Research Article | Vol. 10, No. 6 | 1 Jun 2019 | BIOMEDICAL OPTICS EXPRESS 2684
Biomedical Optics EXPRESS

Graph- and finite element-based total variation models for the inverse problem in diffuse optical tomography

WENQI LU,[1] JINMING DUAN,[1] DAVID ORIVE-MIGUEL,[2,3] LIONEL HERVE,[2] AND IAIN B. STYLES[1,*]

[1] School of Computer Science, University of Birmingham, UK
[2] CEA, LETI, MINATEC Campus, F-38054 Grenoble, France
[3] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France
[*] I.B.Styles@bham.ac.uk

Abstract: Total variation (TV) is a powerful regularization method that has been widely applied in different imaging applications, but is difficult to apply to diffuse optical tomography (DOT) image reconstruction (inverse problem) due to unstructured discretization of complex geometries, non-linearity of the data fitting and regularization terms, and non-differentiability of the regularization term. We develop several approaches to overcome these difficulties by: i) defining discrete differential operators for TV regularization using both finite element and graph representations; ii) developing an optimization algorithm based on the alternating direction method of multipliers (ADMM) for the non-differentiable and non-linear minimization problem; iii) investigating isotropic and anisotropic variants of TV regularization, and comparing their finite element (FEM)- and graph-based implementations. These approaches are evaluated on experiments on simulated data and real data acquired from a tissue phantom. Our results show that both FEM and graph-based TV regularization is able to accurately reconstruct both sparse and non-sparse distributions without the over-smoothing effect of Tikhonov regularization and the over-sparsifying effect of L_1 regularization. The graph representation was found to out-perform the FEM method for low-resolution meshes, and the FEM method was found to be more accurate for high-resolution meshes.

© 2019 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Diffuse optical tomography (DOT) is an important non-invasive imaging technique whose major applications include diagnosing breast cancer [1–3], imaging small animals for the study of disease and analyzing brain function in functional neuroimaging [4–7]. In DOT, near-infrared light (spectral range of 650-900 nm) is injected into the object through optical fibers placed on the surface of the object. The transmitted light is then collected using optical detectors, forming a series of boundary measurements, each of which corresponds to the signal received by a single detector during illumination by a single source. Image reconstruction algorithms are then used to recover the internal distribution of the underlying optical properties of the object from the boundary measurements.

Due to the limited availability of boundary measurements and diffusive nature of near-infrared light propagation, image reconstruction in DOT is an underdetermined, ill-posed and non-linear inverse problem. Regularization is often used to constrain the inverse problem to yield physiologically and anatomically plausible solutions, resulting in the following minimization problem with respect to the optical properties μ

$$\mu^* = \arg\min_{\mu} \left\{ \frac{1}{2} \|\Phi^{\mathrm{M}} - \mathcal{F}(\mu)\|_2^2 + \lambda \mathcal{R}(\mu) \right\}, \tag{1}$$

#357442 https://doi.org/10.1364/BOE.10.002684
Journal © 2019 Received 11 Jan 2019; revised 29 Mar 2019; accepted 17 Apr 2019; published 2 May 2019

Biomedical Optics EXPRESS

where Φ^{M} represents the boundary measurements acquired from the optical detectors, $\mathcal{F}$ is the non-linear operator induced from the forward model [8], $\mathcal{R}$ is a general regularization term, and λ is a weight that determines the extent to which regularization will be imposed on the solution μ^*. The quadratic Tikhonov-type regularization is widely used. However, it promotes smooth solutions and thereby smears sharp features embedded in the image [9]. Regularizations based on the L_1-norm of the solution have been also extensively studied [10–12], as they impose a sparsity constraint on the solution, enabling the recovery of sharp edges of objects in reconstructed images. Eq. (1) with either regularization mentioned results in a convex optimization problem, where highly efficient algorithms [13] are available. Recently, the more general L_p regularization ($0 < p < 1$) that promotes sparsity into the resulting image [14], has been employed for DOT image reconstruction [15, 16]. However, L_p regularization is nonconvex and therefore difficult to optimize.

Regularizations involving the L_1 or L_p-norm of the solution are used under the assumption that the optical properties (representing the image) to be reconstructed are spatially sparse. These regularizations tend to oversparsify the distribution of the optical properties when such an assumption does not hold [13], for example, in the case of multiple activations or complex injuries in the brain, where the features of interest are not spatially localized and the optical properties relative to the background are therefore non-sparse [7]. In order to be able to reconstruct images in which edges are preserved and features are not spatially sparse, a different approach is required. Total variation (TV) regularization, which uses the L_1-norm of the *gradient* of the solution as a regularizing term (the detailed forms will be given in Section 2.1 and 2.2), can be used to overcome the limitations. The gradient operator can transform the solution μ^* to a sparse space where non-zero values only occur at sharp features. As such, TV can perform better than the pure sparsity preserving regularizations at preserving edges of objects in the images that are not sparse. Further, gradient is a highpass operator, which imposes smoothness to the solution. This improves the conditioning of the minimization problem (Eq. (1)), thus enabling a robust numerical solution. Due to these advantages, TV has been adapted from applications in imaging processing [17–19] to various medical image reconstruction problems, including photoacoustic tomography (PAT) [20], bioluminescence tomography (BLT) [21], fluorescence tomography (FT) [22], as well as DOT [23–25].

In most TV-associated imaging problems, the minimization problem is carried out on a Cartesian grid where each element represents a pixel (voxel in three dimensional (3D)) in the image [26]. In this case, the differential operators resulting from minimizing the TV regularization, such as gradient, divergence, Laplacian and curvature, are discretized straightforwardly using the finite difference method (FDM) [26]. In DOT, it is however non-trivial to represent the complex geometry (i.e. the multi-layer head used in our experiment) using a Cartesian grid and the FDM is not always practical. Two representations are often employed to model complex geometries: finite element and graph representations. In the former, the object geometry is represented by a polygon/polyhedron, over which a series of disjoint triangles (tetrahedra in 3D) are usually generated. In the latter, the object geometry is represented by an unstructured graph, defined by vertices, edges and weights. Such a graph is normally constructed by exploring neighborhood relationships between vertices. For each representation, there is a systematic discretization scheme (finite element discretization or graph discretization) for the differential operators, which can be readily applied to the minimization of TV-associated problems. We note that although TV regularization has previously been studied in DOT [23–25], none of these studies have provided detailed information about the discretization schemes they used. The performance of different discretization schemes for TV regularization in DOT has not been systematically investigated.

The minimization of a TV-associated problem can be non-trivial due to the non-linearity and non-differentiability of the TV regularization term. In image processing, many efficient optimization algorithms have been developed for this task, including iteratively reweighed least squares [27],

Biomedical Optics EXPRESS

primal dual [28], split Bregman [29], and fast iterative shrinkage-thresholding algorithm (FISTA) [30,31]. Recently, alternating direction method of multipliers (ADMM) [18,26,32,33] has become increasingly popular. The elegance of ADMM lies in decomposition of the original minimization problem into several simple subproblems, each of which either has a closed-form solution or can be iteratively solved with efficient numerical methods. However, since ADMM-based methods have been implemented mainly for Cartesian grids using a forward-backward FDM [34] and it is not straightforward to generalize them to solve the inverse problem on an unstructured domain. Moreover, the non-linearity of the data fitting term in Eq. (1) further complicates the DOT reconstruction problem, making the minimization process required to solve Eq. (1) difficult.

In this paper, we address these limitations and develop TV regularization approaches for the inverse problem in DOT. More specifically, we make the following three distinct contributions: **(1)** We introduce finite element and graph representations to discretize the TV regularization term in DOT reconstruction enabling the minimization of the inverse problem associated with TV regularization to be carried out on unstructured domains. To the best of our knowledge, this is the first time that finite element-based discretization methods have been provided in detail for DOT image reconstruction with TV regularization. Additionally, we are not aware of any previous work that attempts to formulate the TV-regularized inverse problem using a graph representation. **(2)** We propose an efficient algorithm based on ADMM to minimize the TV-regularized inverse problem. Our algorithm can handle geometries with unstructured grids, and also reduces the computational difficulties arising from the non-differentiability and dual non-linearities in the inverse problem. **(3)** We further investigate the isotropic and anisotropic variants of the TV regularization, and compare their finite element- and graph-based implementations against a baseline Tikhonov model, both qualitatively and quantitatively using extensive numerical experiments.

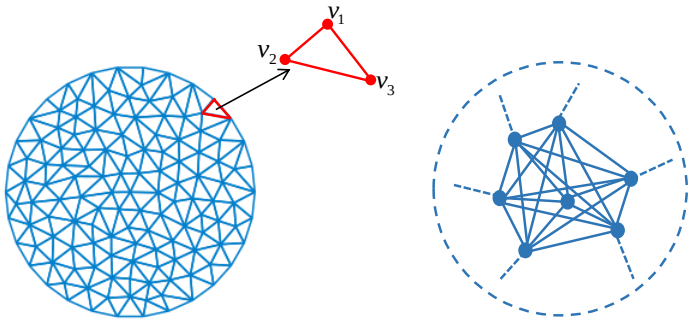

Fig. 1. Modeling a complex geometry using finite element (left) and graph (right) representations.

2. Discretizations using finite element and graph representations

In this section, we first show how an unstructured computational domain can be modelled using finite element (FE) and graph representations. The definitions of TV regularizations (anisotropic and isotropic versions) and corresponding discrete differential operators are then given accordingly for each representation. In Fig. 1, we show a circle discretized with the two representations. For the FE representation (left), the circle is divided by a series of elements joint at different vertices. In the FE representation, we normally term a discrete geometry as a FE mesh. Within the circle mesh, a triangle (representing one element) is highlighted comprising three disjoint vertices. For the graph representation (right), the circle is simply discretized with a

Biomedical Optics EXPRESS

set of vertices and edges. In this representation there is no concept of 'element'. Note that it is easy to convert the FE representation to the graph representation. For example, the FE mesh can be viewed as a graph if we consider only the vertices and edges in it. In this section, two representations are introduced to model the unstructured computational domain of complex DOT geometries.

2.1. Finite element-based discrete differential operators

We apply the Galerkin FE method to the computational domain in DOT [35], the first step of which is to approximate a continuous function by a piecewise-polynomial function. Using the FE representation, let us first discretize an unstructured two dimensional (2D) domain Ω by M triangles jointed at N vertex nodes (e.g. Fig. 1 left). $V = \{V_k\}_{k=1}^{N}$ denotes a finite number of N nodes. Let ϑ_h be the 2D vector space of continuous piecewise-linear functions on the triangles in the FE mesh. The continuous and piecewise-linear function $U(x, y) : \Omega \rightarrow \mathbb{R}$, approximating the optical properties on Ω, can be written in the form of

$$U = \sum_{i=1}^{N} \mu_i \varphi_i. \tag{2}$$

Here $\{\varphi_i\}_{i=1}^{N}$, $\varphi_i \in \vartheta_h$ are linear basis functions defined as $\varphi_j(V_i) = 1$ if $i = j$ and $\varphi_j(V_i) = 0$ if $i \neq j$. $\mu_i : V \rightarrow \mathbb{R}$ is the value of optical property on each vertex in the FE mesh, $i = 1, ..., N$.

Eq. (2) means that the optical property value inside a triangle is associated with the optical property values on all vertices in the mesh. Given three vertices of a triangle T, i.e. $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2)$ and $v_3 = (x_3, y_3)$, there are three linear basis functions φ_i associated with the vertices, which are respectively expressed as

$$\begin{aligned} &\varphi_1(x, y) = a_1 x + b_1 y + c_1 \\ &\varphi_2(x, y) = a_2 x + b_2 y + c_2 : \quad \Omega \rightarrow \mathbb{R}, \\ &\varphi_3(x, y) = a_3 x + b_3 y + c_3 \end{aligned} \tag{3}$$

$a_1 = (y_2 - y_3)/(2A_T)$, $b_1 = (x_3 - x_2)/(2A_T)$, $c_1 = (x_2 y_3 - x_3 y_2)/(2A_T)$, $a_2 = (y_3 - y_1)/(2A_T)$, $b_2 = (x_1 - x_3)/(2A_T)$, $c_2 = (x_3 y_1 - x_1 y_3)/(2A_T)$, $a_3 = (y_1 - y_2)/(2A_T)$, $b_3 = (x_2 - x_1)/(2A_T)$ and $c_3 = (x_1 y_2 - x_2 y_1)/(2A_T)$. (x, y) represents any point inside of the triangle T. A_T denotes the triangular area of T, which is computed as $A_T = |x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)|/2$.

In FE, one starts from a continuous problem and approximates the solution with a piecewise-polynomial function U. As such, we define the following anisotropic and isotropic TV regularizations

$$\int_{\Omega} \left(|\partial_x U| + \left|\partial_y U\right|\right) dxdy = \|D_x \mu\|_1 + \left\|D_y \mu\right\|_1. \tag{4}$$

$$\int_{\Omega} \sqrt{(\partial_x U)^2 + \left(\partial_y U\right)^2} dxdy = \sum_{i=1}^{M} \sqrt{|(D_x \mu)_i|^2 + \left|\left(D_y \mu\right)_i\right|^2}. \tag{5}$$

In Eq. (4) and Eq. (5), the continuous TV regularizations and their resulting discretized versions are shown on the left-hand side and right-hand side, respectively. The two discrete versions respectively are the anisotropic and isotropic definitions of TV regularization. ∂_x and ∂_y are continuous partial derivatives along the x and y directions, respectively. D_x is a matrix of size $M \times N$ which, when applied to μ gives the discrete partial derivative of μ along the x direction. D_y is the derivative matrix along the y direction. $D_x \mu$ and $D_y \mu$ are therefore two vectors of size $M \times 1$, where M is the number of triangles in the mesh. We note that the main idea of FE is to break down the calculation domain Ω onto the local elements individually. Afterwards, the

Biomedical Optics EXPRESS

derived local matrices are assembled element by element to enable the final computation. Eq. (4) and Eq. (5) can be proved by expressing the partial derivatives $\partial_x U$ and $\partial_y U$ in terms of a basis. To illustrate this idea we prove the first term of Eq. (4):

$$\begin{aligned}\int_\Omega |\partial_x U| dxdy &= \sum_{i=1}^{M} \int_{T_i} |\partial_x U| \, dxdy = \sum_{i=1}^{M} \int_{T_i} |\sum_{j=1}^{N} \mu_j \partial_x \varphi_j| dxdy \\ &= \sum_{i=1}^{M} A_{T_i} \left| a_{i,1}\mu_{i,1} + a_{i,2}\mu_{i,2} + a_{i,3}\mu_{i,3} \right| \\ &= \sum_{i=1}^{M} |(D_x \mu)_i| = \|D_x \mu\|_1,\end{aligned} \tag{6}$$

where A_{T_i} denotes the area of triangle T_i and $\{i,1\}$, $\{i,2\}$, $\{i,3\}$ in $\left|a_{i,1}\mu_{i,1} + a_{i,2}\mu_{i,2} + a_{i,3}\mu_{i,3}\right|$ represent the indices of the vertices of the i-th triangle. As $\left|a_{i,1}\mu_{i,1} + a_{i,2}\mu_{i,2} + a_{i,3}\mu_{i,3}\right|$ is a linear combination, we can thus construct the discrete derivative matrix D_x with the following steps:

- Initialize all-zeros matrix D_x of size $M \times N$.
- Loop over M triangles; for each triangle i, compute the coefficients a_1, a_2 and a_3 using the coordinates of the three vertices and fill in the three columns in the i-th row of matrix D_x corresponding to the position of the three vertices in the node sequence.

The discrete derivative matrix D_y can be obtained in a similar way. Note that D_x and D_y are sparse matrices as their most entries are zeros. With D_x and D_y defined, we can therefore minimize the TV regularization (either anisotropic (Eq. (4)) or isotropic (Eq. (5)) version) with the data fidelity term in (Eq. (1)) for DOT reconstruction over 2D unstructured geometries. The corresponding 3D counterparts were also implemented in this paper, as shown in the experiments.

2.2. Graph-based discrete differential operators

In this section, we introduce discrete differential operators on graphs and from these derive the TV regularization terms. First, we discretize an unstructured domain Ω by a weighted graph $G = (V, E, w)$ (e.g. Fig. 1 right). In the graph G, $V = \{V_k\}_{k=1}^{N}$ denotes a finite set of N vertices, and $E \in V \times V$ is a finite set of weighted edges. We assume that G is an undirected simple graph (no multiple edges) in this study. Let $(i, j) \in E$ be an edge of E that connects the vertices i and j in V. Let $\mu_i : V \to \mathbb{R}$ denote the value of the optical properties on i. The DOT reconstruction problem then reduces to finding an optimal value of the optical properties for each of N vertices in V.

The discrete differential operators are defined on the graph based on nonlocal methods [18, 37, 38]. First, we define the *nonlocal gradient operator* ∇_w acting on μ_i

$$\nabla_w \mu_i \triangleq \left(\mu_j - \mu_i\right) \sqrt{w_{ij}} \; : V \to \mathbb{R}. \tag{7}$$

For vertex $i \in V$, $\nabla_w \mu_i$ is a vector with a length of N. The weight $w_{i,j} : V \times V \to \mathbb{R}^+$ represents the similarity between nodes i and j, which is nonnegative and symmetric. The weight function can be determined in many ways. In this study we choose to use the Euclidean distance to define the weight function with $w_{i,j} = 1/d_{i,j}$ where $d_{i,j}$ represents the distance between vertex i and j. We note that an important difference between the FE gradient and the nonlocal gradient is that the former has two directions in 2D or three directions in 3D, whilst the latter is a vector of partial derivatives along all edges connected to the node.

Biomedical Optics EXPRESS

Given a vector function $\boldsymbol{v}_i : V \to \mathbb{R}$, the *nonlocal divergence operator* div_w acting on $\boldsymbol{v}_i$ is given as

$$\mathrm{div}_w\, \boldsymbol{v}_i \triangleq \sum_{j=1}^{N} \left(v_{ij} - v_{ji}\right) \sqrt{w_{ij}} \;: V \to \mathbb{R}, \tag{8}$$

where v_{ij} is the jth index in the vector $\boldsymbol{v}_i$.

Based on Eq. (7) and Eq. (8), the *nonlocal Laplace operator* Δ_w acting on μ_i is written as

$$\Delta_w \mu_i \triangleq \frac{1}{2}\mathrm{div}_w\, (\nabla_w \mu_i) = \sum_{j=1}^{N} \left(\mu_j - \mu_i\right) w_{ij} : V \to \mathbb{R}, \tag{9}$$

which is a linear operator also known as the *graph Laplacian*.

With these discrete differential operators defined on graph, we propose the anisotropic graph TV regularization

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \left|\left(\mu_j - \mu_i\right) \sqrt{w_{ij}}\right|, \tag{10}$$

and the isotropic graph TV regularization

$$\sum_{i=1}^{N} \sqrt{\sum_{j=1}^{N} \left(\mu_j - \mu_i\right)^2 w_{ij}}. \tag{11}$$

In the above definitions, the constructed graph G is assumed to be fully connected, meaning that each vertex is connected to all other vertices in G. In this case, the computational load for the minimizations of Eq. (10) and Eq. (11) will be extremely heavy. Spectral graph theory [39, 40] or nearest neighbour [41, 42] techniques are typically employed to limit the number of edges that are considered. For example, [39] and [40] use spectral approaches and the Nystrom extension method [43] to efficiently calculate the eigen-decomposition of a dense graph Laplacian. In this work, we build the graph by borrowing the positions of the vertices and the connectivity between vertices in the finite element mesh as the vertices and edges in the graph. With this structure, the graph is sparsified and only connected vertices for a given vertice are taken into consideration. In such a case, regularizations (10) and (11) are respectively equivalent to

$$\sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} \left|\left(\mu_j - \mu_i\right) \sqrt{w_{ij}}\right|, \tag{12}$$

and

$$\sum_{i=1}^{N} \sqrt{\sum_{j \in \mathcal{N}_i} \left(\mu_j - \mu_i\right)^2 w_{ij}}, \tag{13}$$

where $\mathcal{N}_i = \{j \in V : (i, j) \in E\}$. We note that the 2D and 3D implementations of these differential operators are identical, making the resulting minimization processes of regularizations (12) and (13) more straightforward than the FE implementation.

3. Minimization of TV-associated DOT inverse problems

Due to the non-linearity of the data fitting term and the non-differentiability of the TV regularizations, it is non-trivial to minimize a TV-regularized inverse problem. It is harder than minimizing the standard L_1-regularized inverse problem [13] because of the existence of the gradient operator. In this section, we propose an efficient algorithm based on ADMM to address this, the idea of

Biomedical Optics EXPRESS

which is to first linearize the non-linear inverse problem and afterwards apply ADMM to the resulting linearized problem. The whole process is then iterated until convergence. We note that due to the use of the differential operators in Sections 2.1 and 2.2, the proposed algorithm can handle complex geometries with unstructured grids, and also can ease the computational difficulties arising from the non-differentiability and non-linearities in the inverse problem. We now describe the details of this algorithm.

3.1. Linearization

Non-linear problems are technically difficult to tackle directly, so iterative linearization can be used to convert a non-linear problem into a series of local linear problems. To do so, Taylor's series expansion is first used to approximate $\mathcal{F}(\mu)$ in the fitting term of Eq. (1) as

$$\mathcal{F}(\mu) \approx \mathcal{F}\left(\mu^{k-1}\right) + J^{k-1}\left(\mu - \mu^{k-1}\right), \tag{14}$$

where J^{k-1} denotes the Jacobian calculated from the $(k-1)$th iteration and is defined as $\partial\mathcal{F}(\mu^{k-1})/\partial(\mu^{k-1})$. The Jacobian in DOT is normally calculated using the adjoint method [44]. With this first order approximation, the non-linear problem Eq. (1) can be converted to

$$\delta\mu^k = \arg\min_{\delta\mu}\left\{\frac{1}{2}\|J^{k-1}\delta\mu - \delta\Phi^{k-1}\|_2^2 + \lambda\mathcal{R}(\delta\mu)\right\}, \tag{15}$$

which is an iterative linearized algorithm. $\delta\Phi^{k-1}$ is the data-model mismatch which is given by $\Phi^{\mathrm{M}} - \mathcal{F}\left(\mu^{k-1}\right)$ and $\delta\mu^k$ is the change in the optical property at the k-th iteration. With a proper initialization μ^0, a local minimizer to Eq. (1) can be found by iteratively updating this step. In Eq. (4), Eq. (5), Eq. (12) and Eq. (13), we have defined the four types of TV regularizations using different representations. We then apply them to $\mathcal{R}(\delta\mu)$ in Eq. (15), resulting in the following four linearized minimization problems in Table 1. However, it remains unclear how to optimize the second term in Eq. (15). We address this using ADMM.

Table 1. Four TV-regularized minimization problems obtained by applying different TV regularizations to Eq. (15). A-FETV, I-FETV, A-GTV and I-GTV respectively represent anisotropic finite element total variation, isotropic finite element total variation, anisotropic graph total variation and isotropic graph total variation.

Name	Formulation
A-FETV	$\delta\mu^* = \arg\min_{\delta\mu}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \lambda\|D_x(\delta\mu)\|_1 + \lambda\left\|D_y(\delta\mu)\right\|_1\right\}$
I-FETV	$\delta\mu^* = \arg\min_{\delta\mu}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \lambda\sum_{i=1}^{M}\sqrt{\left\|(D_x(\delta\mu))_i\right\|^2 + \left\|\left(D_y(\delta\mu)\right)_i\right\|^2}\right\}$
A-GTV	$\delta\mu^* = \arg\min_{\delta\mu}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \lambda\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i}\left\|\left(\delta\mu_j - \delta\mu_i\right)\sqrt{w_{ij}}\right\|\right\}$
I-GTV	$\delta\mu^* = \arg\min_{\delta\mu}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \lambda\sum_{i=1}^{N}\sqrt{\sum_{j\in\mathcal{N}_i}\left(\delta\mu_j - \delta\mu_i\right)^2 w_{ij}}\right\}$

3.2. ADMM implementations

In this section, we introduce ADMM in detail to minimize the A-FETV, I-FETV, A-GTV and I-GTV problems. We begin with the ADMM implementation for A-FETV. Specifically, auxiliary splitting vectors v_x and v_y are introduced to represent $D_x(\delta\mu)$ and $D_y(\delta\mu)$ respectively. Therefore

Biomedical Optics EXPRESS

the A-FETV problem is transformed into the following unconstrained optimization problem:

$$\begin{aligned}\delta\mu^n, v_x^n, v_y^n = \arg\min_{\delta\mu, v_x, v_y}\{&\frac{1}{2}||J\delta\mu - \delta\Phi||_2^2 + \lambda||v_x||_1 + \lambda||v_y||_1 \\ &+ \frac{\theta}{2}||v_x - D_x(\delta\mu) - b_x^{n-1}||_2^2 + \frac{\theta}{2}||v_y - D_y(\delta\mu) - b_y^{n-1}||_2^2\},\end{aligned} \tag{16}$$

where superscript n denotes the n-th ADMM iteration and b_x and b_y are iterative parameters. In order to find the minimizer of Eq. (16), an alternating optimization method is used where Eq. (16) is split into several subproblems with respect to $\delta\mu$, v_x, v_y, b_x and b_y, each of which can be solved separately.

First the iterative minimization approach requires us to solve the subproblem with respect to μ

$$\begin{aligned}\delta\mu^n = \arg\min_{\delta\mu}\{&\frac{1}{2}||J\delta\mu - \delta\Phi||_2^2 + \frac{\theta}{2}||v_x^{n-1} - D_x(\delta\mu) - b_x^{n-1}||_2^2 \\ &+ \frac{\theta}{2}||v_y^{n-1} - D_y(\delta\mu) - b_y^{n-1}||_2^2\},\end{aligned} \tag{17}$$

which has the optimality condition

$$\left(\left(J^T J + \theta\left(D_x{}^T D_x + D_y{}^T D_y\right)\right)\right)\delta\mu^n = J^T\delta\Phi - \theta D_x{}^T\left(b_x^{n-1} - v_x^{n-1}\right) - \theta D_y{}^T\left(b_y^{n-1} - v_y^{n-1}\right). \tag{18}$$

As the inversion matrix of Eq. (18) has size $N \times N$, in order to achieve high efficiency, we use a gradient descent method to optimize the functional iteratively, in which the step size controls how far the iterate moves along the gradient direction during the current iteration. Instead of setting the step size manually, we use a backtracking line search to enforce convergence [31].

The next subproblem with respect to v_x and v_y is given as

$$\begin{aligned}v_x^n, v_y^n = \arg\min_{v_x, v_y}\{&\lambda||v_x||_1 + \lambda||v_y||_1 + \frac{\theta}{2}||v_x - D_x(\delta\mu^n) - b_x^{n-1}||_2^2 \\ &+ \frac{\theta}{2}||v_y - D_y(\delta\mu^n) - b_y^{n-1}||_2^2\}.\end{aligned} \tag{19}$$

It should be noticed that, in A-FETV, there is no coupling between v_x and v_y. We can explicitly compute the optimal value of v_x and v_y using the generalized shrinkage operators

$$\begin{aligned}v_x^n &= \max\left(\left|D_x(\delta\mu^n) + b_x^{n-1}\right| - \frac{\lambda}{\theta}, 0\right)\frac{D_x(\delta\mu^n) + b_x^{n-1}}{\left|D_x(\delta\mu^n) + b_x^{n-1}\right|} \\ v_y^n &= \max\left(\left|D_y(\delta\mu^n) + b_y^{n-1}\right| - \frac{\lambda}{\theta}, 0\right)\frac{D_y(\delta\mu^n) + b_y^{n-1}}{\left|D_y(\delta\mu^n) + b_y^{n-1}\right|},\end{aligned} \tag{20}$$

with the convention that $0/0 = 0$. The last one is to update the iterative parameters b_x and b_y, as

$$\begin{aligned}b_x^n &= b_x^{n-1} + D_x(\delta\mu^n) - v_x^n \\ b_y^n &= b_y^{n-1} + D_y(\delta\mu^n) - v_y^n.\end{aligned} \tag{21}$$

In I-FETV, using the same alternating optimization method, the original minimization problem can be transformed as

$$\begin{aligned}\delta\mu^n, v_x^n, v_y^n = \arg\min_{\delta\mu, v_x, v_y}\{&\frac{1}{2}||J\delta\mu - \delta\Phi||_2^2 + \lambda\left\|(v_x, v_y)\right\|_2 \\ &+ \frac{\theta}{2}||v_x - D_x(\delta\mu) - b_x^{n-1}||_2^2 + \frac{\theta}{2}||v_y - D_y(\delta\mu) - b_y^{n-1}||_2^2\},\end{aligned} \tag{22}$$

Biomedical Optics EXPRESS

where

$$\left\|(v_x, v_y)\right\|_2 = \sum_{i=1}^{M} \sqrt{|(v_x)_i|^2 + \left|(v_y)_i\right|^2}, \tag{23}$$

and M is the number of finite elements. The first subproblem (L_2 component) with respect to $\delta\mu$ is the same as A-FETV. It should be noted that the v_x and v_y variables cannot be decoupled as they were in A-FETV. In order to solve the subproblem with respect to v_x and v_y, we can explicitly solve the minimization problem for $\left(v_x^n, v_y^n\right)$, using a generalized shrinkage formula

$$\begin{aligned} v_x^n &= \max\left(s^n - \frac{\lambda}{\theta}, 0\right) \frac{D_x(\delta\mu^n) + b_x^{n-1}}{s^n} \\ v_y^n &= \max\left(s^n - \frac{\lambda}{\theta}, 0\right) \frac{D_y(\delta\mu^n) + b_y^{n-1}}{s^n}, \end{aligned} \tag{24}$$

with the convention that $0/0 = 0$ and $s^n = \sqrt{\left|D_x(\delta\mu^n) + b_x^{n-1}\right|^2 + \left|D_y(\delta\mu^n) + b_y^{n-1}\right|^2}$. The iterative parameters b_x and b_y are then updated as shown in A-FETV.

The ADMM-based algorithm for A-FETV and I-FETV is given in Algorithm 1, where *inner_loop* is the maximum number of iterations for the ADMM-based algorithm.

Algorithm 1: ADMM-based algorithm for A-FETV and I-FETV.

INPUT: $\delta\Phi$, J, y, *inner_loop*, ϵ_1, regularization parameter $\theta > 0$, $\lambda > 0$

Initialization: $v_x^0 = v_y^0 = b_x^0 = b_y^0 = 0$

for $n = 1 : inner_loop$

1: Update μ^n using Eq. (17)

2: Update v_x^n and v_y^n using Eq. (20) for A-FETV or Eq. (24) for I-FETV

3: Update b_x^n and b_y^n using Eq. (21)

4: Stop if $n = inner_loop$ or $\frac{||\delta\mu^n - \delta\mu^{n-1}||_1}{||\delta\mu^{n-1}||_1} \le \epsilon_1$.

end for

RETURN $\delta\mu^k = \delta\mu^n$

We then propose ADMM-based algorithm to address the minimizations of A-GTV and I-GTV. For A-GTV, we first introduce an auxiliary splitting vector variable v, an iterative parameter b, and a positive penalty parameter θ. The sizes of v and b are both of $N \times N$ where N represents the number of vertices. The A-GTV problem can be reformulated as the following unconstrained optimization problem

$$\delta\mu^n, v^n = \arg\min_{\delta\mu, v}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \lambda\sum_{i=1}^{N}\|v_i\|_1 + \frac{\theta}{2}\sum_{i=1}^{N}\|v_i - \nabla_w(\delta\mu_i) - b_i^{n-1}\|_2^2\right\}. \tag{25}$$

Since Eq. (25) is a multivariate minimization problem, we first solve the subproblem with respect to $\delta\mu$

$$\delta\mu^n = \arg\min_{\delta\mu}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \frac{\theta}{2}\sum_{i=1}^{N}\|v_i^{n-1} - \nabla_w(\delta\mu_i) - b_i^{n-1}\|_2^2\right\}, \tag{26}$$

Biomedical Optics EXPRESS

which gives the the optimality condition

$$\left(J^T J\delta\mu - J^T\delta\Phi\right)_i + \theta \mathrm{div}_w\left(\nu_i^{n-1} - \nabla_w(\delta\mu_i) - b_i^{n-1}\right) = 0, \quad i = 1, ..., N. \tag{27}$$

With the definition of the *nonlocal divergence operator* (Eq. (8)) and the *nonlocal Laplace operator* (Eq. (9)), the point-wise equation system (Eq. (27)) can be equivalently converted to the following matrix-based equation system

$$(J^T J - \theta L)\delta\mu = J^T\delta\Phi - \theta g^{n-1}. \tag{28}$$

L above is the graph Laplacian in matrix form, whose entries are

$$L_{i,j} = \begin{cases} -\sum\limits_{j\in\mathcal{N}_i} w_{ij} & \text{if } i = j \\ w_{ij} & \text{otherwise} \end{cases}.$$

In Eq. (28), the vector $g^{n-1} = \sum\limits_{j\in\mathcal{N}_i}\sqrt{w_{ij}}\left(\nu_{ji}^{n-1} - \nu_{ij}^{n-1}\right) + \sum\limits_{j\in\mathcal{N}_i}\sqrt{w_{ij}}\left(b_{ji}^{n-1} - b_{ij}^{n-1}\right)$. Eq. (28) is a system of linear equations. The solution $\delta\mu^n$ can be acquired iteratively using the same method in A-FETV. Then we minimize the following subproblem with respect to ν

$$\nu^n = \arg\min_{\nu}\left\{\lambda\sum_{i=1}^{N}\|\nu_i\|_1 + \frac{\theta}{2}\sum_{i=1}^{N}\|\nu_i - \nabla_w(\delta\mu_i^n) - b_i^{n-1}\|_2^2\right\}, \tag{29}$$

which has an analytical solution, calculated from the generalized shrinkage formula

$$\nu_{ij}^n = \max\left(\left|\sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) + b_{ij}^{n-1}\right| - \frac{\lambda}{\theta}, 0\right)\frac{\sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) + b_{ij}^{n-1}}{\left|\sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) + b_{ij}^{n-1}\right|}, \tag{30}$$

with the convention that $0/0 = 0$. Lastly, we update the iterative parameter b with

$$b_{ij}^n = b_{ij}^{n-1} + \sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) - \nu_{ij}^n. \tag{31}$$

We can similarly apply ADMM to the minimization of I-GTV, which can be transformed into the following unconstrained problem with the auxiliary splitting vector variable ν, an augmented Lagrangian multiplier b, and a positive penalty parameter θ.

$$\delta\mu^n, \nu^n = \arg\min_{\delta\mu,\nu}\left\{\frac{1}{2}\|J\delta\mu - \delta\Phi\|_2^2 + \lambda\sum_{i=1}^{N}\|\nu_i\|_2 + \frac{\theta}{2}\sum_{i=1}^{N}\|\nu_i - \nabla_w(\delta\mu_i) - b_i^{n-1}\|_2^2\right\}. \tag{32}$$

The L_2 subproblem with respect to $\delta\mu$ is the same as the one in A-GTV and can be computed with Eq. (28). We then fix $\delta\mu$ to minimize the second subproblem with respect to ν:

$$\nu^n = \arg\min_{\nu}\left\{\lambda\sum_{i=1}^{N}\|\nu_i\|_2 + \frac{\theta}{2}\sum_{i=1}^{N}\|\nu_i - \nabla_w(\delta\mu_i^n) - b_i^{n-1}\|_2^2\right\}, \tag{33}$$

which can be solved with the following soft thresholding equation

$$\nu_{ij}^n = \max\left(\sqrt{\sum_{j\in\mathcal{N}_i}\left(\sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) + b_{ij}^{n-1}\right)^2} - \frac{\lambda}{\theta}, 0\right)\frac{\sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) + b_{ij}^{n-1}}{\sqrt{\sum\limits_{j\in\mathcal{N}_i}\left(\sqrt{w_{ij}}\left(\delta\mu_j^n - \delta\mu_i^n\right) + b_{ij}^{n-1}\right)^2}}, \tag{34}$$

Biomedical Optics EXPRESS

Algorithm 2: ADMM-based algorithm for I-GTV and A-GTV.

INPUT: $\delta\Phi$, J, y, $inner_loop$, ϵ_1, regularization parameter $\theta > 0$, $\lambda > 0$

Initialization: $v^0 = b^0 = 0$

for $n = 1 : inner_loop$

1: Update $\delta\mu^n$ using Eq. (28)

2: Update v^n using Eq. (30) for A-GTV or Eq. (34) for I-GTV

3: Update b^n using Eq. (31)

4: Stop if $n = inner_loop$ or $\frac{||\delta\mu^n - \delta\mu^{n-1}||_1}{||\delta\mu^{n-1}||_1} \leq \epsilon_1$.

RETURN $\delta\mu^k = \delta\mu^n$

with the convention that $0/0 = 0$. The update of iterative parameter b is the same as for A-GTV, as shown in Eq. (31). The ADMM-based algorithm for A-GTV and I-GTV is given in Algorithm 2.

Therefore the whole procedure for minimizing the TV-regularized inverse problem (Eq. (15)) is given in Algorithm 3, in which $outer_loop$ represents the maximum number of iterations required for the DOT reconstruction.

Algorithm 3: Algorithm for minimizing the TV-associated inverse problem.

INPUT: Φ^{M}, $\mathcal{F}(\cdot)$, μ^0, $outer_loop$, ϵ_2

for $k = 1 : outer_loop$

1: Compute $\mathcal{F}\left(\mu^{k-1}\right)$ and J^{k-1}

2: Set $\delta\Phi^{k-1} = \Phi^{\mathrm{M}} - \mathcal{F}\left(\mu^{k-1}\right)$

3: Compute $\delta\mu^k$ by introducing $\delta\Phi^{k-1}$ and J^{k-1} to one of **Algorithm 1-2**

4: Update $\mu^k = \delta\mu^k + \mu^{k-1}$

5: Stop if $k = outer_loop$ or $\frac{||\mathcal{F}(\mu^k) - \Phi^{\mathrm{M}}||_2^2 - ||\mathcal{F}(\mu^{k-1}) - \Phi^{\mathrm{M}}||_2^2}{||\mathcal{F}(\mu^{k-1}) - \Phi^{\mathrm{M}}||_2^2} \leq \epsilon_2$.

end for

RETURN μ^k

4. Experiments

In this section, we describe extensive experiments to qualitatively and quantitatively evaluate the performance of finite element and graph representations on the two variants of TV regularization in DOT. We use the finite element representation for the forward modelling in all the experiments and use both the finite element and graph representations to discretize the TV regularization term during the solution of the inverse problem. We first define four evaluation metrics to quantify the quality of the reconstructed images. Then we describe simulated numerical experiments on 2D circle and 3D head samples, and real experiments performed on phantom samples. Fig. 2 shows the unstructured grids of the three computational domains. Red dots represent the vertices in

Biomedical Optics EXPRESS

the computational domain. In 2D, using the finite element representation, the computational domain is discretized with a finite number of triangles (Fig. 2 (d)) while in 3D, tetrahedra are taken as the basic element (Fig. 2 (e)). However the graph representation is the same in both 2D and 3D because the graph method requires only vertices and edges of the mesh. For simulated experiments in which measurement noise was added, ten repeats were performed. In all experiments, the forward model was implemented using the NIRFAST package [45] in Matlab R2017a (Mathworks, Natick, USA). The simulated experiments conducted are all based on single wavelength continuous-wave (CW) measurements where the optical property to be recovered is the tissue absorption coefficient μ_a at that wavelength. We set $inner_loop$ to 100, $outer_loop$ to 40, ϵ_1 to 0.001 and ϵ_2 to 2% for all experiments in this paper.

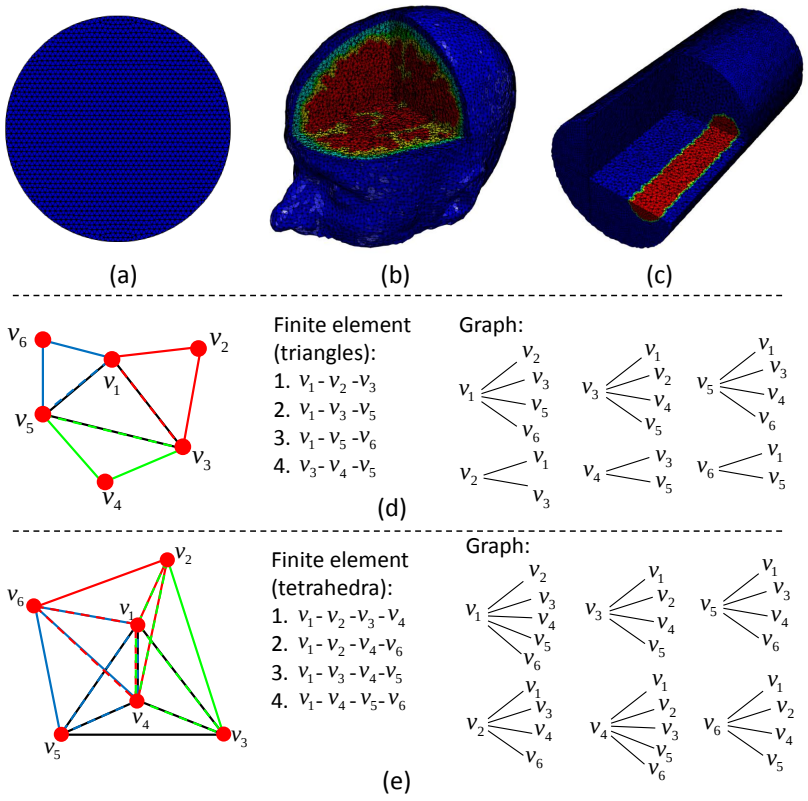

Fig. 2. (a)-(c): Discretized computational domain of the three experimental samples; (d): Detailed mesh composition of 2D geometry in finite element and graph representation respectively; (e): Detailed mesh composition of 3D geometry in finite element and graph representation respectively.

4.1. Quantitative evaluation metrics

The quantitative evaluation is performed using four evaluation metrics: the localization error, average contrast, peak signal-to-noise ratio (PSNR) and relative recovered volume. If the reconstructed image is identical to the ground truth image, the localization error is 0, average contrast and relative recovered volume are both equal to 1. PSNR is higher if the reconstructed image is closer to the ground truth image.

Localization error is defined as the Euclidian distance between the central nodes X_s of the simulated activation region and X_r of the recovered activation region. The recovered activation

Biomedical Optics EXPRESS

regions are selected by thresholding the recovered changes based on 60% of the maximum recovered changes.

$$\text{Localization error} = \|\mathrm{X_s} - \mathrm{X_r}\|_2. \tag{35}$$

The second evaluation metric is the average contrast which is based on the mean value of the region of interest:

$$\text{Average contrast} = \left(\sum_{i=1}^{N_r} \mu^i / N_r\right) / \widetilde{\mu}, \tag{36}$$

where μ^i denotes the recovered optical property on the finite element node i. N_r is the number of nodes in the recovered activation region. $\widetilde{\mu}$ is the ground truth values of the optical property in the recovered activation region.

PSNR is the third evaluation metric, which aims to evaluate the difference between the ground truth image and the recovered image. Larger PSNR values means less difference between these two types of images. PSNR is defined as follows

$$\text{PSNR} = 10 \cdot \log_{10}\left(\mathrm{MAX}_\mu^2 / \mathrm{MSE}\right). \tag{37}$$

Here, MAX_μ is the maximum pixel value of μ and MSE is the mean squared error between the recovered and ground truth images with $\mathrm{MSE} = \sum_{i=1}^{N}\left(\mu^i - \tilde{\mu}^i\right)^2 / N$.

Finally, we measure the relative recovered volume which is given as

$$\mathrm{V}_{RRV} = \mathrm{V}_r / \mathrm{V}_s \times 100\%, \tag{38}$$

where V_r and V_s denote the volume of the recovered activation region and simulated activation region, respectively.

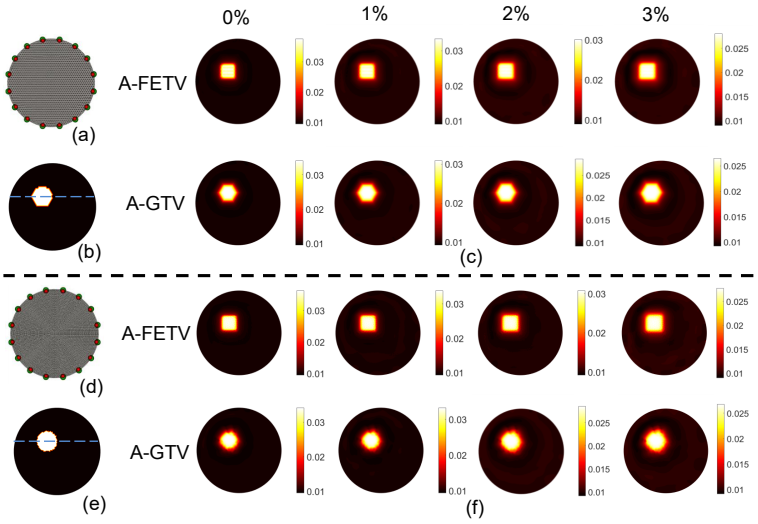

Fig. 3. (a)-(c): Reconstruction on the 2D mesh with low spatial resolution. (d)-(f): Reconstruction on the 2D mesh with high spatial resolution. (a) and (d): 2D reconstruction mesh with sixteen co-located sources and detectors. (b) and (e) give the original target distributions. First row in (c) and (f) represents the results using A-FETV on 0% , 1% , 2% and 3% noisy data while the second row shows the results using A-GTV.

Biomedical Optics EXPRESS

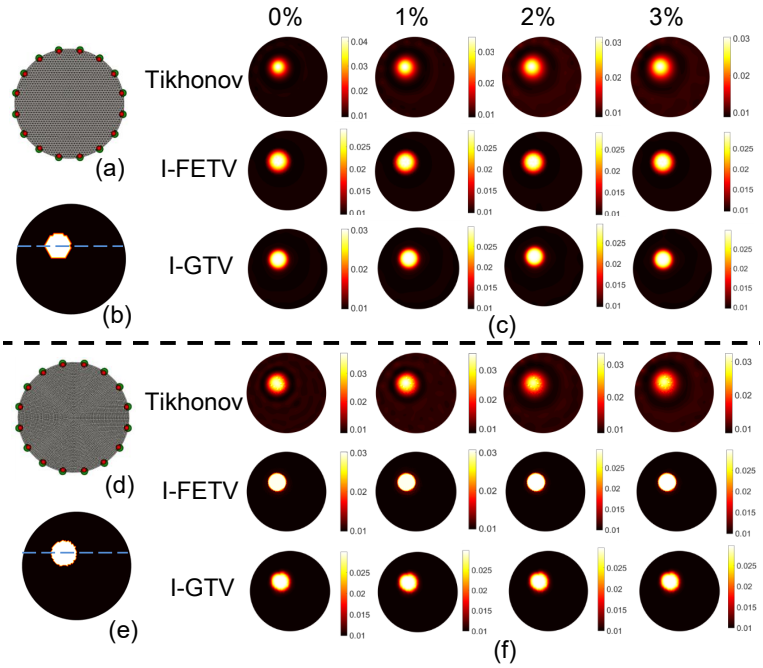

Fig. 4. (a)-(c): Reconstruction on the 2D mesh with low spatial resolution; (d)-(f): Reconstruction on the 2D mesh with high spatial resolution. (a) and (d): 2D reconstruction mesh with sixteen co-located sources and detectors. (b) and (e) give the original target distributions. First row in (c) and (f) represents the results using I-FETV on 0% , 1% , 2% and 3% noisy data while the second row shows the results by I-GTV.

4.2. *Experiments on anisotropic TV regularization*

Anisotropic TV regularization is easy to implement because the partial derivatives along different directions can be decoupled as explained in Section 3.2. It is based on the assumption that the shape of the region of interest is aligned with the coordinate axes. Its minimization favors horizontal and vertical structures, because oblique structures cause the TV regularization to increase [36]. In DOT, this assumption does not necessarily hold as the region of interest is normally random and structures are not normally aligned with the coordinate system. Therefore anisotropic TV regularization seems to be a poor choice for discrete TV in DOT, as it yields 'blocky' artefacts. However no research has been carried out about the relationship between the 'blocky' artefacts and the representation employed to discretize over the unstructured computational domain. In this section we investigate the anisotropic TV regularization in DOT reconstruction and compare their FE- and graph-based implementations. The effect of the representation method adopted on anisotropic TV regularization will be evaluated.

A 2D circular geometry is simulated with one anomaly centered at (-10mm,10mm). The 2D model has a radius of 43mm while the radius of the anomaly is 10mm. Sixteen source-detector fibres are placed equidistant around the external boundary for CW boundary data acquisition. When one fibre as a source is turned on, the rest are used as detectors, leading to 240 total boundary data points per wavelength. All sources were positioned one scattering distance within the outer boundary because the source is assumed to be spherically isotropic. In order to evaluate the effect of mesh resolution on the representation method, two reconstruction meshes are created with different spatial resolutions. The coarser mesh has 1785 nodes and 3418 linear triangle elements with the average element size 1.6977mm^2 (Fig. 3 (a)) while the finer one has 5133 nodes

Biomedical Optics EXPRESS

and 10013 elements with the average element size 0.5801mm^2 (Fig. 3 (d)). The background absorption coefficient μ_a is set as 0.01mm^{-1} and μ_a for the anomaly is set as 0.03mm^{-1} (Fig. 3 (b) and (e)). μ_s remains constant as 1mm^{-1}. To represent various realistic cases, normally distributed randomly generated Gaussian noise ranging from 0% to 3% at 1% intervals was added to the boundary measurements. Reconstructed images of the absorption coefficient are shown in Figs. 3 (c) and (f).

4.3. Experiments on isotropic TV regularization

4.3.1. Two dimensional circular experiments

Using the same reconstruction meshes described in section 4.2, we compare I-FETV, I-GTV against a baseline Tikhonov model. To represent various realistic cases, normally distributed randomly generated noise ranging from 0% to 3% at 1% intervals was added to the amplitude of the boundary data. Reconstructed images of absorption coefficient are shown in Figs. 4 (c) and (f). The 1D cross sections and evaluation metrics comparisons are displayed in Fig. 5 and Fig. 6.

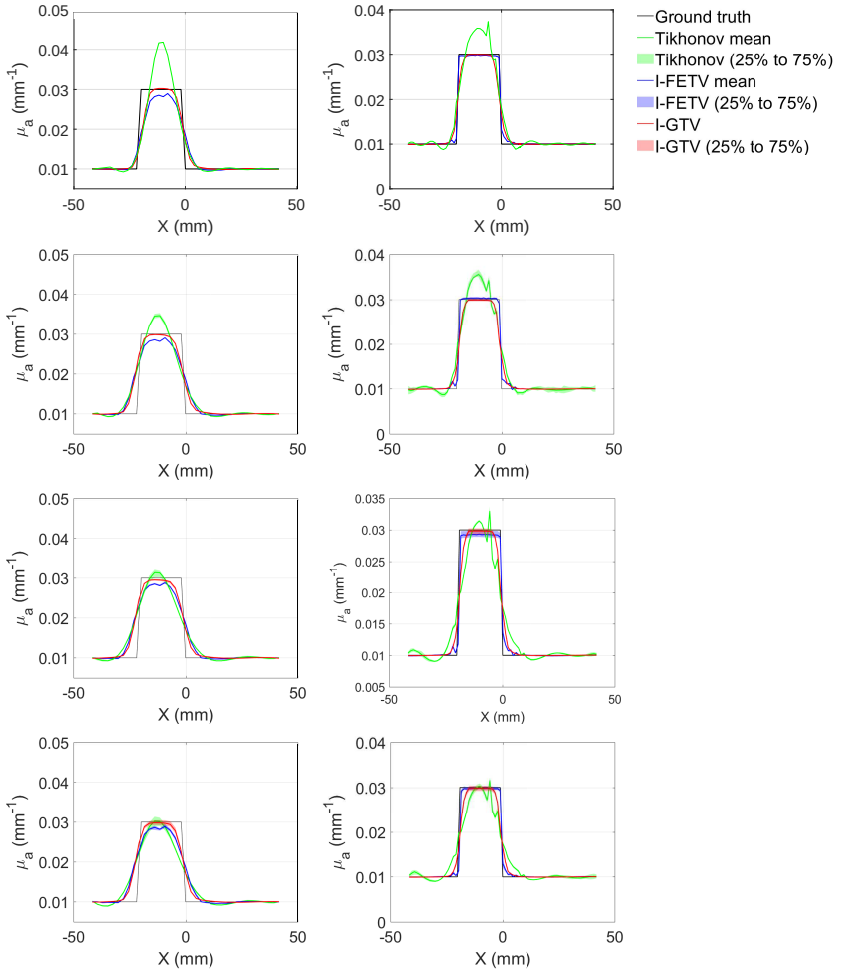

Fig. 5. 1D cross section of images recovered in Fig. 4. First column corresponds to Fig. 4 (c) where the spatial resolution of the reconstruction mesh is lower. Second column corresponds to Fig. 4 (f) where the spatial resolution of the reconstruction mesh is higher. Top to bottom row: 0%, 1%, 2% and 3% added Gaussian noise.

Biomedical Optics EXPRESS

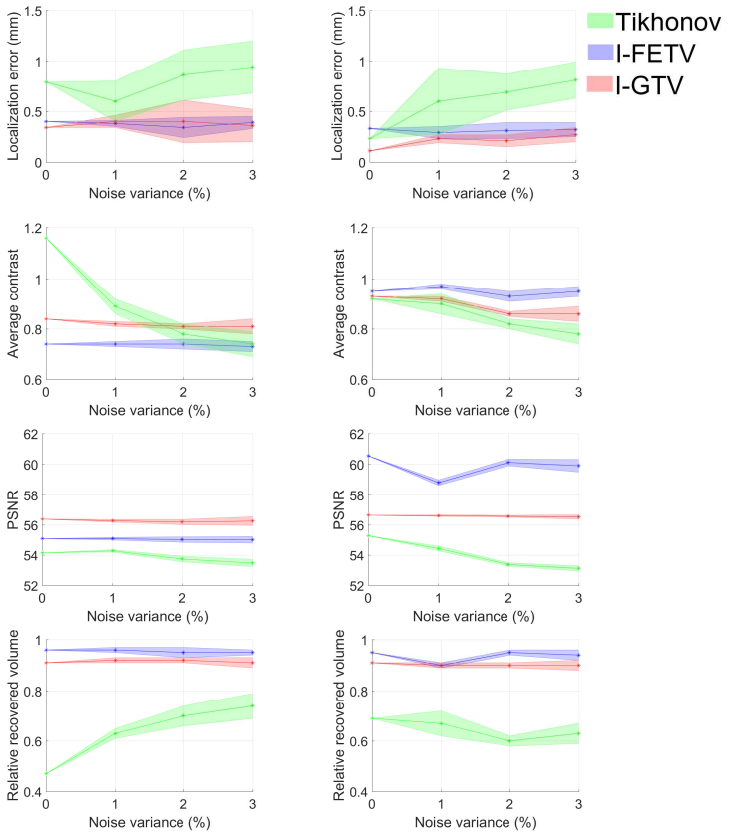

Fig. 6. Evaluation metrics comparing the performance of different methods at four different noise levels. Top to bottom row: localization error index; average contrast index; PSNR index and relative recovered volume. Left column corresponds to the reconstructions in Fig. 4 (c) where the reconsturction mesh resolution is low. Right column corresponds to Fig. 4 (f) where the reconsturction mesh resolution is relatively high.

Table 2. Head tissue optical property for each of five layers.

	Scalp	Skull	CSF	Gray Matter	White Matter
μ_a (mm^{-1})	0.017	0.012	0.004	0.018	0.017
μ_s' (mm^{-1})	0.74	0.94	0.3	0.84	1.19

4.3.2. Three dimensional head numerical experiments

We now evaluate the isotropic TV model with two discrete differential operator definitions on a physically realistic 3D head model. This was created from T1-weighted MPRAGE scans acquired by Eggebrecht *et al* [7] using the process described by Wu et al [46] in which Statistical Parametric Mapping (SPM) software [47] was used to perform parametric segmentation of five tissue types (scalp, skull, cerebrospinal fluid (CSF), gray matter, white matter) based on the pixel intensity probability function distribution. The five layers were processed in NIRFAST to create masks and layered volumetric FEM meshes. The reconstruction mesh consists of 50721 nodes associated with 287547 tetrahedral elements, with the average element size 9.2676mm^3. Each

node is labeled by one of the five segmented head tissue types. Absorption coefficients assigned to each layer are from an in vivo study [48] at 750nm (Table 2).

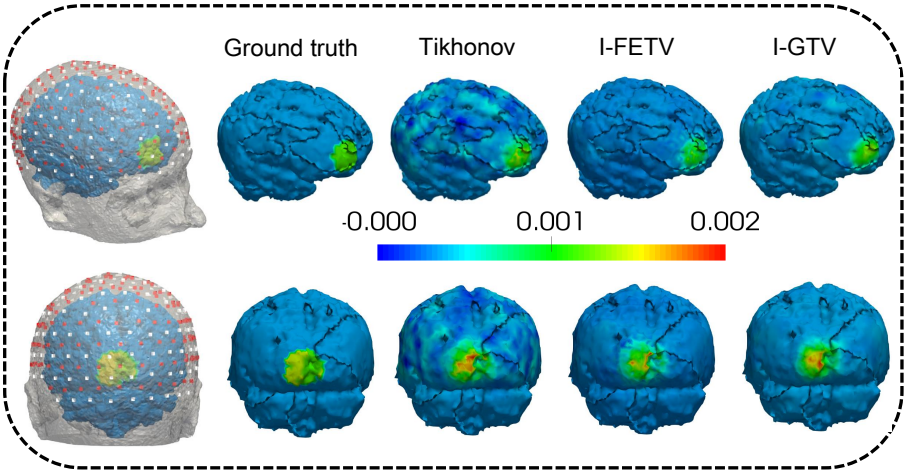

Fig. 7. First column: distribution of the imaging array with 158 sources (red dots) and 166 detectors (white dots) and the positions of the two simultaneous simulated anomalies. Second to final column: Ground truth and reconstructions by Tikhonov, I-FETV and I-GTV.

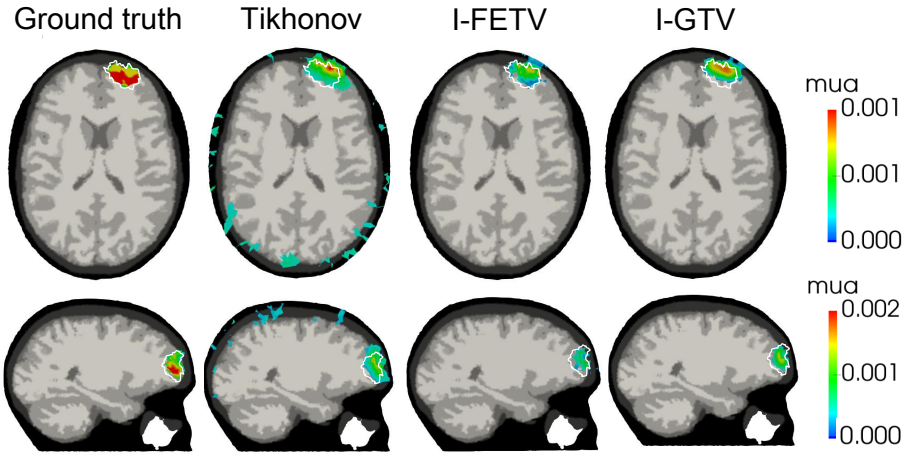

Fig. 8. 2D slices of the reconstructions of the absorption coefficient changes on the forehead anomaly (first row in Fig. 7). The ground truth areas are highlighted in white ellipses.

A high-density (HD) imaging array with 158 sources and 166 detectors (Fig. 7 first column) [7] was placed over the whole head, with source-detector (SD) separation distances ranging from 1.3 to 4.8cm. In this study, 3478 differential measurements per wavelength were used to image hemodynamic changes in the brain. Two distinct anomalies were simulated simultaneously in the brain, with each 15mm radius. In order to simulate the traumatic brain injury (TBI) cases where tissue oxygen saturation (StO_2) is normally between 50% and 75% [49,50], the absorption coefficient in the two anomalies are calculated using Beer's law [45] with 55% StO_2 (Fig. 7 second column). In line with the expected in vivo performance of the imaging system, 0.12%, 0.15%, 0.41% and 1.42% Gaussian random noise was added to first (13mm), second (30mm), third (40mm) and fourth (48mm) nearest neighbor measurements to provide realistic data [51].

Biomedical Optics EXPRESS

Reconstructed absorption coefficients using different model are displayed in the third to fifth column of Fig. 7. Corresponding 2D slices are displayed in Figs. 8 and 9 and the evaluation metrics are presented in Fig. 10.

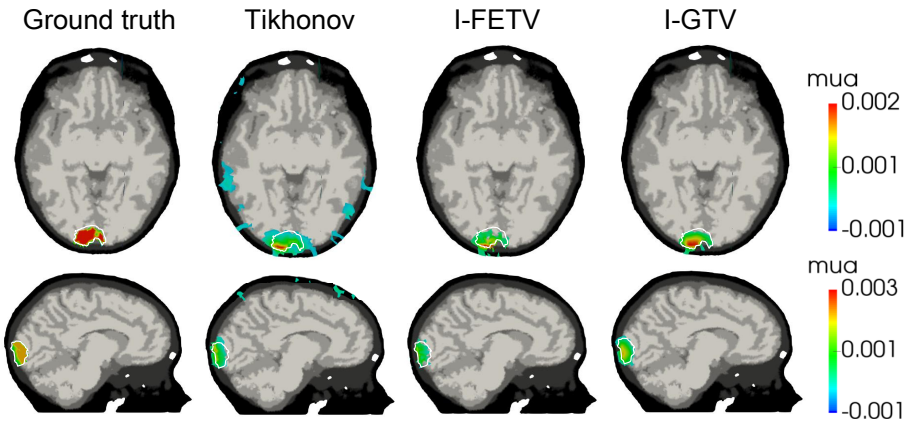

Fig. 9. 2D slices of the reconstructions of the absorption coefficient changes on the back-head anomaly (second row in Fig. 7). The ground truth areas are highlighted in white ellipses.

4.3.3. Experiments with phantom data

In the final experiment we evaluate different methods on real experimental data which is collected from a solid plastic cylindrical phantom using one non-contact CW-DOT system designed for hand imaging [52]. The phantom has size of radius 12.3mm and length 50mm. 35 sources and 99 detectors are positioned on the underside and top of the phantom respectively (Fig. 11(a)). The absorbing dye within the phantom was treated as a chromophore that has unit concentration in the bulk of the phantom. Its extinction coefficient was modelled by the measured absorption coefficient. A cylindrical rod was placed at the depth of 5mm to simulate the heterogeneous version of the phantom (Fig. 11(a)). The rod has radius 3mm and length 50mm and provides a 2:1 contrast in dye concentration compared to background (Fig. 11 (c) top row). Five wavelengths (650nm, 710nm, 730nm 830nm and 930nm) in a transmission setup are used to collect the boundary data. The reconstruction mesh consists of 9082 nodes and 48099 linear tetrahedral elements with the average tetrahedral elements size 0.4218mm^3. Ground truth data and images reconstructed with Tikhonov, I-FETV and I-GTV are shown in Fig. 11(c). The four evaluation metrics in the volume of illumination are given in Table 3. For all the experiments above, the regularization parameter λ is determined using an L-curve method [13].

5. Discussion

The 2D images reconstructed using A-FETV and A-GTV are shown in Fig. 3, together with the original target distributions. The results show that A-FETV keeps reconstructing the target with boundaries that align with the coordinate axes which is same to the assumption of the anisotropic TV regularization. In addition, some artefacts are observed inside the recovered region of interest when the mesh resolution is low. The reconstructions using A-GTV do not feature these artefacts, and the recovered shape is more accurate, with no bias towards the coordinate axes. The experiments reveal that the blocky artefacts of anisotropic TV regularization are associated with the discretization method used. The blocky artefacts are clearly visible in reconstructions based on the finite element representation, but not in the ones based on the graph representation. This is because in the graph representation, the region is discretized along all edge-based directions,

Biomedical Optics EXPRESS

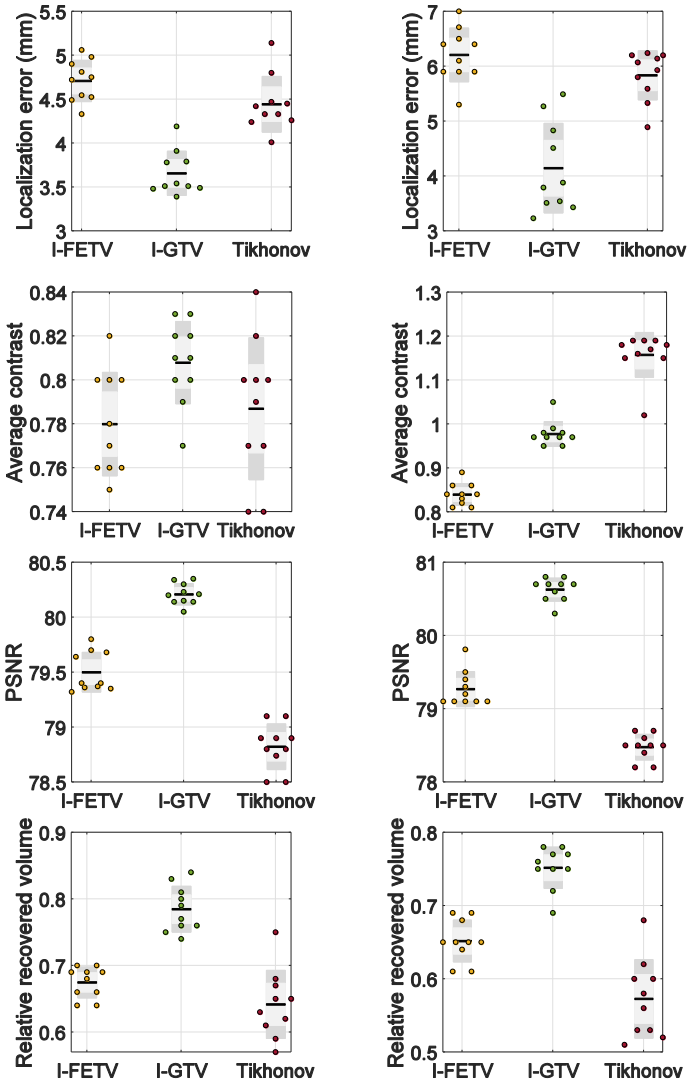

Fig. 10. Evaluation metrics comparing the performance of different methods on a 3D head model. The left column represents the reconstruction of the forehead anomaly (first row in Fig. 7), while the right column gives the reconstruction of the back-head anomaly (second row in Fig. 7).

leading to nearly isotropic solutions. Therefore in DOT, A-GTV can adapt to the ground truth solution. However, it is not a good method to preserve anisotropy if anisotropy is a desired property of the solution.

For the 2D case which uses isotropic TV regularization, Fig. 4, Tikhonov reconstruction over-smooths the results and smears the edges. The results become smoother with increases in measurement noise. Little difference can be visually observed between the reconstruction by I-FETV and I-GTV when the reconstruction mesh resolution is low. However when the reconstruction mesh has higher resolution, Fig. 4 (f), the results by I-FETV is visually closer to the ground truth than the ones by I-GTV. Similar findings are observed from the corresponding 1D cross sections (Fig. 5). Tikhonov reconstruction produces a single peaked distribution in the

Biomedical Optics EXPRESS

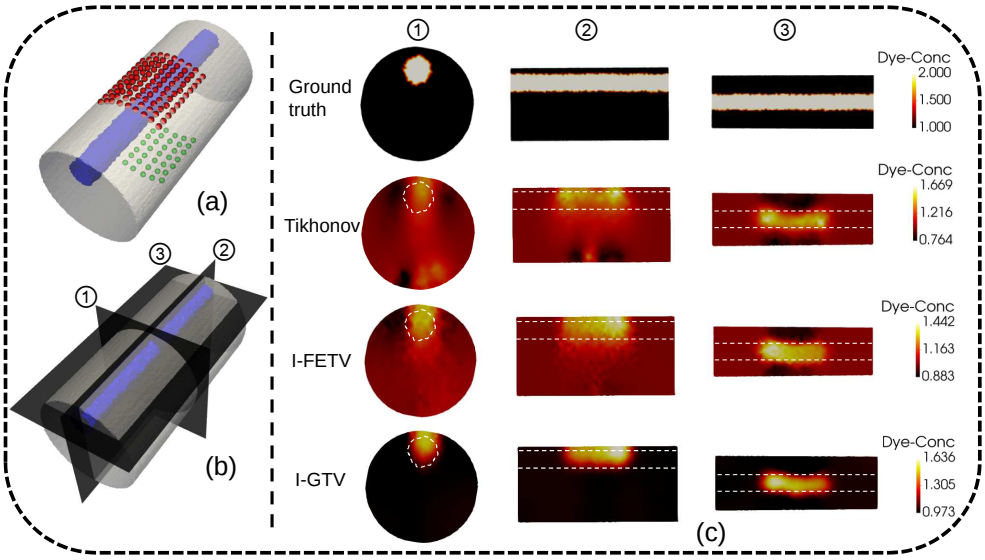

Fig. 11. (a): Distribution of sources and detectors. (b): Illustration of the overall distribution of three slices. (c): Ground truth and reconstruction results with different methods. From top to bottom: ground truth; results with Tikhonov regularization; results with I-FETV regularization and results with I-GTV regularization.

piecewise constant target area, and edges of the objects are over-smoothed. Both TV methods are able to reconstruct a piecewise constant distribution. However when the mesh resolution is lower (first column in Fig. 5), fluctuations in the target regions are observed in the results by I-FETV. When the mesh resolution is higher (second column in Fig. 5), the cross-section from I-FETV reconstruction is almost identical to the ground truth. In Fig. 6, red and blue areas represent 25% to 75% value among the ten repeats' experiments. We see that the performance of I-FETV improves with an increase of mesh resolution: by 25% in localization error, 26% in average contrast and 11% in PSNR, while the performance of I-GTV is relatively unaffected by the mesh resolution. These 2D experiments confirm that the discrete differential operators based on graph representation are not affected by the mesh resolution while the ones based on a finite element representation become more accurate when the reconstruction mesh is finer.

The 3D images reconstructed from the head geometry represent a physically realistic case, in which two anomalies are simulated simultaneously in the brain. From Fig. 7, Tikhonov reconstruction lead to many visible artefacts near the source and detector area. Due to smoothing induced by Tikhonov regularization, sharp features are not present in the image recovered. I-FETV and I-GTV both can eliminate the surface artefacts resulting from Tikhonov regularization and reconstruct tightly localized results. These findings can be clearly observed in the 2D slice images shown in Fig. 8 and Fig. 9. It should be noticed that, in Fig. 8, the colorbar values corresponding to the green and red parts remain 0.001. It is because only three digits are selected after the radix point and in this study we use rounding off to constrain the three digits. From the visualization of the results, there is no obvious difference between the reconstruction performance of I-FETV and I-GTV because both are based on TV regularization. However it can be observed from the evaluation metrics comparison in Fig. 10 that I-GTV achieves the lowest localization error, highest peak signal-to-noise ratio and average contrast much closer to 1. The average relative recovered volume achieved by I-GTV is 77%, compared with I-FETV (66%) and Tikhonov (64%). This experiment confirms the lower performance of I-FETV on reconstruction meshes with low spatial resolution.

Biomedical Optics EXPRESS

In the experiments with phantom data, only the central region is reconstructed in all the cases because the positions of sources and detectors lead to very low sensitivity away from the centre. It can be seen from the second row of Fig. 11(c), that Tikhonov regularization over-smooths the reconstructed images which have much lower image contrast than the ground truth, especially in the first slice image. Artefacts are clearly observed near the source and detector areas. Even though total variation regularization can alleviate the over-smoothing effect caused by Tikhonov regularization, discretization methods still play an important role in the reconstruction performance. It should be noticed that I-FETV can alleviate the artefacts near to the sources and detectors but introduce some artefacts (staircase effect) in the non-anomaly area and does not preserve edges. I-GTV is seen to recover the anomaly with clear edges and high image contrast. It is interesting to compare these results to those of our previous work [13] where L_1 regularization was applied to the phantom data. Reconstructions by L_1 regularization were found to be over-sparsified and over-compact. In this work, TV regularization, which induces sparsity to the gradient of the solution, is seen to effectively alleviate the over-sparsifying effect of L_1 regularization and is therefore suitable for non-sparse coefficient distributions. We calculate the four evaluation metrics in the volume of illumination (Table 3) and these support the same conclusions. Similar localization errors are obtained by the different methods with only 1mm difference. Comparing to I-FETV, I-GTV can obtain the highest average contrast and PSNR values with similar relative recovered volume.

Table 3. Evaluation of different methods for reconstruction on a tissue-simulating phantom.

	Localization error / mm	Average contrast / -	PSNR / -	Relative recovered volume / %
Tikhonov	2.90	0.74	13.74	40
I-FETV	2.81	0.69	14.77	48
I-GTV	3.16	0.79	16.71	46

6. Conclusion

In this paper, we introduce finite element and graph representations to discretize the TV regularization term in DOT reconstruction. Isotropic and anisotropic variants of the TV regularization are also investigated and compared between their FE- and graph-based implementations. The ADMM-based algorithms are proposed for each TV-regularized inverse problem. Experiments on the anisotropic TV regularization reveal that finite element representation yields the 'blocky' artefacts which is the designed in feature in the anisotropic TV regularization. However the graph representation favors the underlying shape of the region of interest so that the 'blocky' artefacts are not realized. Graph discretization on anisotropic TV regularization can adapt to the ground truth solution, but is not a good way to preserve anisotropy.

Numerical experiments on isotropic TV regularization illustrate that, comparing to Tikhonov regularization, TV regularization can alleviate the over-smoothing effect of Tikhonov regularization and localize the anomaly by inducing sparsity of the gradient of the solution. These findings were tested on real experimental data using a tissue-simulating phantom. I-FETV does not perform well on low resolution reconstruction meshes because of the discrete nature of the finite element representation. Because the finite element representation operates on each element, the discretization becomes more accurate when the mesh resolution increases. I-GTV is shown to be more stable and robust to changes in mesh resolution because I-GTV is discretized on the graph directly, having no information of elements. Hence I-GTV can give more accurate discretization when the reconstruction mesh is a coarse mesh which is the usual case in DOT. However, I-FETV will outperform I-GTV when an reconstruction mesh with high resolution is available.

Biomedical Optics EXPRESS

Funding

Horizon 2020 Framework Programme (675332).

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Marie Sklodowska-Curie Innovative Training Networks (ITN-ETN) programme, under grant agreement no 675332, BitMap. We thank Daniel Lighter for providing the experimental data used in Section 4.3.3. Supporting materials can be freely downloaded from https://doi.org/10.25500/eData.bham.00000283.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical tomography," Phys. Med. Biol. **50**(4), 0031–9155 (2005).
2. S. Srinivasan, B. W. Pogue, S. Jiang, H. Dehghani, C. Kogel, S. Soho, J. J. Gibson, T. D. Tosteson, S. P. Poplack, and K. D. Paulsen, "Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured in vivo by near-infrared breast tomography," Proc. Natl. Acad. Sci. **100**(21), 12349–12354 (2003).
3. H. Dehghani, B. W. Pogue, S. P. Poplack, and K. D. Paulsen, "Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results," Appl. Opt. **42**(1), 135–145 (2003).
4. D. A. Boas, K. Chen, D. Grebert, and M. A. Franceschini, "Improving the diffuse optical imaging spatial resolution of the cerebral hemodynamic response to brain activation in humans," Opt. Lett. **29**(13), 1506–1508 (2004).
5. A. Custo, D. A. Boas, D. Tsuzuki, I. Dan, R. Mesquita, B. Fischl, W. Eric. L. Grimson, and W. Wells, "Anatomical atlas-guided diffuse optical tomography of brain activation," Neuroimage **49**(1), 561–567 (2010).
6. H. Niu, S. Khadka, F. Tian, Z. J. Lin, C. Lu, C. Zhu, and H. Liu, "Resting-state functional connectivity assessed with two diffuse optical tomographic systems," J. Biomed. Opt. **16**(4), 046006–046006 (2011).
7. A. T. Eggebrecht, S. L. Ferradal, A. R. Viehoever, M. S. Hassanpour, H. Dehghani, A. Z. Snyder, T. Hershey, and J. P. Culver, "Mapping distributed brain function and networks with diffuse optical tomography," Nat. Photon. **8**(6), 448–454 (2014).
8. S. R. Arridge and J. C. Hebden, "Optical imaging in medicine: ii. modelling and reconstruction," Phys. Med. Biol. **42**(5), 841 (1997).
9. P. K. Yalavarthy, B. W. Pogue, H. Dehghani, C. M. Carpenter, S. Jiang, K. D. Paulsen, "Structural information within regularization matrices improves near infrared diffuse optical tomography," Opt. Express **15**(13), 8043–8058 (2007).
10. C. B. Shaw and P. K. Yalavarthy, "Effective contrast recovery in rapid dynamic near-infrared diffuse optical tomography using l1-norm-based linear image reconstruction method," J. Biomed. Opt. **17**(8), 0860091 (2012).
11. J. C. Baritaux, K. Hassler, M. Bucher, S. Sanyal, and M. Unser, "Sparsity-driven reconstruction for FDOT with anatomical priors," IEEE Trans. Med. Imaging **30**(5), 1143–1153 (2011).
12. V. C. Kavuri, Z. J. Lin, F. Tian, and H. Liu, "Sparsity enhanced spatial resolution and depth localization in diffuse optical tomography," Biomed. Opt. Express **3**(5), 943–957 (2012).
13. W. Lu, D. Lighter, and I. B. Styles, "L1-norm based nonlinear reconstruction improves quantitative accuracy of spectral diffuse optical tomography," Biomed. Opt. Express **9**(4), 1423–1444 (2018).
14. Q. Lyu, Z. Lin, Y. She, and C. Zhang, "A comparison of typical L_p minimization algorithms," Neurocomputing **119**, 413–424 (2013).
15. J. Prakash, C. B. Shaw, R. Manjappa, R. Kanhirodan, and P. K. Yalavarthy, "Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction," IEEE J. Sel. Top. Quantum. Electron. **20**(2), 74–82 (2014).
16. S. Okawa, Y. Hoshi, and Y. Yamada, "Improvement of image quality of time-domain diffuse optical tomography with lp sparsity regularization," Biomed. Opt. Express **2**(12), 3334–3348 (2011).
17. S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," Multiscale Model. Simul. **4**(2), 460–489 (2005).
18. J. Duan, Z. Pan, B. Zhang, W. Liu, and X. C. Tai, "Fast algorithm for color texture image inpainting using the non-local CTV model," J. Glob. Optim. **62**(4), 853–876 (2015).
19. J. Duan, W. Lu, C. Tench, I. Gottlob, F. Proudlock et al., "Denoising optical coherence tomography using second order total generalized variation decomposition," Biomed. Signal Process. Control **24**, 120–127 (2016).
20. L. Yao and H. Jiang, "Enhancing finite element-based photoacoustic tomography using total variation minimization," Appl. Op. **50**(25), 5031–5041 (2011).
21. H. Gao and H. Zhao, "Multilevel bioluminescence tomography based on radiative transfer equation part 1: L_1 regularization," Opt. Express **18**(3), 1854–1871 (2010).

Biomedical Optics EXPRESS

22. M. Freiberger, C. Clason, and H. Scharfetter, "Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting approach," Appl. Opt. **49**(19), 3741–3747 (2010).
23. A. B. Konovalov and V. V. Vlasov, "Total variation based reconstruction of scattering inhomogeneities in tissue from time-resolved optical projections," In PALS 9917: 99170S, International Society for Optics and Photonics (2016).
24. K. D. Paulsen and H. Jiang, "Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization," Appl. Opt. **35**(19), 3447–3458 (1996).
25. J. Tang, B. Han, W. Han, B. Bi, and L. Li, "Mixed total variation and regularization method for optical tomography based on radiative transfer equation," Comput. Math. Methods Med. 2017 (2017).
26. W. Lu, J. Duan, Z. Qiu, Z. Pan, R. W. Liu, and L. Bai, "Implementation of high-order variational models made easy for image processing," Math. Methods Appl. Sci. **39**, 4208–4233 (2016).
27. I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," Commun. Pure Appl. Math. **63**(1), 1–38 (2010).
28. T. Pock, D. Cremers, H. Bischof, and A. Chambolle, "An algorithm for minimizing the Mumford-Shah functional," In *Computer Vision, 2009 IEEE 12th International Conference on*, 1133–1140. IEEE (2009).
29. T. Goldstein and S. Osher, "The split bregman method for L_1-regularized problems," SIAM J. Imaging Sci. **2**(2), 323–343 (2009).
30. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. Imaging Sci. **2**(1), 183–202 (2009).
31. T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a FASTA implementation," arXiv preprint arXiv:1411.3406 (2014).
32. K. Papafitsoros and C. B. Schönlieb, "A combined first and second order variational approach for image reconstruction," J. Math. Imaging Vis. **48**(2), 308–338 (2014).
33. J. Duan, Z. Qiu, W. Lu, G. Wang, Z. Pan, and L. Bai, "An edge-weighted second order variational model for image decomposition," Digit. Signal Process. **49**, 162–181 (2016).
34. J. Duan, W. OC. Ward, L. Sibbett, Z. Pan, and L. Bai, "Introducing diffusion tensor to high order variational model for image reconstruction," Digit. Signal Process. **69**, 323–336 (2017).
35. W. H. Reed and T. R. Hill, "Triangular mesh methods for the neutron transport equation," Technical report, Los Alamos Scientific Lab., N. Mex.(USA) (1973).
36. G. González, V. Kolehmainen, and A. Seppänen, "Isotropic and anisotropic total variation regularization in electrical impedance tomography," Comput. Math. Appl. **74**(3), 564–576 (2017).
37. G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," Multiscale Model. Sim. **7**(3), 1005–1028 (2008).
38. J. Duan, Z. Pan, W. Liu, and X. C. Tai, "Color texture image inpainting using the non local CTV model," JSIP **4**(03), 43 (2013).
39. A. L. Bertozzi and A. Flenner, "Diffuse interface models on graphs for classification of high dimensional data," Multiscale Model. Simul. **10**(3), 1090–1118 (2012).
40. E. Merkurjev, T. Kostic, and A. L. Bertozzi, "An mbo scheme on graphs for classification and image processing," SIAM J. Imaging Sci. **6**(4), 1903–1930 (2013).
41. A. Elmoataz, O. Lezoray, and S. Bougleux, "Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing," IEEE T. Image Process. **17**(7), 1047–1060 (2008).
42. X. Bresson, X. C. Tai, T. F. Chan, and A. Szlam, "Multi-class transductive learning based on ?1 relaxations of cheeger cut and Mumford-Shah-Potts model," J. Math. Imaging Vis. **49**(1), 191–201 (2014).
43. C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the Nystrom method," IEEE Trans. Pattern Anal. Mach. Intell. **26**(2), 214–225 (2004).
44. S. R. Arridge and M Schweiger, "Photon-measurement density functions. part 2: Finite-element-method calculations," Appl. Opt. **34**(34), 8026–8037 (1995).
45. H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, et al., "Near infrared optical tomography using NIRFAST: Algorithms for numerical model and image reconstruction," Int. J. Numer. Method. Biomed. Eng. **25**(6), 711–732 (2009).
46. X. Wu, A. T. Eggebrecht, S. L. Ferradal, J. P. Culver, and H. Dehghani, "Quantitative evaluation of atlas-based high-density diffuse optical tomography for imaging of the human visual cortex," Biomed. Opt. Express **5**(11), 3882–3900 (2014).
47. J. Ashburner and K. J. Friston, "Image segmentation," Human Brain Function 2, (2003).
48. A. T. Eggebrecht, B. R. White, S. L. Ferradal, C. Chen, Y. Zhan, et al., "A quantitative spatial comparison of high-density diffuse optical tomography and FMRI cortical mapping," Neuroimage **61**(4), 1120–1128 (2012).
49. M. Clancy, A. Belli, D. Davies, S. JE. Lucas, Z. Su, and H. Dehghani, "Comparison of neurological NIRS signals during standing valsalva maneuvers, pre and post vasopressor injection," In European Conference on Biomedical Optics, 953817 (Optical Society of America, 2015).
50. C. Ichai, H. Quintard, and J. C. Orban, Metabolic Disorders and Critically Ill Patients: From Pathophysiology to Treatment (Springer, 2017).
51. H. Dehghani, B. R. White, B. W. Zeff, A. Tizzard, and J. P. Culver, "Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography," Appl. Opt. **48**(10), D137–D143 (2009).

Biomedical Optics EXPRESS

52. D. Lighter, A. Filer, and H. Dehghani, "Multispectral diffuse optical tomography of finger joints," In *European Conference on Biomedical Optics*, 104120N (Optical Society of America, 2017).

77.98.150.132 {ts '2022-01-04 11:21:36'}