Photonics **2014**, *1*, 95-109; doi:10.3390/photonics1020095

OPEN ACCESS

photonics

ISSN 2304-6732

www.mdpi.com/journal/photonics

Article

Comparison of Regularization Methods in Fluorescence Molecular Tomography

Dianwen Zhu [1], Yue Zhao [1], Reheman Baikejiang [1], Zhen Yuan [2] and Changqing Li [1,*]

[1] School of Engineering, University of California Merced, Merced, CA 95343, USA;
E-Mails: dzhu2010@gmail.com (D.Z.); yzhao32@ucmerced.edu (Y.Z.); bakijan@gmail.com (R.B.)
[2] Bioimaging Core, Faculty of Health Sciences, University of Macau, Av. Padre Tomás Pereira, Taipa,
Macau SAR, China; E-Mail: zhenyuan@umac.mo

[*] Author to whom correspondence should be addressed; E-Mail: cli32@ucmerced.edu;
Tel.: +1-209-228-4777.

Received: 1 February 2014; in revised form: 14 April 2014 / Accepted: 17 April 2014 /
Published: 29 April 2014

Abstract: *In vivo* fluorescence molecular tomography (FMT) has been a popular functional imaging modality in research labs in the past two decades. One of the major difficulties of FMT lies in the ill-posed and ill-conditioned nature of the inverse problem in reconstructing the distribution of fluorophores inside objects. The popular regularization methods based on L^2, L^1 and total variation (TV) norms have been applied in FMT reconstructions. The non-convex $L^q(0 < q < 1)$ semi-norm and Log function have also been studied recently. In this paper, we adopt a uniform optimization transfer framework for these regularization methods in FMT and compare their individual, as well as the combined effects on both small, localized targets, such as tumors in the early stage, and large targets, such as liver. Numerical simulation studies and phantom experiments have been carried out, and we found that L^q with q near 1/2 performs the best in reconstructing small targets, while joint L^2 and Log performs the best for large targets.

Keywords: reconstruction techniques; optics; fluorescence molecular tomography; optimization transfer

---

1. Introduction

Fluorescence molecular tomography (FMT) is a noninvasive functional imaging modality that has become popular in the past two decades. Its applications include monitoring molecular activities in cancer imaging, drug delivery, *etc.* [1,2]. FMT relies on the modeling of photon transportation at near-infrared wavelengths, which suffers from strong scattering. In addition, usually only a small amount of measurements are available. Therefore, the FMT reconstruction problem is highly ill-posed. Different regularization methods have been proposed to find meaningful solutions of this inverse problem. For example, Tikhonov regularization [3] (a special case of which is the L^2 regularization) has been used to incorporate structural prior information [4–6]. Since in FMT, the targets are usually localized [7,8], L^1 norm and its variants were also studied and were found to yield sparse solutions [9,10]. A comparative study on the L^2 and L^1 regularizations was made by [11], where they found that L^1 is suitable for small targets, while L^2 performs better for large targets. Due to the fact that most of the FMT targets are localized and smooth, total variation (TV) regularization was also employed to promote local smoothness [12]. Furthermore, a joint L^1 and TV regularization approach was taken in [13], where, however, no clear winner was identified in terms of quantitative metrics when the joint L^1 and TV method was compared to L^1, TV or L^2 alone, although the joint L^1 and TV method provided the most natural appearances on reconstructed images. Recently, non-convex $L^q\,(0 < q < 1)$ and Log functioned-based regularizations were also actively studied in optical tomography and were found to have greater power than L^1 in enhancing sparsity [9,14–16].

Inspired by [11], in this paper, we thoroughly compare all regularization methods: L^2, L^1, TV, L^q, Log methods and the combined methods from a smoothing L^2 or TV regularization and a localizing L^1, L^q or Log regularization, in hopes of identifying the best regularization method for FMT for both small target cases and large target cases. For a fair comparison, we use the same simulation settings for both of the small targets, such as tumors in the early stage, and large targets, such as livers or bladders. We also conducted phantom experiments for further validation of these algorithms. The outline of this paper is follows. In Section 2, we introduce the models, optimization transfer algorithms, image quality metrics, as well as simulation and phantom experiment setups. In Section 3, results, including both images and metrics, are detailed for the comparison. Finally, in Section 4, we conclude this paper with discussion.

2. Methodology

2.1. Forward Modeling

Fluorescence molecular tomography (FMT) detects photons on the surface of an object and aims to discover the 3D distribution of a fluorescent source inside the object [17]. Near-infrared fluorescent dye is a typical fluorescent source [18]. A photon's propagation is usually modeled by the radiative transport equation, which is an integro-differential equation that is hard to solve. The diffusion equation was found to be a good approximation [19]. The coupled diffusion equation in the continuous wave domain, along with Robin boundary conditions, can be expressed as:

$$\begin{cases} -\nabla \cdot (D_{ex}(\mathbf{r})\nabla\Phi_{ex}(\mathbf{r})) + \mu_{a,ex}(\mathbf{r})\Phi_{ex}(\mathbf{r}) = S_k(\mathbf{r}) \\ \mathbf{n} \cdot (D_{ex}(\mathbf{r})\nabla\Phi_{ex}(\mathbf{r})) + \alpha_{ex}\Phi_{ex}(\mathbf{r}) = 0 \\ -\nabla \cdot (D_{em}(\mathbf{r})\nabla\Phi_{em}(\mathbf{r})) + \mu_{a,em}(\mathbf{r})\Phi_{em}(\mathbf{r}) = \Delta_j(\mathbf{r}) \\ \mathbf{n} \cdot (D_{em}(\mathbf{r})\nabla\Phi_{em}(\mathbf{r})) + \alpha_{em}\Phi_{em}(\mathbf{r}) = 0 \end{cases} \tag{1}$$

where ∇ denotes the gradient operator, $D_{ex}(\mathbf{r}) = [3(\mu'_{s,ex}(\mathbf{r}) + \mu_{a,ex}(\mathbf{r}))]^{-1}$ and $D_{em}(\mathbf{r}) = [3(\mu'_{s,em}(\mathbf{r}) + \mu_{a,em}(\mathbf{r}))]^{-1}$, with $\mu_{a,ex}(\mathbf{r})$, $\mu_{a,em}(\mathbf{r})$ being the absorption coefficients and $\mu'_{s,ex}(\mathbf{r})$, $\mu'_{s,em}(\mathbf{r})$ being the reduced scattering coefficients at excitation and emission wavelengths, $\Phi_{ex}(\mathbf{r})$, $\Phi_{em}(\mathbf{r})$ the photon densities, $\mathbf{r}$ the location vector, S_k determined by the k^{th} illumination pattern, Δ_j determined by the detector nodes, $\mathbf{n}$ the outward unit normal vector of the boundary, and α_{ex}, α_{em} the Robin boundary coefficients.

The above equations can be solved using the finite element method (FEM) at both the excitation and emission wavelengths, where we arrived at a linear model [20]:

$$\mathbf{b} = \mathbf{A}\mathbf{x} \tag{2}$$

Here, $\mathbf{b} \in \mathbf{R}^{N_m}$ is the measurement of fluorescence intensity acquired from the surface of the object; $\mathbf{A} \in \mathbf{R}^{N_m \times N_n}$ is called the system matrix; $\mathbf{x} \in \mathbf{R}^{N_n}$ is the fluorophore distribution, which equals the product of the unknown fluorescence dye concentration and the quantum yield at each node to be reconstructed. Note that $N_m = N_d * N_s$ is the total number of measurements, N_d, N_s, N_n are the total number of detectors, excitation sources and FEM mesh nodes, respectively.

2.2. Non-Negative Regularized Least Squares

FMT image reconstruction is to solve Equation (2). No direct solution exists, due to the fact that $\mathbf{A}$ is severely ill-conditioned. Equation (2) is instead usually solved as a regularized least squares problem with a non-negativity constraint:

$$\hat{\mathbf{x}} = \min_{\mathbf{x} \geq 0} \Psi(\mathbf{x}) := \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \lambda \cdot \mathbf{R}(\mathbf{x}) \tag{3}$$

where $\mathbf{A}, \mathbf{x}, \mathbf{b}$ are the same as defined in Equation (2), and $\lambda \cdot \mathbf{R}(\mathbf{x})$ is the regularization term. For our comparison studies, we specify in this paper:

$$\lambda \cdot \mathbf{R}(\mathbf{x}) = \frac{1}{2}\lambda_2\|\mathbf{x}\|^2_{L^2} + \lambda_{TV}\|\mathbf{x}\|_{TV} + \lambda_1\|\mathbf{x}\|_{L^1} + \lambda_q\|\mathbf{x}\|^q_{L^q} + \lambda_{Log}\|\mathbf{x}\|_{Log} \tag{4}$$

where $\|\cdot\|_{L^2}$ and $\|\cdot\|_{L^1}$ are the well-known L^2 and L^1 norms,

$$\|\mathbf{x}\|_{TV} = \sum_{i=1}^{n} \sum_{j \in nb(x_i)} |x_i - x_j| := |C(\mathbf{x})| \tag{5}$$

with $nb(x_i)$ denoting all the neighboring points to node x_i, as defined in [13], and the semi-norms:

$$\|\mathbf{x}\|_{\log} = \sum_{i=1}^{n}[\log(|x_i| + \delta_{\log}) - \log\delta_{\log}] \quad \text{and} \quad \|\mathbf{x}\|^q_{L^q} = \sum_{i=1}^{n}(|x_i| + \delta_q)^q \tag{6}$$

as defined in [14]. To investigate an individual regularization method, we set the corresponding λ as nonzero and others as zero; to study the combination of two regularization methods, only the corresponding two λ's are set to be nonzero.

2.3. Optimization Transfer Algorithms

One way to solve the above optimization problem Equation (3) is to directly solve $\nabla_{\mathbf{x}}\Psi = 0$, with $\nabla_x\Psi$ being the gradient of Ψ with respect to $\mathbf{x}$. If L^2 is the only regularization to be considered, then $\nabla_{\mathbf{x}}\Psi$ becomes linear, and if we further disregard the non-negativity requirement, $\nabla_{\mathbf{x}}\Psi = 0$ can be solved by the well-known preconditioned conjugate gradient (PCG) algorithm. However, since the gradient of TV norm, L^q and Log semi-norms are nonlinear, one will have to follow the nonlinear version of PCG, which suffers a very expensive backtracking line search when the non-negativity constraint is enforced.

Another approach toward solving Equation (3) is to use the optimization transfer algorithm, also known as the majorization-minimization algorithm, which optimizes a surrogate function at each iteration. The surrogate function is usually chosen in a way such that its minimization is much easier and can even be analytically solved. This technique has been successfully employed to handle the TV term, as in [13], and also has been applied to cope with the Log and semi-norms in our previous work [14]. We will follow this optimization transfer principle and handle all the regularization methods in a uniform way. The algorithm is as follows:

```
Initialization;
(1) $\mathbf{x}^0 = x_0 * \mathbf{1}$, where $x_0$ is a random number in (0,1) ;
(2) Compute matrix $C$ for the $TV$ regularization;
(3) Compute $\kappa$, with component $\kappa_j = \sum_{i=1}^{N_d*N_s} \frac{a_{ij}^2}{\beta_{ij}}$, where $\beta_{ij} = \frac{a_{ij}}{\sum_{k=1}^{N_n} a_{ik}}$ ;
for $k = 1$ to $N_{max}$ do
    (a) $\kappa_{TV}(\mathbf{x}^k) = 2\lambda_{TV}|C^t| * \zeta(C\mathbf{x}^k)$, and $gradTV(\mathbf{x}^k) = \lambda_{TV}C^t * ((C\mathbf{x}^k).*\zeta(C\mathbf{x}^k))$ ;
    (b) $\mathbf{x}^{k+1} = ((\mathbf{x}^k - \mathbf{A}^t(\mathbf{A}\mathbf{x}^k - \mathbf{b}) + gradTV(\mathbf{x}^k))./(\kappa + \kappa_{TV}(\mathbf{x}^k)))_+$ ;
    (c) $\mathbf{x}^{k+1} = (\mathbf{x}^{k+1} - \lambda_{sp}./(\kappa + \kappa_{TV}(\mathbf{x}^k)))_+$ ;
    (d) $\mathbf{x}^{k+1} = \mathbf{x}^{k+1}./(1 + \lambda_2./(\kappa + \kappa_{TV}(\mathbf{x}^k)))$ ;
    if $\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_{L^2}/\|\mathbf{x}^k\|_{L^2} < \delta_{stop}$ then
        break;
    end
end
```

where $.*$ and $./$ are entry-wise vector multiplication and division respectively, $(\cdot)_+$ denotes the positive part of a vector,

$$\zeta(\mathbf{x}) = \left(\frac{1}{\sqrt{x_1^2+\delta_{TV}}}, \cdots, \frac{1}{\sqrt{x_n^2+\delta_{TV}}}\right), \text{ and } (\lambda_{sp})_j = \begin{cases} \lambda_1; & \text{if } \lambda_1 \neq 0; \\ \frac{\lambda_{Log}}{x_j^k+\delta_{Log}}; & \text{if } \lambda_{Log} \neq 0; \\ \frac{q\lambda_q}{(x_j^k+\delta_q)^{1-q}}, & \text{if } \lambda_q \neq 0, \end{cases}$$

with δ_{TV}, δ_{Log} and δ_q being some empirically chosen small nonnegative numbers, e.g., 1E-9. δ_{stop} is the stopping criteria, e.g., 0.1%. More detailed derivations of the above formulas can be found in [13,14]. All our codes were implemented with MATLAB (2013b).

Algorithm 1: Uniform Optimization Transfer Framework.

2.4. Regularization Parameter Selection and Image Quality Metrics

There are different criteria in optical tomography for quantifying image qualities. A total of 11 methods were listed and compared in [21]. Since no single criterion captures all the features of the reconstructed images, we base our comparison on a mixture of metrics. As proposed in our previous work [14], we will use visual assessment along with four metrics: the volume ratio (VR) [22], the dice similarity coefficient (Dice) [23], the mean squared error (MSE) and the contrast-to-noise ratio (CNR) [24]. Their definitions for a reconstructed image are as follows:

$$VR = \frac{|ROI|}{|tROI|}, \quad Dice = \frac{2*|ROI \cap tROI|}{|ROI|+|tROI|} \tag{7}$$

$$MSE = \frac{1}{N}\sum_{j=1}^{N}(x_j - x_{0j})^2 \tag{8}$$

$$CNR = \frac{Mean(x_{ROI}) - Mean(x_{ROB})}{\sqrt{\omega_{ROI} Var(x_{ROI}) + (1-\omega_{ROI}) Var(x_{ROB})}} \tag{9}$$

where ROI represents the region of interest or the voxels that have intensities greater than 50% of the maximum of the image intensities, ROB the rest voxels, $tROI$ the true target locations, x and x_0 the reconstructed and true image intensity at each voxel and $\omega_{ROI} = |ROI|/|\text{Total image}|$, $|\cdot|$ the number of voxels, *Mean* mean value and *Var* variance. VR measures the sparsity of the reconstructed target. Dice quantifies the shape and location accuracy. MSE is only calculated if the ground-truth is known, and CNR is about the reconstructed image only. Generally, for a reconstructed image to have higher quality than another, its VR and Dice should be closer to one, MSE smaller and CNR larger.

2.5. Numerical Simulations

Following our previous work [14], we conducted numerical simulations based on the Digimouse mesh [25]. In order to get a more uniform mesh, since we do not need all the details from small organs, we extracted the surface mesh and then used Tetgen (http://www.tetgen.org) to regenerate the internal mesh. Our mesh has a total of 32,332 nodes and 161,439 tetrahedral elements. For our first scenario with small targets, we simulated two capillary tubes within the mouse with diameters of 2 mm and lengths of 20 mm as the targets. Additionally, for the second scenario, we selected the liver of the Digimouse as the target. For all the nodes inside the targets, we set the fluorophore concentration to be 1; otherwise, 0. We then chose a total of 60 nodes as laser excitations sources, which are uniformly distributed in 5 rings on the trunk surface of the digimouse. All the 4,020 surface nodes that cover the trunk were set to be detectors. Finally, we added to our simulated measurement data a uniform white noise with a signal-to-noise ratio of 1. Note that the noise level will affect the selection of regularization parameters, which is supposed to balance the measurement error and modeling error. The excitation wavelength was chosen to be 650 nm and the emission wavelength 700 nm. The tissue optical properties are $\mu_a = 0.007\,\text{mm}^{-1}$, $\mu_s' = 0.72\,\text{mm}^{-1}$ at 650 nm and $\mu_a = 0.014\,\text{mm}^{-1}$, $\mu_s' = 0.78\,\text{mm}^{-1}$ at 700 nm.

2.6. Phantom Experiments

We used a cubic phantom to validate our algorithms in the scenario with small targets. The phantom is of a dimension of $32 \times 32 \times 29$ mm^3 and was composed of 1% intralipid, 2% agar and water in the background. The two capillary target tubes are 12 mm in length and 1 mm in diameter.

6.5 μM DiD fluorescence dye solution was injected as the fluorescence sources. Uniformly distributed [^{18}F]-fluoro-2-deoxy-D-glucose (FDG) at an activity level of 100 μCi was also included in the two target tubes for a simultaneous positron emission tomography (PET) scan, whose results will be used to validate our FMT results. The excitation laser at a wavelength of 650 nm scanned the front surface of the phantom at 20 illumination positions. The emission wavelength was 700 nm. The tissue optical properties are $\mu_a = 0.0022\,\text{mm}^{-1}$, $\mu_s' = 1.10\,\text{mm}^{-1}$ at both 650 nm and 700 nm wavelengths, which was determined by a least squares algorithm. Our FMT system used a conical mirror for optical measurement collection, and the conical mirror was inserted into a microPET system for simultaneous PET scans. Measurements from 1,057 detectors on four sides were collected by a CCD camera and then mapped onto the cube surfaces. For more details of the phantom experimental setup, optical property estimation, measurement collection and mapping, please refer to [20]. For the PET images, we threshold at 20% of the maximum FDG concentrations to identify the positions of the capillary tubes. More details on the PET set-up can be found in [26,27]. The FEM mesh for the cubic phantom was again generated by Tetgen and consisted of a total of 8,690 nodes and 47,581 tetrahedral elements.

3. Results

3.1. Simulation Results for Small Targets

The reconstructed images using individual regularizations are shown in Figure 1, and the quantitative image quality metrics (VR, Dice, CNR and MSE) are shown in Table 1. We see that L^2 and TV regularization produce larger targets, while L^1, L^q and Log produce sparser targets. Furthermore, we found that for $q = 1/2$, L^q produces the best overall result. The image has the fewest artifacts by visual assessment, and the corresponding MSE is the lowest, CNR second highest, while VR and Dice not too different from the best either. We also evaluated the combined regularizations using one of the sparsity enhancing regularizations (L^1, L^q or Log) and one of the smoothing regularizations (L^2 or TV). We have not seen any improvements on the image qualities, as indicated by Table 2.

Table 1. Metrics of the best reconstructed images using individual regularizations for simulated small targets, with the best regularization method highlighted in bold. **RegType**, and **Reg** λ are the regularization types and associated regularization parameters, respectively.

RegType	Reg λ	VR	Dice	CNR	MSE
L^2	$\lambda_2 = $ 5.0E-5	6.19	0.26	5.30	3.5E-3
TV	$\lambda_{TV} = $ 1.0E-11	6.05	0.27	5.36	3.5E-3
L^1	$\lambda_1 =$5.9E-4	4.59	0.27	4.99	3.6E-3
$L^{7/8}$	$\lambda_{7/8} = $ 6E-4	1.25	0.31	4.42	3.6E-3
$L^{5/8}$	$\lambda_{5/8} = $ 1.0E-4	3.49	0.34	6.08	3.1E-3
$\mathbf{L^{1/2}}$	$\boldsymbol{\lambda_{1/2}}$ **=6.4E-5**	**2.61**	**0.33**	**5.72**	**3.1E-3**
$L^{3/8}$	$\lambda_{3/8} = $ 3E-5	1.84	0.36	5.52	3.2E-3
$L^{1/8}$	$\lambda_{1/8} = $ 3.2E-6	2.13	0.36	5.38	3.3E-3
Log	$\lambda_{Log} = $ 1.3E-5	2.97	0.34	5.19	3.3E-3

Figure 1. Simulated numerical mouse with two small tubes as targets: coronal sections of the mouse showing (**a**) the simulated truth, and the reconstruction results using regularizations based on (**b**) L^2 with $\lambda_2 = 5.0\text{E-}5$, (**c**) total variation (TV) with $\lambda_{TV} = 1.0\text{E-}11$, (**d**) L^1 with $\lambda_1 = 5.9\text{E-}4$, (**e**) $L^{7/8}$ with $\lambda_{7/8} = 6.0\text{E-}4$, (**f**) $L^{5/8}$ with $\lambda_{5/8} = 1.0\text{E-}4$, (**g**) $L^{1/2}$ with $\lambda_{1/2} = 6.4\text{E-}5$, (**h**) $L^{3/8}$ with $\lambda_{3/8} = 3.0\text{E-}5$, (**i**) $L^{1/8}$ with $\lambda_{1/8} = 3.2\text{E-}6$ and (**j**) Log with $\lambda_{Log} = 1.3\text{E-}5$, respectively.

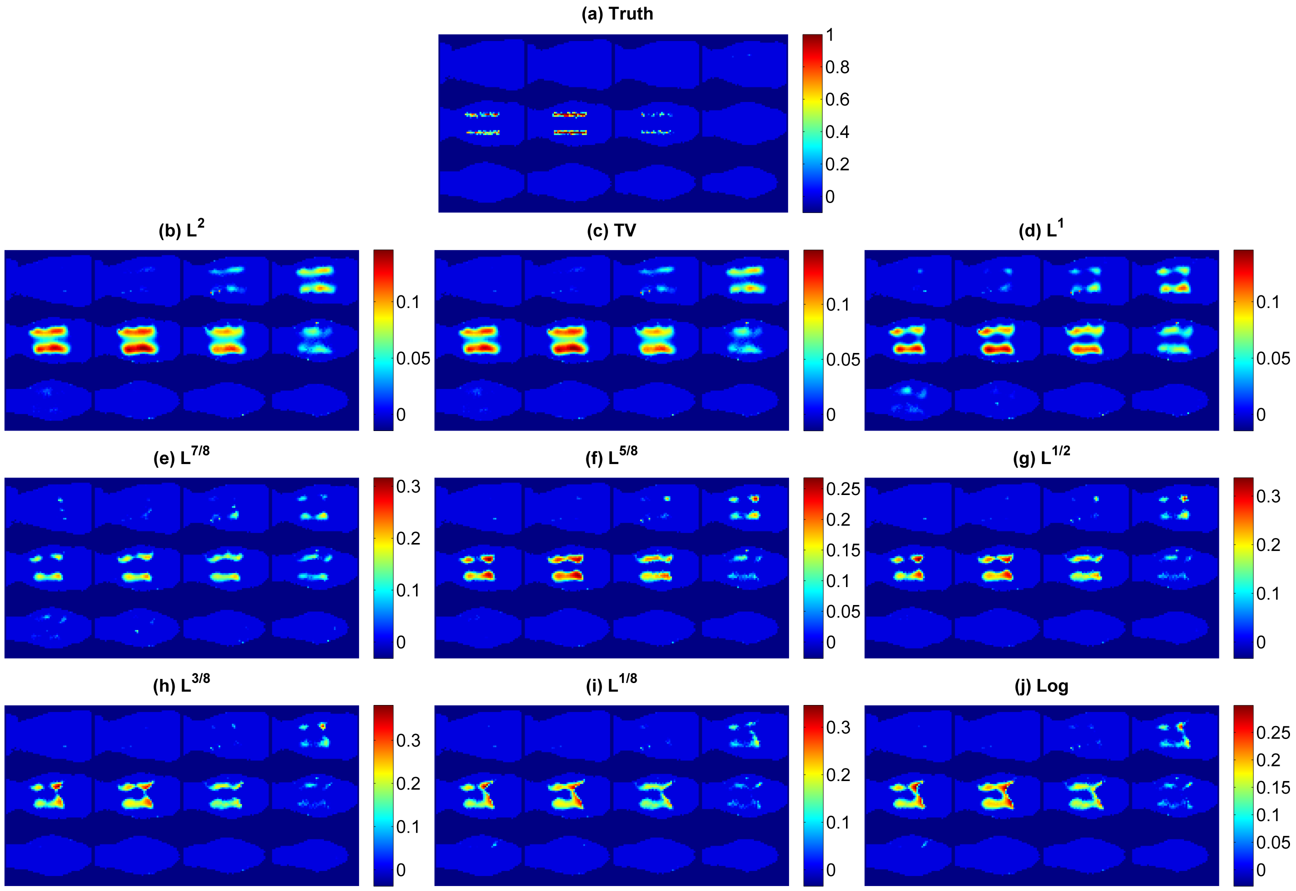


Table 2. Metrics of the best reconstructed images using combined regularizations for simulated small targets.

Reg Type	TV					L^2				
	Reg λ's	**VR**	**Dice**	**CNR**	**MSE**	**Reg λ's**	**VR**	**Dice**	**CNR**	**MSE**
$+L^1$	2.0E-11, 5.9E-4	4.60	0.27	4.99	3.6E-3	5.0E-5, 6.0E-10	4.63	0.26	4.94	3.6E-3
$+L^{7/8}$	3.0E-11, 6.0E-4	1.16	0.30	4.41	3.6E-3	4.9E-5, 6.1E-4	1.16	0.29	4.36	3.6E-3
$+L^{5/8}$	1.0E-11, 1.0E-4	3.45	0.34	6.07	3.1E-3	4.9E-5, 1.1E-4	3.61	0.33	6.04	3.1E-3
$+L^{1/2}$	1.0E-11, 6.4E-5	2.61	0.33	5.72	3.1E-3	4.9E-5, 6.4E-5	2.71	0.32	5.7	3.1E-3
$+L^{3/8}$	2.0E-11, 3.0E-5	1.84	0.36	5.52	3.2E-3	4.9E-5, 3.0E-5	1.99	0.35	5.51	3.2E-3
$+L^{1/8}$	1.0E-11, 3.2E-6	2.13	0.36	5.38	3.3E-3	4.9E-5, 3.2E-6	2.26	0.37	5.35	3.3E-3
$+Log$	3.0E-11, 1.4E-5	2.46	0.34	5.04	3.4E-3	4.9E-5, 1.4E-5	2.63	0.34	5.03	3.4E-3

3.2. Simulation Results for Large Target

For the large target case, the reconstructed images using different individual regularizations are shown in Figure 2 a–i, and the quantitative image quality metrics (VR, Dice, CNR and MSE) are shown in Table 3. Visually, we could only see small differences from the reconstructed images (Figure 2 a–i). Yet, by going through the detailed metrics, we clearly see that Log regularization is the best individual performer. We also examined the results from combined regularizations. Table 4 lists the metrics of the best images we obtained using the combined regularization methods. There are some small gains in image qualities compared with the individual methods. Additionally, the best performance is achieved when Log is combined with L^2. We included the best reconstructed imaging using $L^2 + Log$ and $TV + Log$ in Figure 2 j–k as a reference.

Figure 2. Simulated numerical mouse with its liver as target: coronal sections of the mouse showing reconstruction results using regularizations based on (**a**) L^2 with λ_2 = 1.0E-3, (**b**) TV with λ_{TV} = 1.0E-6, (**c**) L^1 with λ_1 = 1.0E-3, (**d**) $L^{7/8}$ with $\lambda_{7/8}$ = 1.0E-3, (**e**) $L^{5/8}$ with $\lambda_{5/8}$ = 1.0E-3, (**f**) $L^{1/2}$ with $\lambda_{1/2}$ = 1.0E-3, (**g**) $L^{3/8}$ with $\lambda_{3/8}$ = 5.0E-4, (**h**) $L^{1/8}$ with $\lambda_{1/8}$ = 1.0E-3, (**i**) Log with λ_{log} = 1.0E-3, (**j**) $TV + Log$ with λ_{TV} = 5.0E-7 and λ_{log} = 1.0E-3, (**k**) $L^2 + Log$ with λ_2 = 1.0E-3 and λ_{Log} = 1.0E-3, and (**l**) the simulated truth, respectively.

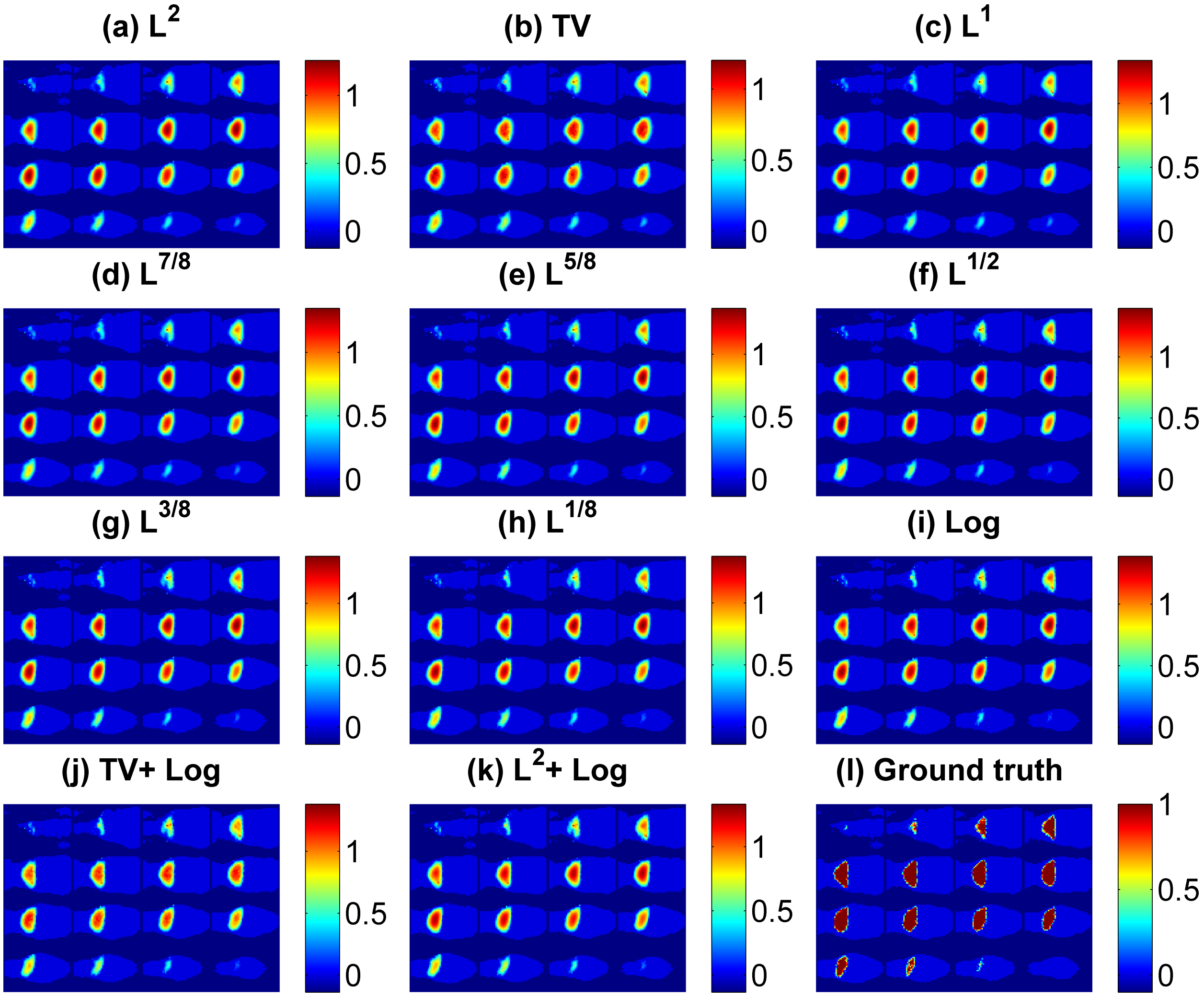


Table 3. Metrics of the best reconstructed images using individual regularizations for simulated large target, with the best regularization method highlighted in bold.

Reg Type	Reg λ	VR	Dice	CNR	MSE
L^2	$\lambda_2 = 1.0\text{E-}3$	0.51	0.68	7.3	9.2E-3
TV	$\lambda_{TV} = 1.0\text{E-}6$	0.50	0.67	7.3	8.9E-3
L^1	$\lambda_1 = 1.0\text{E-}3$	0.61	0.76	6.9	1.0E-2
$L^{7/8}$	$\lambda_{7/8} = 1.0\text{E-}3$	0.62	0.76	7.0	1.0E-2
$L^{5/8}$	$\lambda_{5/8} = 1.0\text{E-}3$	0.63	0.77	7.3	9.1E-3
$L^{1/2}$	$\lambda_{1/2} = 1.0\text{E-}3$	0.64	0.78	7.6	8.3E-3
$L^{3/8}$	$\lambda_{3/8} = 5.0\text{E-}4$	0.65	0.79	7.5	8.4E-3
$L^{1/8}$	$\lambda_{1/8} = 1.0\text{E-}4$	0.67	0.79	7.5	8.4E-3
Log	$\lambda_{\mathbf{Log}} = \mathbf{1.0E\text{-}3}$	**0.69**	**0.81**	**7.7**	**7.7E-3**

Table 4. Metrics of the best reconstructed images using combined regularizations for simulated large target, with the best regularization method highlighted in bold.

Reg Type	TV					L^2				
	Reg λ's	**VR**	**Dice**	**CNR**	**MSE**	**Reg λ's**	**VR**	**Dice**	**CNR**	**MSE**
$+L^1$	1.0E-6, 1.0E-5	0.74	0.83	6.74	1.2E-2	1.0E-3, 1.0E-3	0.68	0.81	6.95	1.1E-2
$+L^{7/8}$	5.0E-7, 1.0E-3	0.67	0.80	7.14	1.0E-2	1.0E-2, 1.0E-4	0.69	0.81	7.74	7.7E-3
$+L^{5/8}$	1.0E-7, 1.0E-3	0.64	0.78	7.37	8.9E-4	1.0E-3, 1.0E-3	0.68	0.81	7.35	9.1E-3
$+L^{1/2}$	5.0E-7, 5.0E-4	0.68	0.81	7.40	9.0E-3	1.0E-4, 1.0E-3	0.64	0.78	7.59	8.2E-3
$+L^{3/8}$	1.0E-7, 5.0E-3	0.66	0.80	7.57	8.2E-3	1.0E-3, 1.0E-3	0.67	0.80	8.12	7.2E-3
$+L^{1/8}$	5.0E-7, 1.0E-4	0.70	0.82	7.65	8.0E-3	1.0E-4, 1.0E-4	0.69	0.80	7.47	8.4E-3
+Log	5.0E-7, 1.0E-3	0.70	0.82	8.01	7.2E-3	**1.0E-3, 1.0E-3**	**0.72**	**0.84**	**7.84**	**7.4E-3**

3.3. Phantom Experimental Results

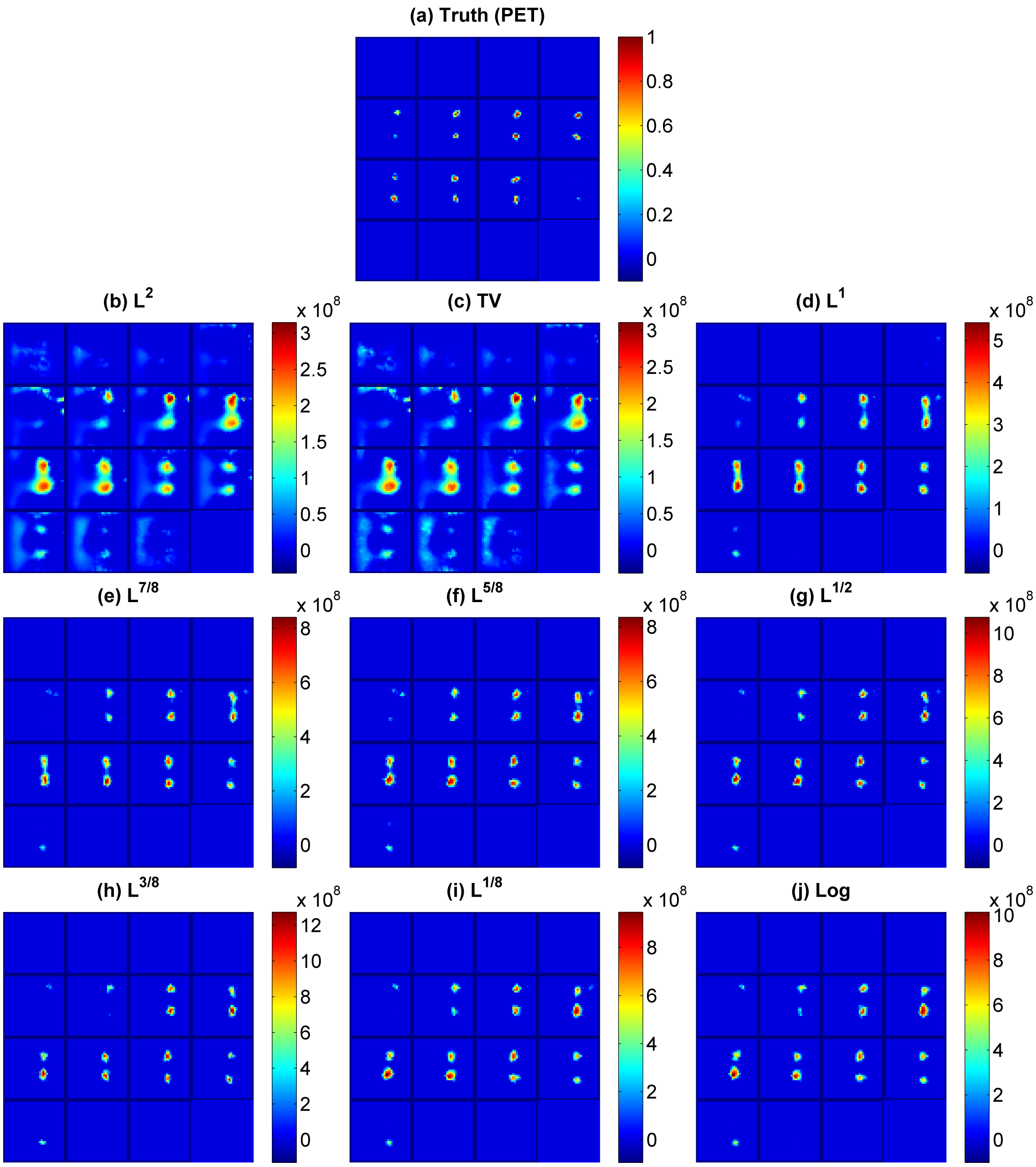
The phantom experimental results are summarized in Figure 3 and Table 5.

Table 5. Metrics for image quality for the cubic phantom, with the best regularization method highlighted in bold.

Reg Type	Reg λ	VR	Dice	CNR
L^2	$\lambda_2 = 1.0\text{E-}6$	4.48	0.28	5.12
TV	$\lambda_{TV} = 3.0\text{E-}9$	4.58	0.27	4.90
L^1	$\lambda_1 = 9.0\text{E+}3$	2.50	0.33	6.62
$L^{7/8}$	$\lambda_{7/8} = 1.0\text{E+}5$	1.67	0.32	6.36
$L^{5/8}$	$\lambda_{5/8} = 4.0\text{E+}6$	2.05	0.43	8.60
$\mathbf{L^{1/2}}$	$\boldsymbol{\lambda_{1/2}} = \mathbf{5.0E\text{+}7}$	**1.52**	**0.39**	**7.74**
$L^{3/8}$	$\lambda_{3/8} = 6.0\text{E+}8$	1.22	0.37	7.17
$L^{1/8}$	$\lambda_{1/8} = 1.0\text{E+}10$	1.92	0.34	7.26
Log	$\lambda_{Log} = 9.0\text{E+}11$	1.73	0.34	7.03

We included the image reconstructed from PET as the true image. Among the four image metrics, MSE is not computed, since we do not have the true image intensity information available. From Figure 3, we clearly see the patterns we identified in the first simulation case, which are (1) L^1, $L^q(0 < q < 1)$, and Log regularizations result in sparser and more accurate reconstructed images than L^2 and TV; and (2) regularization L^q, $q = 1/2$ has produced an image with the fewest artifacts from visual assessment and had a low VR, a high Dice and CNR, as shown in Table 5. Experiments with a combination of the sparse parameter and smooth parameter were run for phantom data, and similar to the simulations we had on small targets, the combination did not bring any improvements.

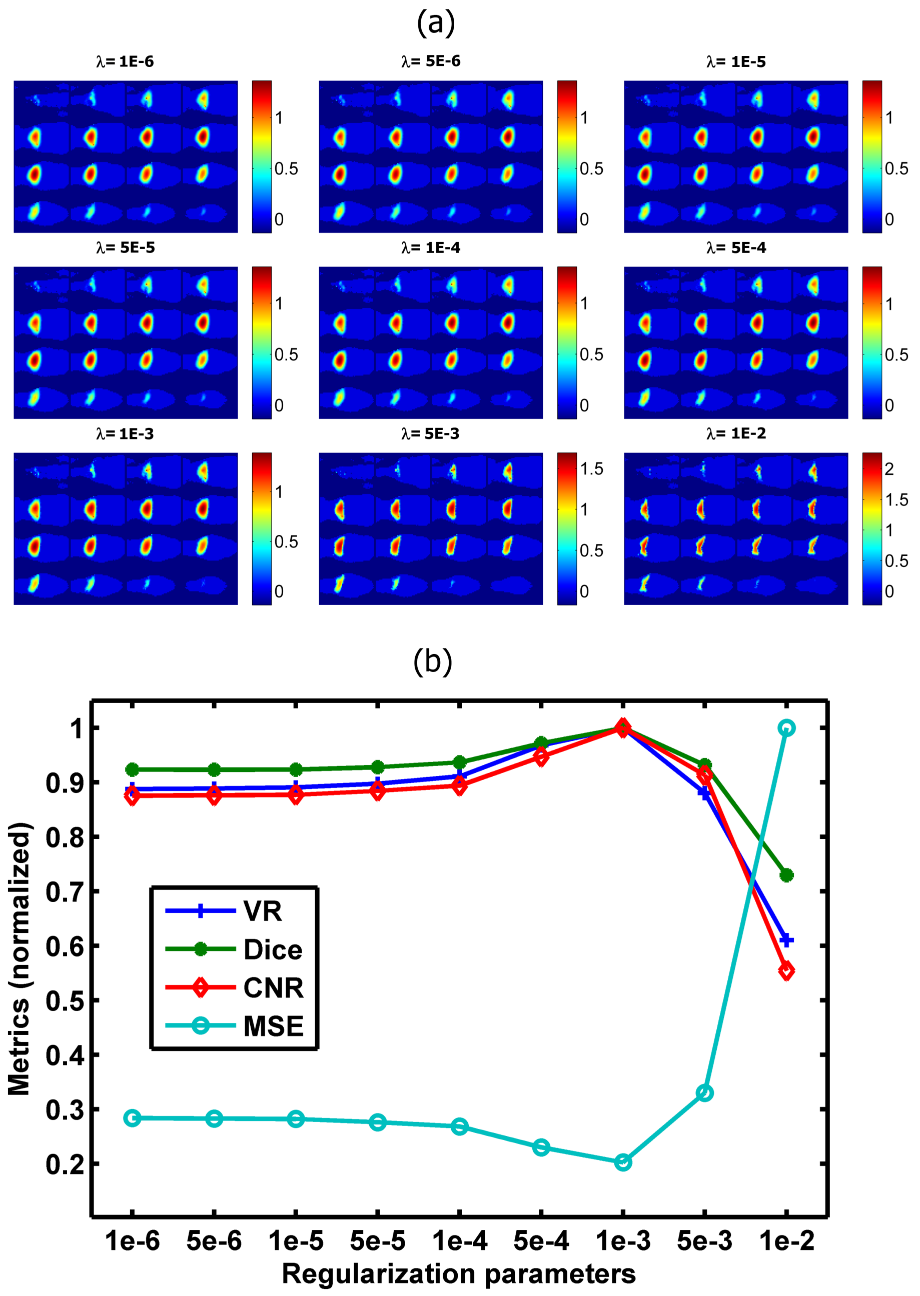
Figure 3. Cubic phantom with two small rods as target: coronal sections of the phantom showing (**a**) PET result as truth, and reconstruction results using regularizations based on: (**b**) L^2 with $\lambda_2 = $ 1.0E-6, (**c**) TV with $\lambda_{TV} = $ 3.0E-9, (**d**) L^1 with $\lambda_1 = $ 9.0E+3, (**e**) $L^{7/8}$ with $\lambda_{7/8} = $ 1.0E+5, (**f**) $L^{5/8}$ with $\lambda_{5/8} = $ 4.0E+6, (**g**) $L^{1/2}$ with $\lambda_{1/2} = $ 5.0E+7, (**h**) $L^{3/8}$ with $\lambda_{3/8} = $ 6.0E+8, (**i**) $L^{1/8}$ with $\lambda_{1/8} = $ 1.0E+10, and (**j**) Log with $\lambda_{log} = $ 9.0E+11, respectively.



3.4. Regularization Parameter Selection

For a better understanding of how the best regularization parameter for each regularization was chosen, we included, as an example, Figure 4 to show the reconstructed images using Log regularization with different regularization parameters, as well as the comparison of the corresponding image quality metrics. A visual assessment gave us λ_{Log} = 5E-4, 1E-3 and 5E-3. Then, their metrics indicate that λ_{Log} = 1E-3 is the best choice, where its MSE is the smallest and VR, Dice and CNR are the largest.

Figure 4. Selection of regularization parameters for Log on a large target, by visual and by metrics, respectively: (**a**) reconstructed fluorescence molecular tomography (FMT) images with Log regularization using different regularization parameters; (**b**) normalized metrics for the images in (a).

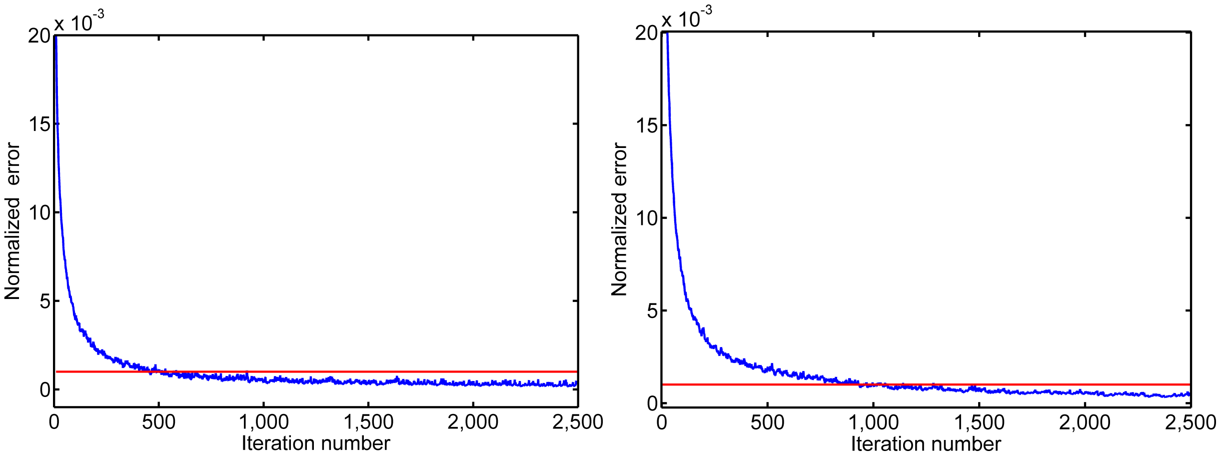


4. Discussion and Conclusions

We thoroughly compared a number of regularization methods in the reconstruction of FMT for both small targets and large targets using numerical simulations. We also included phantom experiments. The optimization transfer algorithm was adopted for a uniform iterative update formula. We started the iterative algorithm from a uniform initial value that can be randomly chosen on interval (0,1) and ran a maximum of 5,000 steps for both the numerical simulations and phantom experimental data or until the normalized error of the iterative update becomes less than 0.1%, which is the preset stopping criteria, δ_{stop}, for our algorithms. We used visual assessment along with four different image quality metrics, VR, Dice, MSE and CNR, to analyze the reconstructed images. For the simulated data, each iteration takes about 1.6 s to finish on an Intel i5 2400 3.1 GHz PC with 8 GB memory; while for the phantom, each iteration takes only about 0.09 s, since its mesh is much smaller. For the small target case, both numerical simulations and the phantom experiment demonstrated that non-convex regularizations, L^q, for q near 1/2 provided VR and Dice values closest to one, while maintaining a small MSE and a high CNR. The combination of smoothing and localizing regularization does not seem to improve the image quality in comparison with the best individual method for this case. However, for the large target case, the joint L^2 and Log performed the best, by providing the best VR, Dice, CNR and MSE at the same time, when compared with the individual or combined regularization methods.

Despite the potentially slow convergence of optimization transfer algorithms [13,28], we found that for all the regularization methods, the normalized error $\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_{L^2}/\|\mathbf{x}^k\|_{L^2}$ became less than 0.1% in less than 1,000 iterations for the large target case; while it may take up to 5,000 iterations before the normalized error reaches the 0.1% level for the small target cases. For example, we see in Figure 5 (left) that the normalized error becomes $< 0.1\%$ in about 500 iterations for the large target case using Log regularization, whereas in Figure 5 (right), the normalized error becomes $< 0.1\%$ only after about 800 iterations for the small target case using $L^q, q = 1/2$ regularization.

Figure 5. Normalized error as the iteration number increases: (**left**) for a large target using Log regularization; and (**right**) for a small target using $L^{1/2}$, respectively. The red line indicates the 0.1% level.



The choice of the stopping criterion is critical. This partially explains why our findings for the large target case are somewhat different from the conclusions in [11]. The sparsity enhancing regularizations, especially the Log, suppress the noise much faster than L^2. Then, in combination with the smoothing effect of L^2, the joint Log and L^2 provided the most satisfactory image. We did notice though that if we require an even smaller stopping criterion, the sparsity enhancing nature of Log will cause the image quality to deteriorate for this large target scenario. We plan to investigate this issue more in the future.

In summary, FMT and other optical tomography modalities have been active research topics in the past two decades. However, due to the high scattering of photons in tissues, reconstructed images still suffer from low spatial resolution. The newly proposed non-convex regularizations were proven to outperform the previously known L^1 regularization in enhancing image sparsity and suppressing noises. Yet, their power on large targets was not clear. In this paper, we have made a comprehensive comparison of $L^2, L^1, TV, L^q(0 < q < 1), Log$ regularizations and the combination of a smoothing L^2 or TV regularization and a localizing L^1, $L^q(0 < q < 1)$ or Log regularization. We found that for small targets, L^q for q around 1/2 performs the best; whereas for large targets, joint L^2 and Log is the best choice. These conclusions should provide useful guidance for selecting the best regularization methods for FMT in either small or large target scenarios.

Acknowledgments

This work is supported by the UC Merced Start-up Fund.

Author Contributions

Dianwen Zhu, Zhen Yuan and Changqing Li conceived the study. Dianwen Zhu and Changqing Li developed the methodology. Dianwen Zhu, Yue Zhao and Reheman Baikejiang performed the analysis. Dianwen Zhu wrote the first draft of the manuscript. All authors contributed and approved the final manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Cherry, S.R. *In vivo* molecular and genomic imaging: New challenges for imaging physics. *Phys. Med. Biol.* **2004**, *49*, R13–R48.
2. Ntziachristos, V.; Ripoll, J.; Wang, L.V.; Weissleder, R. Looking and listening to light: The evolution of whole-body photonic imaging. *Nat. Biotechnol.* **2005**, *23*, 313–320.
3. Tikhonov, A.; Arsenin, V.Y. *Solutions of Ill-Posed Problems*; WH Winston: Washington, DC, USA, 1977; p. 330.
4. Hyde, D.; Miller, E.L.; Brooks, D.H.; Ntziachristos, V. Data specific spatially varying regularization for multimodal fluorescence molecular tomography. *IEEE Trans. Med. Imaging* **2010**, *29*, 365–374.

5. Axelsson, J.; Svensson, J.; Andersson-Engels, S. Spatially varying regularization based on spectrally resolved fluorescence emission in fluorescence molecular tomography. *Opt. Express* **2007**, *15*, 574–13.
6. Pogue, B.W.; McBride, T.O.; Prewitt, J.; Österberg, U.L.; Paulsen, K.D. Spatially variant regularization improves diffuse optical tomography. *Appl. Opt.* **1999**, *38*, 2950–2961.
7. Bloch, S.; Lesage, F.; McIntosh, L.; Gandjbakhche, A.; Liang, K.; Achilefu, S. Whole-body fluorescence lifetime imaging of a tumor-targeted near-infrared molecular probe in mice. *J. Biomed. Opt.* **2005**, *10*, 054003–054003.
8. Weissleder, R.; Tung, C.H.; Mahmood, U.; Bogdanov, A. *In vivo* imaging of tumors with protease-activated near-infrared fluorescent probes. *Nat. Biotechnol.* **1999**, *17*, 375–378.
9. Shi, J.; Zhang, B.; Liu, F.; Luo, J.; Bai, J. Efficient L1 regularization-based reconstruction for fluorescent molecular tomography using restarted nonlinear conjugate gradient. *Opt. Lett.* **2013**, *38*, 3696–3699.
10. Han, D.; Tian, J.; Zhu, S.; Feng, J.; Qin, C.; Zhang, B.; Yang, X. A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization. *Opt. Express* **2010**, *18*, 8630–8646.
11. Yi, H.; Chen, D.; Li, W.; Zhu, S.; Wang, X.; Liang, J.; Tian, J. Reconstruction algorithms based on L1-norm and L2-norm for two imaging models of fluorescence molecular tomography: A comparative study. *J. Biomed. Opt.* **2013**, *18*, 056013, doi:10.1117/1.JBO.18.5.056013.
12. Behrooz, A.; Zhou, H.M.; Eftekhar, A.A.; Adibi, A. Total variation regularization for 3D reconstruction in fluorescence tomography: Experimental phantom studies. *Appl. Opt.* **2012**, *51*, 8216–8227.
13. Dutta, J.; Ahn, S.; Li, C.; Cherry, S.R.; Leahy, R.M. Joint L^1 and total variation regularization for fluorescence molecular tomography. *Phys. Med. Biol.* **2012**, *57*, 1459–1476.
14. Zhu, D.; Li, C. Nonconvex regularizations in fluorescence molecular tomography for sparsity enhancement. *Phys. Med. Biol.* **2014**, *59*. (In Press.)
15. Prakash, J.; Shaw, C.; Manjappa, R.; Kanhirodan, R.; Yalavarthy, P.K. Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction. *IEEE J. Sel. Top. Quantum Electron.* **2014**, *20*, 6800609, doi:10.1109/JSTQE.2013.2278218.
16. Okawa, S.; Hoshi, Y.; Yamada, Y. Improvement of image quality of time-domain diffuse optical tomography with L_q sparsity regularization. *Biomed. Opt. Express* **2011**, *2*, 3334–3348.
17. Weissleder, R.; Ntziachristos, V. Shedding light onto live molecular targets. *Nat. Med.* **2003**, *9*, 123–128.
18. Shu, X.; Royant, A.; Lin, M.Z.; Aguilera, T.A.; Lev-Ram, V.; Steinbach, P.A.; Tsien, R.Y. Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* **2009**, *324*, 804–807.
19. Arridge, S.R. Optical tomography in medical imaging. *Inverse Probl.* **1999**, *15*, R41–R93.
20. Li, C.; Mitchell, G.S.; Dutta, J.; Ahn, S.; Leahy, R.M.; Cherry, S.R. A three-dimensional multispectral fluorescence optical tomography imaging system for small animals based on a conical mirror design. *Opt. Express* **2009**, *17*, 7571–7585.

21. Correia, T.; Gibson, A.; Schweiger, M.; Hebden, J. Selection of regularization parameter for optical topography. *J. Biomed. Opt.* **2009**, *14*, 034044, doi:10.1117/1.3156839.
22. Tian, F.; Alexandrakis, G.; Liu, H. Optimization of probe geometry for diffuse optical brain imaging based on measurement density and distribution. *Appl. Opt.* **2009**, *48*, 2496–2504.
23. Dice, L.R. Measures of the amount of ecologic association between species. *Ecology* **1945**, *26*, 297–302.
24. Song, X.; Pogue, B.W.; Jiang, S.; Doyley, M.M.; Dehghani, H.; Tosteson, T.D.; Paulsen, K.D. Automated region detection based on the contrast-to-noise ratio in near-infrared tomography. *Appl. Opt.* **2004**, *43*, 1053–1062.
25. Dogdas, B.; Stout, D.; Chatziioannou, A.F.; Leahy, R.M. Digimouse: A 3D whole body mouse atlas from CT and cryosection data. *Phys. Med. Biol.* **2007**, *52*, 577–587.
26. Li, C.; Wang, G.; Qi, J.; Cherry, S.R. Three-dimensional fluorescence optical tomography in small-animal imaging using simultaneous positron-emission-tomography priors. *Opt. Lett.* **2009**, *34*, 2933–2935.
27. Li, C.; Yang, Y.; Mitchell, G.S.; Cherry, S.R. Simultaneous PET and multispectral 3-dimensional fluorescence optical tomography imaging system. *J. Nucl. Med.* **2011**, *52*, 1268–1275.
28. Hunter, D.R.; Lange, K. A tutorial on MM algorithms. *Am. Stat.* **2004**, *58*, 30–37.