fluorescent signal at the emission wavelength. A cooled CCD camera is then used to capture the filtered data images. The imaged fluorescent signal is normalized by the transillumination signal to calibrate the measurements into the same scale in a Born normalization fashion [35]. An intralipid-20% liquid tissue phantom [36] that mimics the optical properties of biological tissue is used for experimentation. India ink is added to diluted intralipid-20% with a concentration of 0.012% to match tissue absorption. The phantom has a scattering coefficient of 3.6 mm^{-1} and an absorption coefficient of 0.03 mm^{-1} at excitation and emission wavelengths. The phantom has a thickness of 14 mm, and a width and a length of 180 mm. We use 36 source

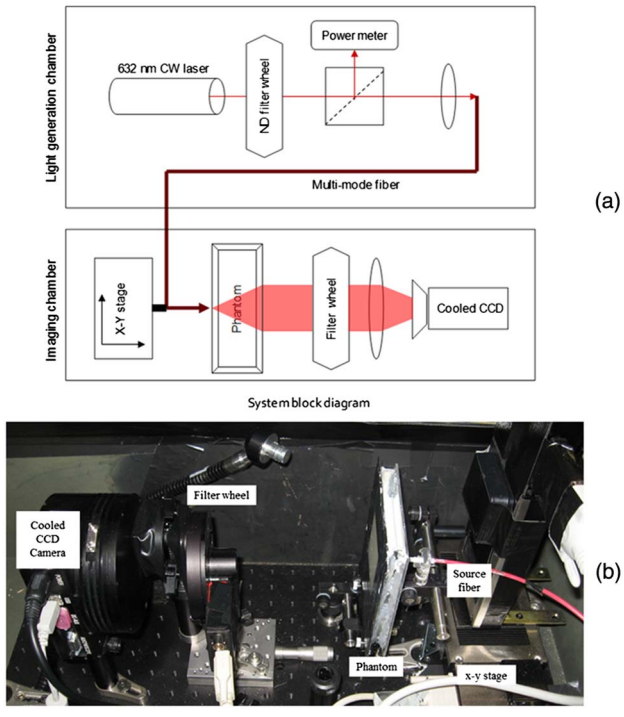

Fig. 6. (Color online) (a) Schematic diagram of the fluorescent tomography setup in the transillumination geometry. (b) Photograph of the interior of the imaging chamber.

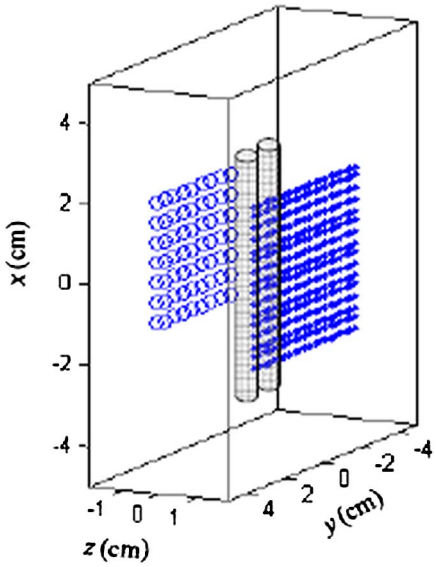

Fig. 7. (Color online) Configuration for the experimental phantom based fluorescence tomography. Two fluorescent tubes are inserted in an intralipid-20% liquid phantom that is excited at 36 source positions (circles) and imaged by a CCD camera that yields 81 data points (dots).

locations and 81 detector locations (as shown in Fig. 7) for illumination and data acquisition, respectively. A region of interest in the central part of the phantom with area of 8 cm × 8 cm is selected for reconstruction. A tetrahedral mesh with 153,945 nodes and 876,712 voxels is used to discretize the selected volume. The fluorescent dye used for experimentation is a dimethyl sulfoxide-based 100 μM solution of Oxazine 750 Perchlorate from Exciton, Inc. (excitation at 632 nm, emission at 700 nm). The dye solution is placed in two thin glass tubes with thickness of 1 mm. The tubes are immersed in the tissue phantom with vertical orientation at 3, 6, and 9 mm depth from the front surface of the phantom imaged by the cooled CCD camera. The camera is cooled down to −10°C to minimize dark current noise. Dark frame images (images with laser being off) are taken along with each measurement and subtracted from the data images to minimize stray light and other unwanted signals.

Figure 8 depicts the 3D reconstructions from experimental data using L_2 regularization and the two implementations for our proposed TV regularization. As the depth of the fluorescent inclusions increases, the measured emission signal becomes weaker and its SNR decreases. Therefore, stronger regularization is required for cases with deeper inclusions. Figure 8(a) depicts the performance of L_2 regularization in reconstructing the two fluorescent tubes at three different depths. While for lower depths with higher SNR, the reconstructions do not possess much spreading, the oversmoothing and spreading becomes very strong at higher depths with low SNR as the need for strong regularization increases. ROF-based TV regularization, as depicted in Fig. 8(b), performs better in reconstructing deeper inclusions (from low SNR data). The reconstructed tubes from

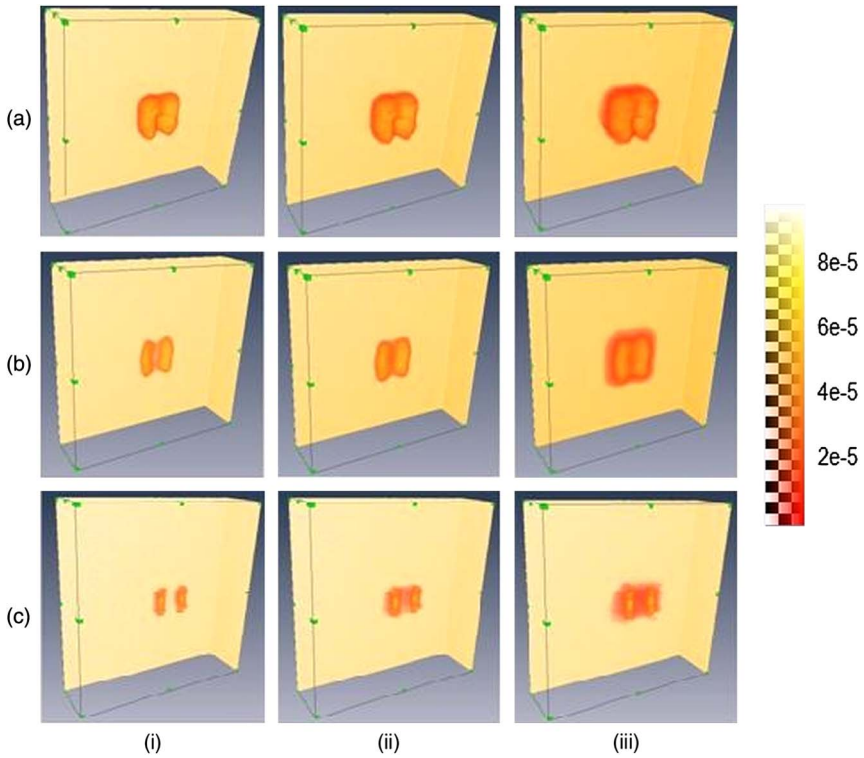

Fig. 8. (Color online) Reconstructed fluorophore distributions from experimental data where fluorophore tubes are located at (i) 3 mm, (ii) 6 mm, and (iii) 9 mm depth using (a) L_2 regularization, (b) time marching ROF-based TV regularization, and (c) iterative Bregman-based TV regularization.

ROF-based TV regularization are better resolved and localized around their corresponding positions. Reconstructions from Bregman-based TV regularization are depicted in Fig. 8(c). When compared to L_2 regularization, results from the Bregman-based TV regularization have less spreading, and the reconstructed tubes are distinctly separated. The results in Fig. 8 reveal the advantages of TV regularization over L_2 regularization in resolving fluorescent inclusions. While results from both methods possess artifacts, TV regularization shows a better promise in recovering and separating lesions and inclusions.

6. Discussion and Conclusions

The numerical and experimental results presented in this paper demonstrate that TV regularization has the potential of offering higher resolution and robustness compared to conventional L_2 regularization algorithms and ART. As depicted in Figs. 3 and 8, 2D and 3D reconstructions for both implementations of our proposed TV regularization algorithm are better localized around their corresponding coordinates and possess less spread. The nonspreading edge-preserving nature of TV regularization for ill-posed problems has been a subject of several studies [26,37]. Advantages of TV regularization over L_2 regularization are analogous in nature to advantages of wavelet-based de-noising [38] over low-pass filtering in image processing. While low-pass filtering can filter out the oscillatory noise in the image, it also diffuses the edges and sharp transitions in the image making it blurry and poorly resolved. De-noising through wavelet thresholding, on the other hand, only filters out highly oscillatory components in the image while preserving the edges and sharp transitions.

The optimization problem in TV regularization is not as easily tractable as L_2 regularization and hence the convergence time for TV regularization is higher in almost all cases presented in this paper. Among the two implementations presented, split Bregman iteration provides a faster convergence than ROF. ROF is a time-marching algorithm, and therefore it takes a hundred or more iterations to converge even when started with a good initial guess while split Bregman can converge within 20 iterations. However, the time required for each step of ROF is significantly less than that in split Bregman, since we need to implement a preconditioned conjugate gradient in each iteration of split Bregman. Depending on the regularization parameter, the preconditioned conjugate gradient may converge fast or slow, as its convergence depends mainly on the condition number [30] of the symmetric positive definite linear operator in Eq. (13). In other words, in strongly regularized cases we find that split Bregman iteration is quite faster than ROF while for weakly regularized cases they have the same convergence time order. The difference between ROF and split Bregman iteration implementation lies in the degrees of freedom associated with the gradient of the reconstructed image. Split Bregman iteration does not prefer a certain direction for the edges as it deals with each gradient component separately. However, the ROF implementation prefers circular and isotropic shapes as it deals with the magnitude of the gradient vector not the components separately. The ROF reconstructions tend to be more circular or cylindrical and thus isotropic, while split Bregman reconstructions have edges and borders of various shapes. While this property makes split Bregman implementation better capable of reconstructing various fluorescence maps, it makes it more prone to edge distortions resulting from noise and errors. This inherent difference between ROF and split Bregman is why the results of these two implementations are widely different in our numerical and experimental studies.

Based on the results reported in this work we conclude that for 3D reconstruction in FT, L_2 regularization is remarkably noise-robust and faster than TV regularization, but it offers poorer resolution. ROF-based TV regularization provides stable results with higher resolution than L_2 regularization for high SNR data, but its accuracy diminishes with increase in noise or error level. Split Bregman-based TV regularization performs faster than ROF and offers better resolution than L_2 regularization. However, for low SNR data its performance is not as robust as either L_2 regularization or ROF-based TV regularization. Therefore, for high SNR data split Bregman-based TV regularization can provide fast reconstructions with improved resolution while ROF-based TV regularization can be used for improvements in reconstruction resolution for lower SNR data.

In summary, we demonstrated two implementations of the TV regularization for FT. The main advantage of the TV regularization over more conventional L_2 regularization is that it does not need to sacrifice resolution for stability. L_2 regularization methods compromise resolution for stability; weakly regularized solutions have artifacts such as oscillations and impulses around the reconstructed maps, and the strongly regularized solutions have smooth overspread reconstructed maps as depicted in Figs. 3(i) and 7(a). In TV regularization, we observe that the edge preserving nature of the algorithm does not allow spreading in the solution. However, the locations of the edges in the solution may move as noise and error levels increase. In other words, strongly TV regularized solutions from noisy data with relatively low SNR (SNR < 30 dB) do not possess spreading and retain edges and sharp transitions while the presence of noise can affect the location of the edges and make them deviate from the original distribution. This can be seen particularly in Figs. 3(iii) and 3(iv). Therefore, the types of artifacts present in TV regularization for FT are inherently different than L_2 regularization. While damping out artifacts in L_2 regularization results in poor resolution, in TV regularization it results in distortion of the shape and geometry of the reconstructed fluorescence maps. In addition, the form

of noise-induced distortions in the solution of TV regularization differs between the two implementations. The ROF-based implementation prefers circularly shaped reconstructions and is thereof less prone to edge distortion, as depicted in Figs. 3(iii) and 8(b). Split Bregman iteration is more affected byedge distortions and especially for low SNR measurements, the noise-induced edge distortions may severely impair the shape and borders of the reconstructed maps as depicted in Figs. 3(iv) and 8(c). The resolution offered by these two implementations, however, is higher than L_2 regularization for high or low SNR measurements, as TV regularization does not compromise resolution in securing stability.

Appendix A: Finite Element Modeling

Two coupled diffusion equations describe the dynamics of a fluorescence tomographic scan, one for the diffusion of excitation photons and one for the diffusion of fluorescent photons. The coupled equations are formulated as below,

$$-\nabla \cdot D(r)\nabla\Phi_{\text{exc}}(r) + \mu_a(r)\Phi_{\text{exc}}(r) = q_{\text{exc}}(r), \quad \text{(A1)}$$

$$-\nabla \cdot D(r)\nabla\Phi_{\text{em}}(r) + \mu_a(r)\Phi_{\text{em}}(r) = \eta\mu_{\text{fl}}c(r)\Phi_{\text{exc}}(r), \quad \text{(A2)}$$

where $\Phi_{\text{exc}}(r)$ is the average fluence of excitation photons at location r, $q_{\text{exc}}(r)$ is the power density of the excitation laser or LED source used for illumination of the tissue at location r (as a result, $q_{\text{exc}}(r)$ is zero inside the tissue and non-zero at the boundary source locations), $\Phi_{\text{em}}(r)$ is the average fluence of the fluorescent light at location r, η is the dimensionless quantum efficiency of the fluorescent dye, μ_{fl} is the per molar fluorescent absorption coefficient, and $c(r)$ is the molar concentration of the fluorescent dye at location r. As formulated in the right-hand side of Eq. (A2), the source intensity term for the emission diffusion equation is the product of the quantum efficiency, the per molar fluorescent dye absorption, the fluorescent dye concentration, and the excitation fluence. The boundary conditions accompanying Eqs. (A1) and (A2) are modified Robin-type boundary conditions [39] expressed as

$$\Phi(\xi) + 2AD(\xi)\hat{n} \cdot \nabla\Phi(\xi) = 0, \quad \text{(A3)}$$

where ξ is any given point on the boundary, $\hat{n}$ is the normal vector to the boundary surface at ξ and A is a dimensionless constant that accounts for the index mismatch and internal reflection at the boundary.

Mathematically, the fluence function $\Phi(r)$ is approximated by the FE basis functions denoted ψ_j, $j = 1, \ldots, N$, (where N is the number of mesh nodes and hence the number of basis functions) as follows,

$$\Phi^h(r) = \sum_1^N \Phi_j\psi_j(r), \quad \text{(A4)}$$

where Φ_j denotes the weight of the j-th basis function $\psi_j(r)$ in the expansion and $\Phi^h(r)$ is the projection of $\Phi(r)$ onto the space spanned by the FE basis function. In applying the Galerkin approach to diffusion equation, the weak formulation of diffusion equation can be derived for each node indexed $j = 1, \ldots, N$ as below [39]

$$\int \psi_j(r)(-\nabla \cdot D(r)\nabla + \mu_a(r))\Phi^h(r) = \int \psi_j(r)q(r). \quad \text{(A5)}$$

Integration by parts and substitution of $\Phi^h(r)$ from Eq. (A4) transforms Eq. (A5) to the following discrete matrix equation [39]

$$[K(D) + C(\mu_a)]\Phi = Q - \beta, \quad \text{(A6)}$$

where

$$K_{ij} = \int D(r)\nabla\psi_i(r) \cdot \nabla\psi_j(r)\mathrm{d}\Omega, \quad \text{(A7)}$$

$$C_{ij} = \int \mu_a(r)\psi_i(r)\psi_j(r)\mathrm{d}\Omega, \quad \text{(A8)}$$

$$\beta_i = \int \psi_i(r)\Gamma(r)\mathrm{d}(\delta\Omega), \quad \text{(A9)}$$

$$Q_i = \int \psi_i(r)q(r)\mathrm{d}\Omega, \quad \text{(A10)}$$

and $d\Omega$ and $d(\delta\Omega)$ denote the differential elements for volume and boundary surface of the medium. Also, $\Gamma(r)$ is the surface exitance expressed mathematically as below and simplified using Robin boundary condition [39]

$$\Gamma(\xi) = -cD(\xi)\hat{n} \cdot \nabla\Phi(\xi) = \frac{c\Phi(r)}{2A}, \quad \text{(A11)}$$

which when substituted in Eq. (A6) yields

$$[K(D) + C(\mu_a) + F(A)]\Phi = Q, \quad \text{(A12)}$$

where

$$F_{ij} = \frac{-c}{2A}\int \psi_i(r)\psi_j(r)\mathrm{d}(\delta\Omega). \quad \text{(A13)}$$

Therefore, the expansion coefficients populated in the vector Φ can be obtained by solving the linear system expressed in Eq. (A11) and then substituted in Eq. (A5) to get $\Phi^h(r)$ which approximates the desired $\Phi(r)$. In the case of FT, where there are two coupled diffusion equations for every source position, the FE formulation yields two matrix equations as below

$$Z_e \Phi_e^{(i)} = Q_e^{(i)}, \tag{A14}$$

$$Z_m \Phi_m^{(i)} = Q_m^{(i)}, \tag{A15}$$

where $\Phi_e^{(i)}$ ($\Phi_m^{(i)}$) and $Q_e^{(i)}$ ($Q_m^{(i)}$) are fluence and source vectors at the excitation (emission) wavelength, respectively, when i-th source location is used. In FT only one source location is illuminated at a time. The matrix Z_e (Z_m) is the system matrix obtained from the finite element method (FEM) formulation at the excitation (emission) wavelength. The excitation source vector, $Q_e^{(i)}$, is nonzero only on the mesh nodes that neighbor the i-th point source location. Equation (A14) is the discrete FEM version of the forward problem and its solution for any given source location constitutes a Green's function.

The emission source vector, $Q_m^{(i)}$, can be mathematically expressed as follows:

$$Q_m^{(i)} = \eta \mu_{\mathrm{fl}} \mathrm{diag}(\Phi_e^{(i)}) x, \tag{A16}$$

where for any $N \times 1$ vector g, diag(g) is defined as an $N \times N$ diagonal matrix with elements of g populating its diagonal entries. Also, x represents the projection of fluorophore concentration function $c(r)$ onto the FEM basis functions and therefore lists the fluorophore concentration at each node of the mesh. Let N_s and N_d denote the number of source locations and detector locations, respectively. Hence for every source location, there are N_d measurements of $\Phi_m^{(i)}$ on the boundary of the medium. Let y denote the $N_s N_d \times 1$ vector that lists the boundary measurements of $\Phi_m^{(i)}$ for all source locations $i = 1, \ldots, N_s$. From Eqs. (A14), (A15), and (A16), we have

$$y = \begin{bmatrix} \bar{Z}_m^{-1} \eta \mu_{\mathrm{fl}} \mathrm{diag}(Z_e^{-1} Q_e^{(1)}) \\ \vdots \\ \bar{Z}_m^{-1} \eta \mu_{\mathrm{fl}} \mathrm{diag}(Z_e^{-1} Q_e^{(N_s)}) \end{bmatrix} x, \tag{A17}$$

where $\bar{Z}_m^{-1}$ is a submatrix of Z_m^{-1} that only includes the N_d rows that correspond to the detector locations. As a result, Eq. (A17) establishes a linear relationship between boundary detector measurements of emitted fluorescent signal and fluorophore distribution in the turbid medium. Considering detector noise, shot noise, and modeling errors present in FT, the linear model in Eq. (A16) can be expressed as below:

$$y = Mx + n, \tag{A18}$$

where e is the additive $N_s N_d \times 1$ error vector encompassing the modeling errors and noise, M is the $N_s N_d \times K$ system matrix (K being the number of mesh nodes) formulated in Eq. (A17), x is the $K \times 1$ fluorophore concentration vector, and y is the $N_s N_d \times 1$ measurement vector. Solving for x in Eq. (A18) from measurements of y, prior knowledge of M, and statistical properties of n constitutes the inverse problem of FT.

Appendix B: Computational Details of TV Regularization Algorithms

Iterative ROF TV regularization algorithm pseudo-code is as follows:

1. *Initialize* $u^{(0)} = T * (M * M + \lambda^2 I)^{-1} M * y$
2. *While* $\frac{\|u^{n+1} - u^n\|}{\|u^n\|} < \mathrm{eps}$
   - 2.1 *Map u to a cubic mesh by interpolation*
   - 2.2 *Approximate* $[u_x u_y u_z u_{xy} u_{xx} u_{yy} u_{xz} u_{yz} u_{zz}]$ *using finite difference*
   - 2.3 *Set* $d = u_x^2 u_{yy} + u_x^2 u_{zz} + u_y^2 u_{xx} + u_y^2 u_{zz} + u_z^2 u_{xx} + u_z^2 u_{yy} - 2(u_x u_y u_{xy} + u_x u_z u_{xz} + u_z u_y u_{zy}) / (u_x^2 + u_y^2 + u_z^2)^{1.5}$
   - 2.4 *Map u to a tetrahedral meshby interpolation*
   - 2.5 $u^{(n+1)} = u^{(n)} + \Delta t (T * M * (y - MT(u)) + \lambda d)$
   - 2.6 $u^{(n+1)} = \max(0, u^{(n+1)})$
3. *End*
4. $x^{\mathrm{final}} = T u^{(n+1)}$

Iterative split Bregman TV regularization algorithm pseudo-code is as follows:

1. $u_{L2} = T * (M * M + \lambda^2 I)^{-1} M * y$
2. *Initialize* $u^0 = T * M * y$
3. *While* $\frac{\|u^{k+1} - u^k\|}{\|u^k\|} < \mathrm{eps}$
   - 3.1 $u^{k+1} = PCG(\mu T * M * MT(u) + \alpha I - \beta \Delta, \mu T * M * f + \alpha u_{L2} + \beta \nabla_x^T (d_x^k - b_x^k) + \beta \nabla_y^T (d_y^k - b_y^k) + \beta \nabla_z^T (d_z^k - b_z^k))$
   - 3.2 *Map u to a cubic mesh by interpolation*
   - 3.3 $s^k = \sqrt{|b_x^k + \nabla_x u^k|^2 + |b_y^k + \nabla_y u^k|^2 + |b_z^k + \nabla_z u^k|^2}$
   - 3.4 $d_x^{k+1} = \max(s^k - \frac{1}{\lambda}, 0) \frac{\nabla_x u^k + b_x^k}{s^k}$
   - 3.5 $d_y^{k+1} = \max(s^k - \frac{1}{\lambda}, 0) \frac{\nabla_y u^k + b_y^k}{s^k}$
   - 3.6 $d_z^{k+1} = \max(s^k - \frac{1}{\lambda}, 0) \frac{\nabla_z u^k + b_z^k}{s^k}$
   - 3.7 $b_x^{k+1} = b_x^k - d_x^{k+1} + \nabla_x u^{k+1}$
   - 3.8 $b_y^{k+1} = b_y^k - d_y^{k+1} + \nabla_y u^{k+1}$
   - 3.9 $b_z^{k+1} = b_z^k - d_z^{k+1} + \nabla_z u^{k+1}$
   - 3.10 *Map u to a tetrahedral mesh by interpolation*
4. *End*
5. $x^{\mathrm{final}} = T u^{(k+1)}$

References

1. V. Ntziachristos, "Fluorescence molecular imaging," Annu. Rev. Biomed. Eng. **8**, 1–33 (2006).
2. V. Ntziachristos, C. Bremer, E. E. Graves, J. Ripoll, and R. Weissleder, "*In vivo* tomographic imaging of near-infrared fluorescent probes," Mol. Imaging **1**, 82–88 (2002).
3. P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion* (SIAM, 1997).
4. S. C. Davis, H. Dehghani, J. Wang, S. Jiang, B. W. Pogue, and K. D. Paulsen, "Image-guided diffuse optical fluorescence tomography implemented with Laplacian-type regularization," Opt. Express **15**, 4066–4082 (2007).
5. A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, and S. R. Arridge, "Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans," Opt. Express **15**, 6696–6716 (2007).
6. M. J. Eppstein, D. J. Hawrysz, A. Godavarty, and E. M. SevickMuraca, "Three-dimensional, Bayesian image reconstruction from sparse and noisy data sets: near-infrared

fluorescence tomography," Proc. Natl. Acad. Sci. USA **99**, 9619–9624 (2002).
7. Y. Lin, H. Yan, O. Nalcioglu, and G. Gulsen, "Quantitative fluorescence tomography with functional and structural a priori information," Appl. Opt. **48**, 1328–1336 (2009).
8. A. X. Cong and G. Wang, "A finite-element-based reconstruction method for 3D fluorescence tomography," Opt. Express **13**, 9847–9857 (2005).
9. X. Song, D. Wang, N. Chen, J. Bai, and H. Wang, "Reconstruction for free-space fluorescence tomography using a novel hybrid adaptive finite element algorithm," Opt. Express **15**, 18300–18317 (2007).
10. J. C. Baritaux, K. Hassler, and M. Unser, "An efficient numerical method for general L_p regularization in fluorescence molecular tomography," IEEE Trans. Med. Imag. **29**, 1075–1087 (2010).
11. P. Mohajerani, A. A. Eftekhar, J. Huang, and A. Adibi, "Optimal sparse solution for fluorescent diffuse optical tomography: theory and phantom experimental results," Appl. Opt. **46**, 1679–1685 (2007).
12. D. Han, J. Tian, S. Zhu, J. Feng, C. Qin, B. Zhang, and X. Yang, "A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization," Opt. Express **18**, 8630–8646 (2010).
13. D. Han, X. Yang, K. Liu, C. Qin, B. Zhang, X. Ma, and J. Tian, "Efficient reconstruction method for L1 regularization in fluorescence molecular tomography," Appl. Opt. **49**, 6930–6937 (2010).
14. J. C. Baritaux, K. Hassler, M. Bucher, S. Sanyal, and M. Unser, "Sparsity-driven reconstruction for FDOT with anatomical priors," IEEE Trans. Med. Imag. **30**, 1143–1153 (2011).
15. R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography," J. Theor. Biol. **29**, 471–481 (1970).
16. R. Schultz, J. Ripoll, and V. Ntziachristos, "Experimental fluorescence tomography of tissues with noncontact measurements," IEEE Trans. Med. Imag. **23**, 492–500 (2004).
17. X. Intes, V. Ntziachristos, J. P. Culver, A. Yodh, and B. Chance, "Projection access order in algebraic reconstruction technique for diffuse optical tomography," Phys. Med. Biol. **47**, N1–N10 (2002).
18. M. Freiberger, C. Clason, and H. Scharfetter, "Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting approach," Appl. Opt. **49**, 3741–3747 (2010).
19. J. Dutta, S. Ahn, C. Li, S. R. Cherry, and R. M. Leahy, "Joint L1 and total variation regularization for fluorescence molecular tomography," Phys. Med. Biol. **57**, 1459–1476 (2012).
20. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, 1978).
21. H. Jiang, "Frequency-domain fluorescent diffusion tomography: a finite-element based algorithm and simulations," Appl. Opt. **37**, 5337–5343 (1998).
22. C. L. Lawson and R J. Hanson, *Solving Least Squares Problems* (Prentice-Hall, 1974).
23. J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction: Spatial-invariant tomographs," IEEE Trans. Image Process. **9**, 1346–1358 (1996).
24. H. Gao and H. K. Zhao, "Multilevel bioluminescence tomography based on radiative transfer equation. Part 2: total variation and l1 data fidelity," Opt. Express **18**, 2894–2912 (2010).
25. P. Kisilev, M. Zibulevsky, and Y. Zeevi, "Wavelet representation and total variation regularization in emission tomography," in *2001 International Conference on Image Processing* (IEEE, 2001), Vol. **1**, pp. 702–705.
26. D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," Inverse Probl. **19**, S165–S187 (2003).
27. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D. **60**, 259–268 (1992).
28. T. Goldstein and S. Osher, "The Split Bregman method for L1-regularized problems," SIAM J. Imaging Sci. **2**, 323–343 (2009).
29. J. F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," SIAM J. Multisc. Model. Simul. **8**, 337–369 (2009).
30. R. Courant, K. Friedrichs, and H. Lewy, "Über die partiellen Differenzengleichungen der mathematischen Physik," Math. Ann. **100**, 32–74 (1928).
31. P. C. Hansen, "The *L*-curve and its use in the numerical treatment of inverse problems," in *Computational Inverse Problems in Electrocardiology*, P. D. Johnston, ed. (WIT Press, 2001), pp. 119–142.
32. X. Liu and L. Huang, "Split Bregman iteration algorithm for total bounded variation regularization based image deblurring," J. Math. Anal. Appl. **372**, 486–495 (2010).
33. C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations* (SIAM, 1995).
34. A. Michelson, *Studies in Optics* (University of Chicago, 1927).
35. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media using a normalized born approximation," Opt. Lett. **26**, 893–895 (2001).
36. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "A solid tissue phantom for photon migration studies," Phys. Med. Biol. **42**, 1971–1979 (1997).
37. R. Acar and C. R. Vogel, "Analysis of bounded variation penalty methods for ill-posed problems," Inverse Probl. **10**, 1217–1229 (1994).
38. S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," IEEE Trans. Image Proces. **9**, 1532–1546 (2000).
39. M. Schweiger, S. R. Arridge, M. Hiraoka, and D. T. Delpy, "The finite element method for the propagation of light in scattering media: boundary and source conditions," Med. Phys. **22**, 1779–1792 (1995).