

# Split operator method for fluorescence diffuse optical tomography using anisotropic diffusion regularisation with prior anatomical information

Teresa Correia,<sup>1,\*</sup> Juan Aguirre,<sup>2</sup> Alejandro Sisniega,<sup>2</sup>  
Judith Chamorro-Servent,<sup>2</sup> Juan Abascal,<sup>2</sup> Juan J. Vaquero,<sup>2</sup>  
Manuel Desco,<sup>2,3</sup> Ville Kolehmainen,<sup>4</sup> and Simon Arridge<sup>1</sup>

<sup>1</sup>Department of Computer Science, University College London, Gower Street, London WC1E6BT, UK

<sup>2</sup>Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Spain

<sup>3</sup>Unidad de Medicina y Cirugía Experimental, Hospital General Universitario Gregorio Marañón, CIBER de Salud Mental (CIBERSAM), Madrid, Spain

<sup>4</sup>Department of Applied Physics, University of Kuopio, PO BOX 1627, 70211 Kuopio, Finland

\*[tcorreia@cs.ucl.ac.uk](mailto:tcorreia@cs.ucl.ac.uk)

**Abstract:** Fluorescence diffuse optical tomography (fDOT) is an imaging modality that provides images of the fluorochrome distribution within the object of study. The image reconstruction problem is ill-posed and highly underdetermined and, therefore, regularisation techniques need to be used. In this paper we use a nonlinear anisotropic diffusion regularisation term that incorporates anatomical prior information. We introduce a split operator method that reduces the nonlinear inverse problem to two simpler problems, allowing fast and efficient solution of the fDOT problem. We tested our method using simulated, phantom and *ex-vivo* mouse data, and found that it provides reconstructions with better spatial localisation and size of fluorochrome inclusions than using the standard Tikhonov penalty term.

© 2011 Optical Society of America

**OCIS codes:** (100.3190) Inverse problems; (170.6960) Tomography.

---

## References and links

1. V. Ntziachristos, "Fluorescence molecular imaging," *Ann. Rev. Biomed. Eng.* **8**, 1–33 (2006).
2. V. Ntziachristos, C. Bremer, and R. Weissleder, "Fluorescence imaging with near-infrared light: new technological advances that enable in vivo molecular imaging," *Eur. Radiol.* **13**, 195–208 (2003).
3. G. Zacharakis, H. Kambara, H. Shih, J. Ripoll, J. Grimm, Y. Saeki, R. Weissleder, and V. Ntziachristos, "Volumetric tomography of fluorescent proteins through small animals *in vivo*," *Proc. Natl. Acad. Sci. U.S.A.* **12**, 18252–18257 (2005).
4. D. S. Keshpore, S. L. Gibbs-Strauss, J. A. OHara, M. Hutchins, N. Mincu, F. Leblond, M. Khayat, H. Dehghani, S. Srinivasan, and B. W. Pogue, "Imaging of glioma tumor with endogenous fluorescence tomography," *J. Biomed. Opt.* **14**, 030501 (2009).
5. A. Martin, J. A. J. A. Sarasa-Renedo, D. Tsoukatou, A. Garofalakis, H. Meyer, C. Mamalaki, J. Ripoll, and A. M. Planas, "Imaging changes in lymphoid organs in vivo after brain ischemia with three-dimensional fluorescence molecular tomography in transgenic mice expressing green fluorescent protein in T lymphocytes." *Molec. Imaging* **7**, 157–167 (2008).

6. A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, "Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans," *Opt. Express* **15**, 6696–6716 (2007).
7. V. Ntziachristos, A. G. Yodh, M. Schnall, and B. Chance, "Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement," *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2767–2772 (2000).
8. S. van de Ven, A. Wiethoff, T. Nielsen, B. Brendel, M. van der Voort, R. Nachabe, M. V. der Mark, M. V. Beek, L. Bakker, L. Fels, S. Elias, P. Luijten, and W. Mali, "A novel fluorescent imaging agent for diffuse optical tomography of the breast: First clinical experience in patients," *Molec. Imaging Biol.* **12**, 343–248 (2010).
9. S. R. Arridge and J. C. Schotland, "Optical tomography: forward and inverse problems," *Inv. Probl.* **25**, 123010 (2009).
10. A. Joshi, W. Bangerth, and E. M. Sevick-Muraca, "Adaptive finite element based tomography for fluorescence optical imaging in tissue," *Opt. Express* **12**, 5402–5417 (2004).
11. A. D. Zacharopoulos, P. Svenmarker, J. Axelsson, M. Schweiger, S. R. Arridge, and S. Andersson-Engels, "A matrix-free algorithm for multiple wavelength fluorescence tomography," *Opt. Express* **17**, 3025–3035 (2009).
12. A. Ale, R. B. Schulz, A. Sarantopoulos, and V. Ntziachristos, "Imaging performance of a hybrid x-ray computed tomography-fluorescence molecular tomography system using priors," *Med. Phys.* **37**, 1976–1986 (2010).
13. S. C. Davis, H. Dehghani, J. Wang, S. Jiang, B. W. Pogue, and K. D. Paulsen, "Image-guided diffuse optical fluorescence tomography implemented with Laplacian-type regularization," *Opt. Express* **15**, 4066–4082 (2007).
14. Y. Lin, H. Yan, O. Nalcioglu, and G. Gulsen, "Quantitative fluorescence tomography with functional and structural *a priori* information," *Appl. Opt.* **48**, 1328–1336 (2009).
15. C. R. Vogel, *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics, 2002).
16. M. Freiberger, H. Egger, and H. Scharfetter, "Nonlinear inversion in fluorescence optical tomography," *IEEE Trans. Biomed. Eng.* **57**, 2723–2729 (2010).
17. M. Freiberger, C. Clason, and H. Scharfetter, "Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting approach," *Appl. Opt.* **49**, 3741–3747 (2010).
18. A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Num. Math.* **76**, 167–188 (1997).
19. A. Douiri, M. Schweiger, J. Riley, and S. Arridge, "Local diffusion regularization method for optical tomography reconstruction by using robust statistics," *Opt. Lett.* **30**, 2439–2441 (2005).
20. A. Douiri, M. Schweiger, J. Riley, and S. R. Arridge, "Anisotropic diffusion regularization methods for diffuse optical tomography using edge prior information," *Meas. Sci. Technol.* **18**, 87–95 (2007).
21. S. R. Arridge, V. Kolehmainen, and M. J. Schweiger, "Reconstruction and regularisation in optical tomography," in "Interdisciplinary Workshop on Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)," (Pisa, Italy, 2007), pp. 1–18.
22. J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, and E. Somersalo, "Inverse problems with structural prior information," *Inv. Probl.* **15**, 713–729 (1999).
23. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," *Opt. Lett.* **26**, 893–895 (2001).
24. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Fluorescence lifetime imaging in turbid media," *Optics Letters* **21**, 158–160 (1996).
25. A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent tomography in the presence of heterogeneities: study of the normalized Born ratio," *IEEE Trans. Med. Imaging* **24**, 1377–1386 (2005).
26. T. J. Rudge, V. Y. Soloviev, and S. R. Arridge, "Fast image reconstruction in fluorescence optical tomography using data compression," *Opt. Lett.* **35**, 763–765 (2010).
27. F. Catteé, P. Lions, J. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Num. Anal.* **29**, 182–193 (1992).
28. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 629–639 (1990).
29. J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. Image Process.* **7**, 398–410 (1998).
30. M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Trans. Image Process.* **7**, 421–432 (1998).
31. M. Schweiger, S. R. Arridge, and I. Nissilä, "Gauss-Newton method for image reconstruction in diffuse optical tomography," *Phys. Med. Biol.* **50**, 2365–2386 (2005).
32. H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Mathematics and its applications (Kluwer Academic Publishers, 2000).
33. O. Scherzer and C. Groetsch, "Inverse scale space theory for inverse problems," in "Scale-Space and Morphology in Computer Vision," , vol. 2106 of *Lecture Notes in Computer Science*, M. Kerckhove, ed. (Springer, 2006), pp. 317–325.
34. B. Kaltenbacher, "Some Newton-type methods for the regularization of nonlinear ill-posed problems," *Inv. Probl.* **13**, 729 (1997).

35. M. Hanke, "A regularizing Levenberg - Marquardt scheme, with applications to inverse groundwater filtration problems," *Inv. Probl.* **13**, 79 (1997).
  36. P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.* **14**, 1487–1503 (1993).
  37. T. Correia, A. Gibson, M. Schweiger, and J. Hebden, "Selection of regularization parameter for optical topography," *J. Biomed. Opt.* **14**, 034044 (2009).
  38. Q. Fang and D. Boas, "Tetrahedral mesh generation from volumetric binary and gray-scale images," in "Proceedings of IEEE International Symposium on Biomedical Imaging," (2009), pp. 1142–1145.
  39. B. Dogdas, D. Stout, A. F. Chatziioannou, and R. M. Leahy, "Digimouse: a 3D whole body mouse atlas from CT and cryosection data," *Phys. Med. Biol.* **52**, 577 (2007).
  40. G. Alexandrakis, F. R. Rannou, and A. F. Chatziioannou, "Tomographic bioluminescence imaging by use of a combined optical-PET (OPET) system: a computer simulation feasibility study," *Phys. Med. Biol.* **50**, 4225 (2005).
  41. J. Aguirre, A. Sisniega, J. Ripoll, M. Desco, and J. J. Vaquero, "Design and development of a co-planar fluorescence and X-ray tomograph," in "Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE," (2008), pp. 5412–5413.
  42. D. Boas, "Diffuse photon probes of structural and dynamical properties of turbid media: theory and biomedical applications," Ph.D. thesis, University of Pennsylvania, Philadelphia (USA) (1996).
- 

## 1. Introduction

Fluorescence diffuse optical tomography (fDOT) is a relatively new optical imaging modality that uses fluorescent markers that accumulate in specific regions, to monitor cellular and subcellular functional activity. Although it is a technique mostly used for small animal research [1–5], it is a promising technique with several medical applications such as detection, diagnosis and monitoring of human neoplasms, in particular breast tumours [6–8].

In fDOT, a near-infrared (NIR) light source at excitation wavelength is projected onto the subject under study at different positions. The excitation light propagates diffusely into the tissue and some of the photons are absorbed by fluorochromes, which re-emit part of the energy at a longer wavelength. Then, fluorescence, and possibly, excitation light intensities are measured by detectors placed around the object, for example, opposite the source for transmission geometry. This measurement process is repeated for several source positions to obtain a data set that is used for the reconstruction of the fluorochrome concentration. In modern fDOT systems, detection is performed using a charged-coupled device (CCD) camera, not in direct contact with the subject. This leads to large data sets, making the problem largescale and (nominally) overdetermined. However, due to the diffusive nature of light propagation in biological tissue the image reconstruction is an ill-posed inverse problem [9], hence, in order to obtain a stable and meaningful solution one has to use regularisation methods.

Prior assumptions on the solution, for example its smoothness, can be incorporated in the image reconstruction [10, 11]. Another possibility is to include structural prior information, which can be obtained from other imaging methods such as X-ray computed tomography (XCT) [12] and Magnetic Resonance Imaging (MRI) [13]. The most common regularisation methods are the Tikhonov-type methods, which are L<sub>2</sub>-norm based and filter high-frequency noise and produce smooth images. Tikhonov regularisation with a Laplacian-type operator has been used in a few fDOT studies. Davis *et al* [13] used a Laplacian-type prior that incorporated structural information derived from MRI images. Their phantom and simulation studies showed significant improvements in the reconstruction of the fluorescence distribution when the structural prior was used. Lin *et al* [14] observed that, when both functional and structural Laplacian-type prior information is incorporated in the reconstruction, the recovered fluorochrome concentration is more accurate. Ale *et al*. [12] used different Laplacian-type matrices in the regularisation functional, and tested the reconstruction method on simulated, phantom and mouse data. Results showed that reconstructions were more accurate when their method was used compared to the standard Tikhonov method.

If the forward model is linearised, by using for example a Born approximation, and the regularisation term is L2-norm based, then the reconstruction of the fluorochrome concentration is reduced to a linear inverse problem. Nevertheless, in order to preserve edge information in the reconstructed image, non-linear regularisation methods, such as total variation (TV) [15], can be advantageous. Total variation (TV) regularisation is a L1-norm based method that not only penalises highly oscillating solutions, but also preserves edge information in the reconstructed image. Image reconstruction using non-linear priors is computationally expensive, and convergence can be slow and not always ensured. Freiberger *et al.* [16] analysed reconstruction schemes for fDOT based on the non-linear formulation of the forward model and non-linear priors, namely, total variation and level-set type priors. They found, through simulations performed using a point source-detection measurement geometry, that the reconstructions performed by solving the non-linear problem using a Newton-type method were more accurate than the reconstructions obtained using linearised forward models and linear regularisation methods. In a previous work, Freiberger *et al.* [17] introduced an alternating direction minimisation method to solve non-linear regularisation methods. This method splits the reconstruction problem into a Gauss-Newton step and a TV minimisation step.

The disadvantage of TV based methods is that in the presence of noise they can lead to the staircase effect, i.e., smooth regions appear as piecewise constant regions [18]. Anisotropic diffusion regularisation with anatomical edge prior has been used in diffusion optical tomography (DOT) [19–21] and electrical impedance tomography (EIT) [22] image reconstruction of simulated data. This non-linear regularisation technique smooths image noise whilst preserving edges, and does not yield staircase effects.

This paper proposes a split operator method to reconstruct fluorescence images fast and efficiently, using anisotropic diffusion regularisation and anatomical prior information obtained from another medical imaging modality. The image reconstruction algorithm splits the non-linear problem into a linear problem and a nonlinear anisotropic diffusion problem, which are straightforward to implement and solve. The performance of the algorithm is assessed using simulations, experimental phantom and mouse data. Prior anatomical information is obtained from XCT images, which were acquired sequentially with fDOT. Furthermore, the influence of using different edge-preserving functions in the anisotropic diffusion prior term is analysed. The choice of an edge-preserving function is an important topic, since it defines in which regions (homogeneous regions/discontinuities) diffusion is isotropic/anisotropic.

## 2. Methods

### 2.1. Formulation of the problem

In fDOT the continuous-wave forward model is described by a set of coupled diffusion equations at the excitation and emission wavelength,  $\lambda_e$  and  $\lambda_f$ , respectively:

$$[-\nabla \cdot \kappa(\mathbf{r}, \lambda_e) \nabla + \mu_a(\mathbf{r}, \lambda_e)] U(\mathbf{r}, \lambda_e) = q(\mathbf{r}, \lambda_e), \quad \mathbf{r} \in \Omega \quad (1)$$

$$[-\nabla \cdot \kappa(\mathbf{r}, \lambda_f) \nabla + \mu_a(\mathbf{r}, \lambda_f)] U(\mathbf{r}, \lambda_f) = U(\mathbf{r}, \lambda_e) h(\mathbf{r}, \lambda_f), \quad \mathbf{r} \in \Omega \quad (2)$$

where  $\kappa(\mathbf{r}, \lambda) = [3(\mu'_s(\mathbf{r}, \lambda) + \mu_a(\mathbf{r}, \lambda))]^{-1}$  is the diffusion coefficient at position  $r$  and at wavelength  $\lambda$ ,  $\mu_a$  and  $\mu'_s$  are the absorption and reduced scattering coefficients, respectively. The fluorophore is excited with light at  $\lambda_e$  emitted by a source  $q(\mathbf{r}, \lambda_e)$ . The emission photon density  $U(\mathbf{r}, \lambda_e)$  over the domain  $\Omega$  is obtained by solving Eq. (1). The fluorescence yield coefficient  $h$  is related to the quantum yield of the fluorophore  $\eta$  and its absorption coefficient at  $\lambda_f$ . The excitation photon density  $U(\mathbf{r}, \lambda_f)$  is calculated from Eq. (2). The Robin boundary condition is

applied to both equations:

$$U(\mathbf{r}, \lambda) + 2\zeta \kappa(\mathbf{r}, \lambda) \frac{\partial U(\mathbf{r}, \lambda)}{\partial n} = 0, \quad \mathbf{r} \in \partial\Omega \quad (3)$$

where  $\zeta$  is a boundary term that incorporates the refractive index mismatch and  $n$  is the outward normal at position  $\mathbf{r}$  on the boundary  $\partial\Omega$ . For weak fluorochrome concentrations the contribution of the fluorochrome to the optical parameters can be neglected, which yields the first-order Born approximation [23, 24]. Therefore, the measured fluorescence photon density  $\mathbf{y}_f$  and excitation photon density  $\mathbf{y}_e$ , due to a source at  $\mathbf{r}_s$  and a detector at  $\mathbf{r}_d$ , can be written as:

$$\begin{aligned} U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e) &= \int_{\Omega} q(\mathbf{r}, \lambda_e) G(\mathbf{r}_s, \mathbf{r}_d, \lambda_e) d\mathbf{r}, \\ \mathbf{y}_e &= \Theta_e(\mathbf{r}_s, \mathbf{r}_d) U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e), \end{aligned} \quad (4)$$

$$\begin{aligned} U(\mathbf{r}_s, \mathbf{r}_d, \lambda_f) &= \int_{\Omega} h(\mathbf{r}, \lambda_f) U(\mathbf{r}_s, \mathbf{r}, \lambda_e) G(\mathbf{r}, \mathbf{r}_d, \lambda_f) d\mathbf{r}, \\ \mathbf{y}_f &= \Theta_f(\mathbf{r}_s, \mathbf{r}_d, \lambda_f) U(\mathbf{r}_s, \mathbf{r}_d, \lambda_f), \end{aligned} \quad (5)$$

where  $G$  represents the Green's function solution and  $\Theta$  are the unknown source and detector coupling coefficients. Normalising the measured fluorescence photon density  $\mathbf{y}_f$  by the measured excitation photon density  $\mathbf{y}_e$  reduces the effects of  $\Theta$  [23, 25]:

$$\frac{\mathbf{y}_f}{\mathbf{y}_e} = \frac{U(\mathbf{r}_s, \mathbf{r}_d, \lambda_f)}{U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e)} \quad (6)$$

$$= \frac{1}{U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e)} \int_{\Omega} h(\mathbf{r}, \lambda_f) U(\mathbf{r}_s, \mathbf{r}, \lambda_e) G(\mathbf{r}, \mathbf{r}_d, \lambda_f) d\mathbf{r}, \quad (7)$$

where  $\Theta_f = \Theta_e$  is assumed. Note that in CCD based measurement systems, such as used in this paper, the previous equation is modified to:

$$\frac{\mathbf{y}_f}{\mathbf{y}_e} = \frac{P(\partial\Omega \rightarrow \Sigma) U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e)}{P(\partial\Omega \rightarrow \Sigma) U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e)}, \quad (8)$$

where the operator  $P$  represents the projection from the domain boundary  $\partial\Omega$  to the camera  $\Sigma$ .

The forward problem defined by the set of diffusion Eqs. (1) and (2) is computed numerically using the finite element method (FEM) on a tetrahedral mesh based on the geometry being considered. However, when using an image based prior, such as the anisotropic diffusion prior, it is convenient to map the solution into a regular grid. Therefore, Eq. (6) is discretised into  $N$  volume elements of size  $v$ :

$$\frac{\mathbf{y}_f}{\mathbf{y}_e} = \frac{1}{U(\mathbf{r}_s, \mathbf{r}_d, \lambda_e)} \sum_{i=1}^N h(\mathbf{r}_i, \lambda_f) U(\mathbf{r}_s, \mathbf{r}_i, \lambda_e) G(\mathbf{r}_i, \mathbf{r}_d, \lambda_f) v. \quad (9)$$

For a set of  $M_s$  source-camera projections, where the detector is a camera with  $M_x \times M_y$  pixels, the fDOT problem can be written as a linear system:

$$\hat{\mathbf{y}} = \frac{\mathbf{y}_f}{\mathbf{y}_e} = \mathbf{J}\mathbf{h}, \quad (10)$$

where here  $\mathbf{y}_f$  and  $\mathbf{y}_e$  are vectors of size  $M_s \times M_x \times M_y$ ,  $\mathbf{J}$  is the Jacobian matrix of dimensions  $(M_s \times M_x \times M_y) \times N$ , where each row is obtained from Eq. (9) for each source-detector pair, and  $\mathbf{h}$  is a vector of size  $N$ .

The estimation of the fluorescence yield  $\mathbf{h}$  from Eq. (10) is often computationally infeasible due to the large dimensions of the Jacobian matrix. Therefore, data compression is used to reduce the dimensions of the data for computational efficiency [26], while simultaneously reducing the redundancy of the data. The 2D wavelet transform of a projection image  $\mathbf{y}_i$  captured by a CCD camera can be expanded in the wavelet basis as:

$$\mathbf{y}_i = \sum_k^{M_x \times M_y} c_{ik} \varphi_{ik} \quad (11)$$

where  $c$  are the wavelet coefficients and  $\varphi$  the wavelet basis. If only the largest  $M_\varphi$  wavelet coefficients are kept, which contains most of the relevant information, then the dimensions of the vectorised compressed data  $\tilde{\mathbf{y}}$  are reduced to  $M_s \times M_\varphi$ . Similarly, the size of the compressed Jacobian  $\tilde{J}$  is reduced to  $(M_s \times M_\varphi) \times N$ . The resulting compressed forward problem is:

$$\tilde{\mathbf{y}} = \tilde{J}\mathbf{h}. \quad (12)$$

## 2.2. Inverse problem methods in image denoising

Many regularisation schemes in inverse problems originate from image denoising applications. One general approach in this application is to convolve an image with a Gaussian kernel, which smooths out the noise and edges, and then an anisotropic diffusion filter is applied, which recovers the image edges and smooths homogeneous regions in the image [27].

### 2.2.1. Anisotropic diffusion

The concept of anisotropic diffusion was introduced in image processing by Perona-Malik [28]. It is an efficient technique that preserves and enhances edges present in the images while simultaneously removing image noise. Following is a brief analysis of the principle behind this method.

In the image space  $h \in \mathbb{R}^3$  the anisotropic diffusion process is given by:

$$\frac{\partial h}{\partial t} = \nabla \cdot [g(|\nabla h|) \nabla h], \quad (13)$$

where  $|\nabla h| = \sqrt{(\frac{\partial h}{\partial x})^2 + (\frac{\partial h}{\partial y})^2 + (\frac{\partial h}{\partial z})^2}$ ,  $g(|\nabla h|) = \frac{\psi'(|\nabla h|)}{|\nabla h|}$  is the diffusivity, and  $\psi$  is a potential function that controls the diffusion process. This term can be written in terms of the gradient direction and tangent directions. The unit vector in the direction of the gradient is  $\eta = \frac{\nabla h}{|\nabla h|}$  and the two unit vectors in the tangent direction, i.e., orthogonal to the gradient, are  $\xi_1$  and  $\xi_2$ . Using the divergence product rule of a scalar function  $f$  and vector  $\mathbf{v}$ :  $\nabla \cdot (f\mathbf{v}) = f(\nabla \cdot \mathbf{v}) + \nabla f \cdot \mathbf{v}$ , the anisotropic diffusion operator can be written as:

$$\nabla \cdot \left[ \frac{\psi'(|\nabla h|)}{|\nabla h|} \nabla h \right] = \psi'(|\nabla h|) \nabla \cdot \left( \frac{\nabla h}{|\nabla h|} \right) + \nabla (\psi'(|\nabla h|)) \cdot \frac{\nabla h}{|\nabla h|}. \quad (14)$$

After some manipulation we have the directional anisotropic diffusion:

$$\nabla \cdot \left[ \frac{\psi'(|\nabla h|)}{|\nabla h|} \nabla h \right] = \frac{\psi'(|\nabla h|)}{|\nabla h|} (h_{\xi_1 \xi_1} + h_{\xi_2 \xi_2}) + \psi''(|\nabla h|) h_{\eta \eta}, \quad (15)$$

where  $h_{\xi_1 \xi_1}$ ,  $h_{\xi_2 \xi_2}$  and  $h_{\eta \eta}$  are the second derivatives of  $h$  in the  $\xi_1$ ,  $\xi_2$  and  $\eta$  directions, respectively. This formulation can help us understand the effect of the anisotropic diffusion regularisation in the normal and tangential directions. In order to preserve edges the regularisation should locally penalise magnitudes of the directional derivative across an edge much less

than along an edge. Thus, the function  $\psi$  should be constructed such that the diffusion along the normal direction  $\eta$  is small compared to the tangential directions  $\xi_1$  and  $\xi_2$ .

The isotropic diffusion (Tikhonov) regularisation, i.e., when  $\psi(|\nabla h|) = \frac{1}{2}|\nabla h|^2$ , does not preserve edges since it acts in all directions:

$$\nabla \cdot \left[ \frac{\psi'(|\nabla h|)}{|\nabla h|} \nabla h \right] = h_{\xi_1 \xi_1} + h_{\xi_2 \xi_2} + h_{\eta \eta}. \quad (16)$$

### 2.2.2. Discretisation

The discrete formulation of the anisotropic diffusion in Eq. (13) using an explicit scheme, at the  $k$  th iteration, is given by [29]:

$$\frac{h_i^{k+1} - h_i^k}{\Delta t} = \sum_{j \in \mathfrak{N}(i)} \frac{g_j^k + g_i^k}{2d^2} (h_j^k - h_i^k), \quad (17)$$

where  $g_i$  is the diffusivity of pixel  $i$ ,  $\mathfrak{N}(i)$  are the neighbours of pixel  $i$ , and  $\Delta t$  is the time step and  $d$  is the discretisation interval. For  $m$ -dimensions, the previous equation can be written in a matrix-vector notation as [29]:

$$\frac{h_i^{k+1} - h_i^k}{\Delta t} = \sum_{l=1}^m \mathcal{L}_l(h^k) h^k \quad (18)$$

$$h^{k+1} = \left( I + \Delta t \sum_{l=1}^m \mathcal{L}_l(h^k) \right) h^k, \quad (19)$$

where the matrix  $\mathcal{L}_l(h^k) = [\ell_{ij}(h^k)]$  corresponds to the diffusivity derivatives along the  $l$ th coordinate axis:

$$\ell_{ij}(h^k) = \begin{cases} \frac{g_j^k + g_i^k}{2dx^2} & j \in \mathfrak{N}(i) \\ -\sum_{j \in \mathfrak{N}(i)} \frac{g_j^k + g_i^k}{2dx^2} & j = i \\ 0 & \text{else} \end{cases}$$

Note that only two neighbours of pixel  $i$  are considered for each direction. The drawback of the explicit scheme is that it is only conditionally stable. This means that the step size needs to be small for the solution to be stable, and therefore, a large number of iterations are required. For example, in 2D  $\Delta t < 1/4$ . The allowed step size becomes smaller for higher dimensions.

Alternatively, we use the semi-implicit additive operator splitting (AOS) method, which assumes the following expression [29]:

$$h^{k+1} = \frac{1}{m} \sum_{l=1}^m \left( I - m\Delta t \mathcal{L}_l(h^k) \right)^{-1} h^k, \quad (20)$$

This scheme is unconditionally stable. Thus,  $\Delta t$  can be relatively large without causing numerical instability. Since it is an additive method one can treat each direction separately. Therefore, the AOS scheme involves solving  $m$  linear systems  $Dh^{k+1} = h^k$ , where the matrix  $D = (I - \Delta t m \mathcal{L}_l(h^k))$  is tridiagonal. Note that for each direction there is one tridiagonal matrix. An efficient way of solving this is to use the Thomas algorithm, which is a Gaussian elimination algorithm for tridiagonal systems [29]. This method avoids the need of directly determining the inverse of large matrices, since the problem is reduced to computing a series of multiplications, divisions, additions and subtractions.

Table 1. Edge-Preserving Functions

Perona- Malik (Cauchy or Lorentzian) [30]	$\psi( \nabla h ) = \frac{T^2}{2} \log \left( 1 + \left( \frac{ \nabla h }{T} \right)^2 \right)$ $g( \nabla h ) = \frac{1}{1 + \left( \frac{ \nabla h }{T} \right)^2}$
Perona- Malik 2 (Welsh) [30]	$\psi( \nabla h ) = \frac{T^2}{2} \left[ 1 - \exp \left( - \left( \frac{ \nabla h }{T} \right)^2 \right) \right]$ $g( \nabla h ) = \exp \left( - \left( \frac{ \nabla h }{T} \right)^2 \right)$
Total variation approximation [21]	$\psi( \nabla h ) = T \sqrt{ \nabla h ^2 + T^2} - T$ $g( \nabla h ) = \frac{T}{\sqrt{ \nabla h ^2 + T^2}}$
Huber [21, 30]	$\psi( \nabla h ) = \begin{cases} T \nabla h  - \frac{T^2}{2} &  \nabla h  > T \\ \frac{ \nabla h ^2}{2} &  \nabla h  \leq T \end{cases}$ $g( \nabla h ) = \begin{cases} \frac{T}{\sqrt{ \nabla h ^2 + T^2}} &  \nabla h  > T \\ 1 &  \nabla h  \leq T \end{cases}$
Tukey [30]	$\psi( \nabla h ) = \begin{cases} \frac{T^2}{6} \left[ 1 - \left[ 1 - \left( \frac{ \nabla h }{T} \right)^2 \right]^3 \right] &  \nabla h  < T \\ \frac{T^2}{6} &  \nabla h  \geq T \end{cases}$ $g( \nabla h ) = \begin{cases} \left[ 1 - \left( \frac{ \nabla h }{T} \right)^2 \right]^2 &  \nabla h  < T \\ 0 &  \nabla h  \geq T \end{cases}$
Exceedance	$\psi( \nabla h ) = \frac{ \nabla h ^2}{2} (1 - P( \nabla h  \leq X))$ $g( \nabla h ) = 1 - P( \nabla h  \leq X) = P( \nabla h  > X)$

### 2.2.3. Edge-preserving functions

As mentioned previously, the diffusion regularisation should be isotropic in homogeneous regions and in the presence of an edge the diffusion regularisation should be significantly smaller in the normal direction than in the tangential direction. Table 1 shows a few functions with these edge preserving properties [21, 30]. In the isotropic diffusion case we have the first-order Tikhonov regularisation, where  $\psi(|\nabla h|) = \frac{1}{2}|\nabla h|^2$  and  $g(|\nabla h|) = 1$ .

The parameter  $T$  is the threshold. If this parameter is too large it leads to an oversmoothing of the image. Whereas, if  $T$  is too small, smoothing is not applied and the resulting image is similar to the initial one. The parameter  $T$  can be selected using a method based on normalised cumulative histogram (NCH) of the gradient [28]. This method is commonly used in edge detection problems. The NCH indicates the probability  $P$  of a gradient taking on a value less than or equal to the value  $X$  that the bin represents. It increases monotonically and the smoothness/sharpness of the curve indicates how smooth/sharp the edges are. The threshold can be



calculated from the NCH by setting the threshold at, for example, 90 percent. It can be kept fixed or it could be updated at every iteration. Note that for each edge-preserving function the parameter  $T$  is scaled, so that all functions preserve edges at the same threshold.

The exceedance function can be thought of as the probability of an edge of interest being present and it can be calculated using the NCH of the image gradient. The NCH gives the probability  $P(|\nabla h| \leq X)$ . Consider  $X = T$ , then  $P(|\nabla h| \leq X) \simeq 0$  and  $g \simeq 1$ , on the other hand NCH tends to 1 as the bin values increase, and therefore,  $g$  tends to 0. The exceedance function has a similar behaviour than the other edge-preserving functions, but does not require a direct estimation of  $T$ .

### 2.3. fDOT inversion using anisotropic diffusion regularisation

We now define the inverse problem of fDOT as minimising the following objective function:

$$\text{Minimise } E(\mathbf{h}) = \frac{1}{2} \|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}\|^2 + \alpha \Psi(\mathbf{h}), \quad (21)$$

where  $\alpha$  is the regularisation parameter, and  $\Psi(\mathbf{h})$  is the prior functional given by:

$$\Psi(\mathbf{h}) = \int_{\Omega} \psi(|\nabla \mathbf{h}|) d\Omega. \quad (22)$$

The minimum of  $E(\mathbf{h})$  can be calculated using the gradient:

$$\nabla E(\mathbf{h}) = \tilde{\mathbf{J}}^T (\tilde{\mathbf{J}}\mathbf{h} - \tilde{\mathbf{y}}) + \alpha \mathcal{L}(\mathbf{h})\mathbf{h}, \quad (23)$$

where  $\mathcal{L}$  is a linear operator given by:

$$\mathcal{L}(\mathbf{h}) = -\nabla \cdot [g(|\nabla \mathbf{h}|) \nabla]. \quad (24)$$

#### 2.3.1. Incorporation of an anatomical prior

If prior information about the image structure is available, it can be used in the prior term to weight the diffusion function. Consider  $x_{ref}$  to be an anatomical image, which has been filtered previously by a Gaussian or anisotropic diffusion filter. Then, in order to obtain the image weighting factor  $W(|\nabla x_{ref}|)$  one can apply one of the diffusion functions in Table 1 to the anatomical image, which will return an image where homogeneous regions have a large weight and close to edges the weight approaches zero. The prior term takes the form:

$$\Psi_W(\mathbf{h}) = \int_{\Omega} W(|\nabla x_{ref}|) \psi(|\nabla \mathbf{h}|) d\Omega, \quad (25)$$

which leads to

$$\mathcal{L}_W(\mathbf{h}) = \nabla \cdot [W(|\nabla x_{ref}|) g(|\nabla \mathbf{h}|) \nabla]. \quad (26)$$

Instead of keeping the threshold parameter  $T$  (see §2.2.3) constant over the entire image we use a spatially variant  $T$ , which is updated at each iteration. The normalised image gradient  $\nabla h_n^k$  and normalised structural weighting factor  $W_n(|x_{ref}|)$  can be used to weight the threshold value found from the NCH. Note that both quantities are normalised so that they are equal to 1 in flat regions and tend to 0 near the edges. The spatially variant threshold is calculated as:  $T(x, y, z) = T + T [g_n(x, y, z) - 1/2]$ , where  $T$  is the threshold found using the NCH and  $g_n(x, y, z) = (\nabla h_n^k + W_n(|x_{ref}|)) / 2$ . Therefore, in the presence on an edge in either the optical image or the anatomical image, the parameter  $T$  is small and the edges are preserved, since the diffusion process is stopped. In homogeneous regions  $T$  is large, which results in a stronger smoothing effect. The variant  $T$  is particularly useful for restoring features in the image with different contrasts.

## 2.4. Image reconstruction algorithm

The Gauss-Newton approach can be used to obtain the solution iteratively [31]:

$$\mathbf{h}^{k+1} = \mathbf{h}^k + \tau \left( \tilde{J}^T \tilde{J} + \alpha \Psi''(\mathbf{h}^k) \right)^{-1} \tilde{J}^T \left( (\tilde{\mathbf{y}} - \tilde{J}\mathbf{h}^k) - \alpha \Psi'(\mathbf{h}^k) \right), \quad (27)$$

where  $\tau$  is the step length. In the lagged-diffusivity Gauss-Newton method the Hessian of the prior is calculated from a linear diffusion operator:  $\Psi''(\mathbf{h}) = \mathcal{L}(\mathbf{h})$ . However, the Gauss-Newton method Eq. (27) requires a computationally expensive matrix inversion. In order to overcome this issue and aiming at having fast runtime performance, a split operator algorithm is introduced in this paper.

We want to find an update step for Eq. (21). We note that  $\mathcal{L}(\mathbf{h})$  is a symmetric positive definite regularisation matrix, and we assume both a decomposition and inverse to exist such that:

$$L^T L = \mathcal{L}(\mathbf{h}), \quad \Gamma = [\mathcal{L}(\mathbf{h})]^{-1} = L^{-1} (L^{-1})^T. \quad (28)$$

Since  $\mathcal{L}(\mathbf{h})$  is a differential operator, then  $\Gamma$  is a smoothing operator. Now consider the transformation:

$$\tilde{J} = \tilde{J}L^{-1}, \quad \check{\mathbf{h}} = L\mathbf{h}, \quad (29)$$

such that Eq. (21) becomes:

$$\text{Minimise } E(\check{\mathbf{h}}) = \frac{1}{2} \|\tilde{\mathbf{y}} - \tilde{J}\check{\mathbf{h}}\|^2 + \alpha \|\check{\mathbf{h}}\|^2. \quad (30)$$

An iterative descent method (the Landweber method) [32] for the previous regularised least squares problem (Eq. (30)) is:

$$\check{\mathbf{h}}^{k+1} = \check{\mathbf{h}}^k - \tilde{J}^T \left( \tilde{\mathbf{y}} + \tilde{J}\check{\mathbf{h}}^k \right) \quad (31)$$

$$\Rightarrow \mathbf{h}^{k+1} = \mathbf{h}^k + \Gamma \tilde{J}^T \left( \tilde{\mathbf{y}} - \tilde{J}\mathbf{h}^k \right). \quad (32)$$

We can interpret Eq. (31) as a gradient descent step for the likelihood (first term is Eq. (21)), followed by a smoothing step to regularise the update:

$$\mathbf{h}^{k+1/2} = \mathbf{h}^k + \tilde{J}^T \left( \tilde{\mathbf{y}} - \tilde{J}\mathbf{h}^k \right) \quad (33)$$

$$\mathcal{L}(\mathbf{h}^{k+1})\mathbf{h}^{k+1} = \mathbf{h}^{k+1/2} \quad (34)$$

In order to achieve a faster convergence we may use a Levenberg-Marquardt (LM) step instead of a descent step:

$$\mathbf{h}^{k+1/2} = \mathbf{h}^k + \tilde{J}^T \mathbf{B} \left( \tilde{\mathbf{y}} - \tilde{J}\mathbf{h}^k \right), \quad (35)$$

where  $\mathbf{B}$  is a symmetric positive definite matrix. Here we define:

$$\mathbf{B} = \left( \tilde{J}\tilde{J}^T + \lambda I \right)^{-1}, \quad (36)$$

which is cheap to invert because we compressed the row space of the Jacobian using wavelets. The parameter  $\lambda$  in the LM method is the damping factor, which changes at each iteration

Full inversion of the smoothing step is expensive, so we replace the regularisation step by the semi-implicit AOS method (Eq. (20)), which can be solved efficiently using the Thomas algorithm. We may now relax the requirement that an inverse of  $\mathcal{L}(\mathbf{h})$  exists, and we do not need to consider the decomposition used in Eq. (28).

A simpler way to infer this split operator method is by considering Eq. (23) as the update  $\Delta \mathbf{h}^k$  of the gradient descent method. Since  $\Delta \mathbf{h}^k$  consists of two terms it can be split into a two-step iteration.

This reconstructions method involves solving two systems of equations, where the first step is simply the LM method [32] and the second step is the semi-implicit form of the anisotropic diffusion method, where  $\Delta t$  is the time step, which plays the role of  $\alpha$  in Eq. (21) [33]. Therefore, the first step reconstructs the fluorescence distribution image and the second step applies the prior, by deblurring/denoising the reconstructed image.

The parameter  $\lambda$  in the LM method serves a role similar to the step size  $\tau$ , since it controls the magnitude and direction of the iteration update  $\Delta \mathbf{h}^k$ . The LM method switches between the Gauss-Newton method and the gradient descent method. When  $\lambda \rightarrow 0$  the method is basically the same as Gauss-Newton. On the other hand when  $\lambda \rightarrow \infty$  the method becomes a gradient descent approach. Since the LM method is a special case of Tikhonov regularisation [34, 35] and, furthermore,  $\tilde{\mathbf{h}}^0 = \{0, \dots, 0\}$ , then  $\lambda^0$  can be calculated by the L-curve method  $\lambda_{lc}$  [36, 37]. For the following iterations, if the least squares norm  $\|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^{k+1/2}\|^2 < \|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^k\|^2$  then  $\lambda^{k+1}$  is decreased by a certain percentage of the current value,  $p_\lambda$ , since the Gauss-Newton method converges quickly near a minimum. If  $\|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^{k+1/2}\|^2 > \|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^k\|^2$  then  $\lambda^{k+1}$  is greater than  $\lambda^k$ , since the gradient descent method is more efficient when the procedure is far from the minimum.

The image reconstruction is summarised in the following algorithm:

---

**Algorithm 1** Two-Step Image Reconstruction Algorithm

---

```

 $\alpha^0 = \lambda_{lc} \text{trace}(JJ^T)$ 
 $u = 0$ 
 $\varepsilon = \text{tolerance value}$ 

while  $\|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^k\|^2 < \varepsilon \vee k < N_k$  do
     $\mathbf{h}^{k+1/2} = \mathbf{h}^k + \tilde{\mathbf{J}}^T (\tilde{\mathbf{J}}\tilde{\mathbf{J}}^T + \lambda I)^{-1} (\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^k)$ 
AOS-Thomas algorithm

    for  $i = 1, \dots, n$  do
        for  $l = 1, \dots, m$  do
             $v = (I + m\Delta t \mathcal{L}_l(\mathbf{h}^{k+1/2}))^{-1} \mathbf{h}^{k+1/2}$ 
             $u = u + v$ 
        end for
         $\mathbf{h}^{k+1/2} = \frac{u}{m}$ 
    end for
     $\mathbf{h}^{k+1} = \mathbf{h}^{k+1/2}$ 

    if  $\|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^{k+1/2}\|^2 < \|\tilde{\mathbf{y}} - \tilde{\mathbf{J}}\mathbf{h}^k\|^2$  then
         $\lambda^{k+1} = \lambda^k - \lambda^k * p_\lambda$ 
    else
         $\lambda^{k+1} = \lambda^k + \lambda^k * p_\lambda$ 
    end if
end while

```

---

From our experience, for the prior step, it is preferable to perform a small number of iterations  $n$  and use a relatively large  $\Delta t$ , than just a single iteration and a very large  $\Delta t$ .

## 2.5. Evaluation

The performance of our image reconstruction algorithm is evaluated through simulations, experimental phantom measurements and *ex vivo* mouse measurements.

The image reconstruction algorithm implementation is based on the software package Time-resolved Optical Absorption and Scattering Tomography (TOAST) developed at University College London (UCL). TOAST is a finite element method (FEM) based software and therefore requires a suitable computational mesh. Meshes were generated from the XCT images of the phantom and mouse. These images were contaminated by noise and were smoothed using a simple Gaussian filter. The images were segmented by thresholding, and the segmented images were used to generate a tetrahedral mesh using the Matlab package Iso2mesh [38].

The simulation procedure and experimental setup are described below. A tolerance value  $\varepsilon = 1e-4$  and maximum iteration number  $N_k = 30$  were used as the stopping criteria in Algorithm 1.

### 2.5.1. Simulations

The Digimouse atlas [39] was used to generate a mouse mesh with 76 973 nodes, 450 829 elements and dimensions 32.5 mm  $\times$  19.5 mm  $\times$  88 mm. A fluorescent target was used to simulate a tumour in the liver. The target is a sphere of radius 1.75 mm and contrast  $h = 1 \text{ mm}^{-1}$  ( $h = 0 \text{ mm}^{-1}$  outside). We used a simplified two-tissue model, where different optical properties were assigned to the liver ( $\mu_a = 0.035 \text{ mm}^{-1}$  and  $\mu'_s = 0.68 \text{ mm}^{-1}$ ) and other tissue ( $\mu_a = 0.01 \text{ mm}^{-1}$  and  $\mu'_s = 0.8 \text{ mm}^{-1}$ ). The optical properties were calculated using the parameters provided by Alexandrakis *et al* [40] at a wavelength of 670 nm. The source and detector (CCD camera), which was placed opposite the source, were considered to rotate around the mouse. Projections were calculated for 16 evenly-spaced positions over the full 360° range. Data consisted of fluorescence and excitation projections with 1% additive Gaussian random noise. For the reconstruction a total of 128 wavelet coefficients were used.

The peak signal-to-noise ratio (PSNR) in decibels (dB) is used to evaluate the quality of the reconstructed images. The PSNR between the true image  $h_{true}$  of dimensions  $X \times Y \times Z$  and the reconstructed image  $h_{recon}$  is defined as follows:

$$PSNR = 10 \log_{10} \frac{\max(h_{true})}{MSE}, \quad (37)$$

where MSE is the mean squared error:

$$MSE = \frac{\sum_{i=1}^X \sum_{j=1}^Y \sum_{k=1}^Z (h_{true} - h_{recon})^2}{X \times Y \times Z}. \quad (38)$$

A higher PSNR indicates a greater fidelity to the original image.

### 2.5.2. Experimental setup

**2.5.2.1. fDOT-XCT system** The fDOT-XCT rotating system was developed and assembled at Universidad Carlos III de Madrid (Spain) [41]. The fDOT and CT components were mounted on a rotating gantry, whose rotation movement was controlled by a Newport RV350PP motorised rotation stage (Newport Co, Irvine, CA, USA). The object to be imaged was placed in the horizontal position on a bed specifically designed for the experiment purposes, which was mounted on a motorised linear translation stage (LTM80-300-HsM, OWIS GmbH, Staufen, DE) to accurately position the object in the field of view (FOV) of the system, and was gently compressed between two methacrylate transparent plates to produce planar surfaces.

Note that in our studies the gantry did not rotate during the experiment, and only the source position changes. The CCD camera (ORCA II, Hammamatsu, Japan) and objective lenses

(Nikon) was placed parallel to the anti-reflective plates. The laser diode (Monocrom, Barcelona, Spain) with an emission wavelength of  $675 \pm 5$  nm was placed opposite the CCD camera. The output power was controlled via pulse width modulation (PWM) using transistor – transistor logic (TTL) levels. The laser light was focused onto the sample at predefined points via two mirrors moved by galvanometers (Scancube 7, ScanLab AG, Puchheim, Germany). For every source position, the laser power that gives an optimal number of counts recorded by the CCD camera was obtained using an automated algorithm. Fluorescence measurements were recorded by placing a 10 nm bandwidth filter centred at 700 nm in front of the objective of the CCD camera, whereas for excitation measurements a 10 nm bandwidth filter centred at 675 nm was used. A motorised filter wheel (Luxiflux V2, Cyberstar, Echirolles, France) was used to quickly switch between filters. All the acquisition processes were controlled via a software written in C++ and IDL code.

The micro-CT consisted of a modified Hamamatsu L9631 microfocus X-ray source (Hamamatsu Photonics, Japan), in which the X-ray control unit was removed from its housing in order to reduce size and weight of the device attached to the rotating gantry, and a flat-panel X-ray detector (Hamamatsu C7940DK-02, Hamamatsu Photonics, Japan). The X-ray source has an output power of 50 W, a maximum peak energy of 110 keV and a maximum anode current of 0.8 mA. The detector pixel size was 0.05 mm and up to  $4 \times 4$  pixels could be combined through pixel binning, in order to speed image acquisition (at the expense of image resolution). The image integration time was 125 ms when  $4 \times 4$  binning was used.

The detector was mounted on a motorised linear stage (LTM80-100-HSM, OWIS GmbH, Staufen, DE), which allows the X-ray detectors to be moved radially to set the desired FOV size and magnification factor (that determines the resolution of the acquired images). Images were acquired with a fixed integration time of 125 ms, for a total of 360 projections over an angular range of  $360^\circ$ . The resulting images had a voxel size of  $(0.144 \text{ mm})^3$ .

**2.5.2.2. Phantom** A slab phantom made of a polyester resin was built with  $\mu_a = 0.01 \text{ mm}^{-1}$  and  $\mu'_s = 0.8 \text{ mm}^{-1}$  [42]. The optical properties of the phantom were considered to be the same at both measurement wavelengths. The phantom dimensions were  $50 \times 50 \times 10$  mm. A rectangular mesh with slighter larger dimensions than the phantom was generated ( $75 \times 75 \times 10$  mm), which contained 65 596 nodes and 303 750 elements. A capillary of 1 mm diameter was inserted into the phantom, close and parallel to the imaging surface, and its tip was filled with 2  $\mu\text{L}$  of Alexa Fluor 680 with fluorochrome concentration 20  $\mu\text{Molar}$  ( $\mu\text{M}$ ). The capillary was located approximately 1.8 mm below the imaging surface and at the centre ( $x = 25$  mm) of the phantom. The extremities of the capillary tube were at  $y = 0$  mm and  $y = 31$  mm. A total of 42 projections were obtained (the CCD camera, placed opposite the source, was static and only the source position varied). Sources were equally spaced within a square region of  $10 \times 10$  mm. Images were reconstructed with 128 wavelet coefficients.

**2.5.2.3. Mouse** Similar to the phantom experiment, a capillary with 2  $\mu\text{L}$  Alexa Fluor 680 with fluorochrome concentration 20  $\mu\text{M}$  was inserted into the esophagus of an euthanised adult nude mouse. A total of 56 equally spaced sources scanned a  $10 \times 10$  mm region of the upper torso of the animal. A finite element mesh was generated, from the XCT images, with 115 764 nodes and 608 319 elements. For simplicity, optical properties were assumed to be homogeneous ( $\mu_a = 0.01 \text{ mm}^{-1}$  and  $\mu'_s = 0.8 \text{ mm}^{-1}$ ) and the same for both excitation and emission wavelengths. A total of 128 wavelet coefficients was used in the data in the image reconstruction.

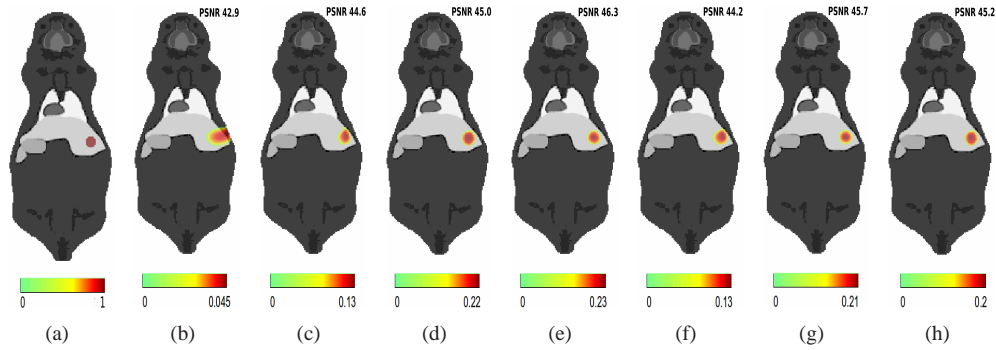


Fig. 1. Reconstructions of  $h$  ( $\text{mm}^{-1}$ ) from simulated data (image dimensions  $32.5 \text{ mm} \times 88 \text{ mm}$ ) and PSNR values (dB): (a) the target, (b) reconstruction with Tikhonov prior, (c) reconstruction with anatomical prior and total variation, (d) reconstruction with anatomical prior and Perona-Malik function, (e) reconstruction with anatomical prior and Perona-Malik 2 function, (f) reconstruction with anatomical prior and Huber function, (g) reconstruction with anatomical prior and Tukey function, and (h) reconstruction with anatomical prior and exceedance function.

### 3. Results

#### 3.1. Simulations

The simulated fluorescent inclusion in the mouse liver and reconstructed images of the fluorescence yield coefficient (in  $\text{mm}^{-1}$ ), overlaid on the anatomical atlas, are shown in Fig. 1. For comparison with the images reconstructed using the proposed method, Fig. 1(b) shows the image reconstructed using a simple Tikhonov regularisation. Figures 1(c)–1(h) show the reconstructed images using the anatomical prior and different edge-preserving functions, with spatially variant threshold  $T$ . The PSNR value for each reconstruction is also displayed. Images reconstructed using the Perona-Malik 2, Tukey and exceedance functions have the highest PSNR values (in decreasing order).

Figure 2 shows the weighted Perona-Malik edge prior at the first and last iteration of the reconstruction. The prior at the first iteration only shows the anatomical edges. As the number of iterations increase, the edges of the liver weight the influence of the edges in the reconstructed image. Therefore, diffusion is blocked across boundaries where the weighted prior has a small value, thus enhancing the edges of the reconstructed image, as well as smoothing the remaining regions where the weighted prior has a larger value.

Figure 3 shows the normalised profiles across the fluorescent target, for the longitudinal and lateral directions, for the images reconstructed using Tikhonov regularisation and our method with the edge-preserving function that returned the highest PSNR value.

#### 3.2. Phantom

The phantom consisted of a resin slab with a small capillary tube filled with a fluorescent dye, which was located close to the imaging surface. The XCT images were used to extract the capillary structure, i.e., the structural prior  $W_n(|x_{ref}|)$ . Figure 4(a) shows the fluorochrome concentration (in  $\mu\text{M}$ ) reconstructed using Tikhonov regularisation. The reconstructed concentrations using the structural prior and the different potential functions are shown in Figs. 4(b)–4(e). Figure 4(b) shows the reconstruction with total variation, Fig. 4(c) shows the reconstruction with Perona-Malik function, Fig. 4(d) shows the reconstruction with Perona-Malik 2 function and

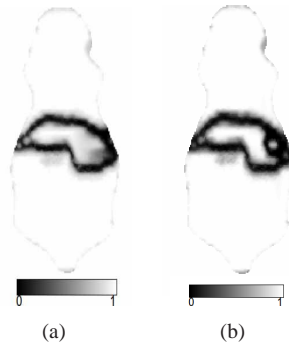


Fig. 2. Weighted Perona-Malik prior (a) at the first iteration, where only the anatomical edges are visible, and (b) at the last iteration, where both anatomical and fluorescence target edges can be seen.

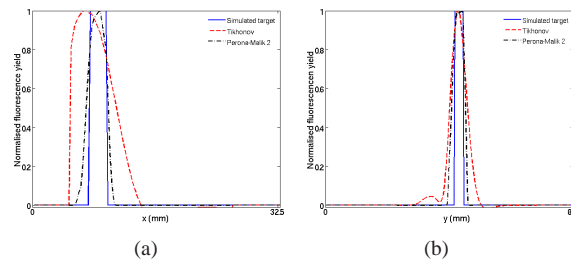


Fig. 3. Normalised profile plots across the fluorescent target: (a) lateral direction and (b) longitudinal direction.

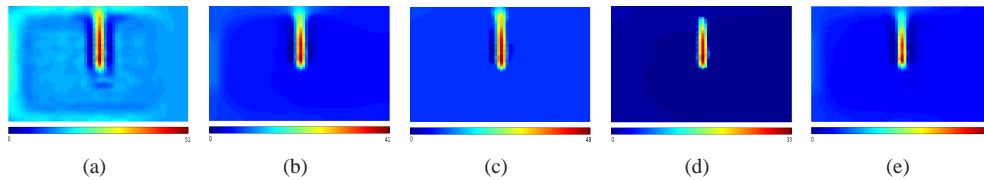


Fig. 4. Fluorochrome concentrations ( $\mu\text{M}$ ) reconstructed from phantom data using : (a) Tikhonov prior, (b) structural prior and total variation, (c) structural prior and Perona-Malik function, (d) structural prior and Perona-Malik 2 function, and (e) structural prior and exceedance function. Images have dimensions  $150 \text{ mm} \times 150 \text{ mm}$ .

Fig. 4(e) shows the reconstruction with the exceedance function. Similar reconstruction were obtained using the Huber and Tukey methods, and therefore are not shown.

### 3.3. Mouse

The XCT images of the mouse with a capillary inserted in the esophagus were used to obtain the anatomical prior, where the main visible structures were the bones and the capillary. Figure 5 shows the reconstructed images of the fluorochrome in the capillary, overlaid on the XCT image. Figure 5(a) shows the recovered fluorochrome concentration (in  $\mu\text{M}$ ) obtained using the Tikhonov method and Fig. 5(b) shows the image reconstructed using the Perona-Malik 2 method weighted by the anatomical prior.

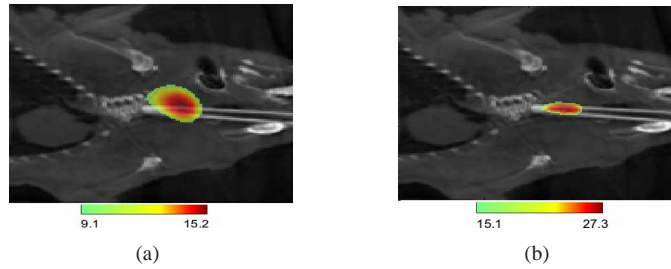


Fig. 5. Fluorochrome concentrations ( $\mu\text{M}$ ) reconstructed from mouse data using : (a) Tikhonov prior, (b) structural prior and Perona-Malik 2. Images have dimensions  $32 \text{ mm} \times 32 \text{ mm}$ .

The computational time of a single iteration was approximately 1.7 s on a 2.66 GHz Intel dual core processor. The anisotropic diffusion step was computed using the AOS-Thomas algorithm in 0.92 s. It took approximately 70 s to run the algorithm until the stopping criterion was reached. However, the reconstructions obtained after the first 5 iterations were already quite good, which is achieved in less than 9 s.

#### 4. Discussion and conclusion

In this work we introduced an operator splitting method for solving the fDOT inverse problem using non-linear anisotropic diffusion regularisation. A Levenberg-Marquardt minimisation step alternates with an anisotropic diffusion filtering step, which can be thought of as an image reconstruction step followed by an image denoising/deblurring step. We presented a fast and efficient algorithm for the anisotropic diffusion step, based on the semi-implicit additive operator splitting method. The performance of our reconstruction method was tested on simulated, experimental phantom data and *ex-vivo* mouse data. We also analysed the performance of a few different edge-preserving functions that had been proposed in the literature. It is important to choose an adequate edge-preserving function  $g$  and threshold value  $T$  to be used in the anisotropic diffusion prior, since it determines which gradients are true outliers, and hence, if smoothing should be inhibited in order to preserve edges. We have introduced a  $T$  that varies according to the gradient of the reconstructed optical image and anatomical weighting factor, and is updated at each iteration.

Simulations showed that our proposed splitting method with anisotropic diffusion regularisation provides better reconstructions than the standard Tikhonov method. The different edge-preserving functions have very similar effects on the reconstructions, since they all are monotonically decreasing functions, except the exceedance functional, which penalise gradients smaller than the threshold value, and for larger values these functions are almost flat and tend to zero, meaning that these gradients are preserved in the images. However, we observed that edge-preserving functions that are zero (or very closer to zero) for the gradients considered to be outliers (Tukey and Perona-Malik 2 functions) provide better localised images, which are also qualitatively more accurate, than other functions that, even though in a small amount, still penalise the outliers (Huber and total variation functions). The exceedance functional has the advantage that it does not depend on the parameter  $T$ , since it is based on the probability of a gradient being an edge.

In the phantom study, even though the capillary was placed quite close to the imaging surface and the reconstruction obtained using the Tikhonov prior is already quite good, it is clear that our method with the anisotropic diffusion prior provides better images by removing the noise



in the homogeneous regions and enhancing the capillary edges. In the mouse study, the effect of the anatomical prior on the reconstruction is quite evident. The anatomical prior blocks diffusion across the edges, hence the reconstructed fluorochrome concentration appears confined within the capillary boundaries.

Our results show good estimates of the dimension and location of the fluorophore distribution, which was the main goal of this work. This information is essential in applications such as cancer treatment using radiotherapy, so that radiation is directed towards the tumour while healthy surrounding tissue receives a minimum amount of radiation possible. It is important to note that the fluorochrome is not visible in the reference anatomical image.

An important feature of our method is that it reflects both structural information present in the reference image and also that present in the fluorescent image. Reconstruction do not constrain the fluorophore to be confined by the structures in the reference image, but allow for it to be distributed across tissue boundaries.

The computation of the Jacobian can be time consuming depending on the number of measurements, number of wavelet coefficients used and number of mesh nodes. For example, it took 165 s per source to built the Jacobian for our simulation study. This calculation can be twice as fast if half the wavelet coefficients are used for the data compression. For a mesh with 20% of the nodes of the mesh used in our simulations and 128 wavelets, it takes about 17 s. Alternatively, the Jacobian matrix could be calculated using a matrix-free algorithm [11] or by parallelisation of the Jacobian calculation.

The underestimation of  $h$  observed in the simulations needs to be further investigated. A possible explanation for these results is the partial volume effect, which is known to lead to an underestimation of the fluorescence yield. Furthermore, the simulated target contrast in the nodal basis is lower than in the voxel basis.

In summary, our operator splitting method is straightforward to implement and is computationally fast. The images reconstructed using the anisotropic diffusion prior with anatomical prior show superior quality than those reconstructed by solving a simpler linear problem using Tikhonov regularisation. The weighted anisotropic diffusion prior combines the advantage of preserving/enhancing the edges of the reconstructed optical image while reducing the noise and stopping diffusion through boundaries defined by the anatomical prior. However, further investigation is required to evaluate the adequacy of our method for fDOT image reconstruction of *in-vivo* animal data.

## Acknowledgments

This work has also been supported in part by funding from ECs seventh framework programme FMT-XCT under Grant Agreement No. 201792 and by Comunidad de Madrid and European Regional Development Fund ARTEMIS S2009/DPI-1802.