Local diffusion regularization method for optical tomography reconstruction by using robust statistics

Abdel Douiri, Martin Schweiger, Jason Riley, and Simon Arridge

Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK

Received March 14, 2005; revised manuscript received May 10, 2005; accepted May 11, 2005

We formulate a solution to the diffuse optical tomography (DOT) inverse problem as the minimization of an energy functional of the solution and the data. For the solution prior we introduce a local diffusion regularization potential with a threshold based on robust statistics (the Hubert function). We compare results on simulated data for the Hubert function and two other standard regularization functionals, Tikhonov and total variation. © 2005 Optical Society of America

OCIS codes: 170.5280, 100.6950, 100.3190, 100.3020.

Diffuse optical tomography (DOT) is an emerging medical imaging modality that aims to image the optical properties of biological tissue, particularly the peripheral muscle, breast and the brain.[1–3] It is noninvasive, portable, and can produce images of clinically relevant parameters, such as blood volume and oxygenation, that cannot be obtained by other diagnostic methods.[4] Medical optical tomography has received increasingly intense research interest worldwide over recent years due to advances both in measurement technology and in theoretical and practical understanding of the nature of the image reconstruction problem. Several reconstruction algorithms have been developed to date, and methods for their improvement are an active area of research.

In the frequency domain, the forward problem of light propagation for DOT can be modeled as a diffusion equation,[5] which is given by $r \in \Omega \subset \mathbb{R}^N (N=2,3)$:

$$\left(-\nabla \cdot \kappa(r)\nabla + \mu_a(r) + \frac{i\omega}{c}\right)\Phi(r,\omega) = q(r,\omega), \quad (1)$$

with a Robin-type boundary condition, $m \in \partial\Omega$,

$$\Phi(m,\omega) + 2\kappa(m)A\frac{\partial\Phi(m,\omega)}{\partial n} = 0. \quad (2)$$

Here $\omega \in \mathbb{R}^+$ is the frequency modulation, Φ is the photon density, c is the velocity of light, q is an isotropic source of light in the medium, A is a boundary term that incorporates the refractive index mismatch at the tissue–air boundary, n is the outward normal at $\partial\Omega$, and κ and μ_a are the diffusion and absorption coefficients, respectively. We define $\kappa = 1/[3(\mu_a + \mu'_s)]$, where μ_s' is the reduced scattering coefficient.

The differential equation and the boundary condition of the DOT forward problem can be replaced by the form of a general nonlinear forward equation

$$F(x) = y, \quad (3)$$

where the nonlinear forward operator F maps a distribution of the object $x = (\mu_a, \mu_s')$ into a data image y.

The reconstruction problem consists in finding an estimate for the object x that is mapped by the operator F to perfect data y. The measured data g approximates the perfect data y within an error estimate:

$$\|g - y\| \leqslant \delta, \quad \text{where } \delta > 0. \quad (4)$$

The diffusive nature of the DOT forward model F makes the reconstruction problem ill posed. One solution to overcome this difficulty is to incorporate a regularization term constraining the space of unknowns. The interpretation of the regularization term has a statistical meaning within the Bayesian framework.[6] Under the usual assumptions of a system corrupted by multivariate Gaussian additive noise and according to the maximum *a posteriori* principle, a commonly used solution of problem (3) can be determined by minimizing the following relaxed potential objective function:

$$E(x) := \frac{1}{2}\|g - F(x)\|_{\mathsf{R}}^2 + \tau\int_\Omega \psi(|\nabla x|) \to \min, \quad (5)$$

where x is the original object to be recovered, R is the data-space correlation matrix, $\| \|_{\mathsf{R}}$ is the data-space norm, $\tau > 0$ is a positive regularization parameter, and ψ is the prior function on the space of the solution. The gradient of the objective function $E(x)$ is calculated as follows:

$$\text{grad } E(x) = F'^T(x)\mathsf{R}(F(x) - g) + \tau\mathcal{L}(x)x, \quad (6)$$

where

$$\mathcal{L}(x) = -\nabla \cdot \left(\frac{\psi'(|\nabla x|)}{|\nabla x|}\nabla\right).$$

The term $\psi'(|\nabla x|)/|\nabla x|$ is called the diffusion function.[7] Given the size of the data, an explicit expression of the whole Hessian may become unrealistic for most mainframe computer systems, so we use an efficient approximation instead. Dropping the second-order derivative of the forward map, $F''(x)$, corresponds to the Gauss–Newton approach, and dropping the second-order derivative of the prior, $\mathcal{L}'$, corresponds to the lagged-diffusivity approach.[8] Fur-

0146-9592/05/182439-3/$15.00

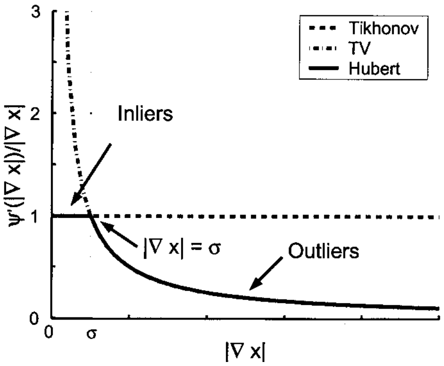


Fig. 1. Illustration of the choice of σ for the Hubert function with respect to the diffusion function.

thermore we use a Krylov method to represent the approximated Hessian implicitly.[9] The approximated Hessian is given by

$$\widetilde{\text{Hess}}\, E(x) = F'^{T}(x)\mathrm{R}F'(x) + \tau\mathcal{L}(x). \tag{7}$$

The choice of an adequate prior function is a wide field of research in itself. A popular choice that is used extensively in reconstruction problems is the Tikhonov function and its variants[10–12]. These regularization methods reduce high-frequency noise and are suitable for homogeneous objects, but for a heterogenous medium, such as occurs in medical applications of DOT, these methods are predisposed to favor oversmooth solutions. In particular, they penalize discontinuities present in the object. To overcome this difficulty, a technique based on the total variation (TV) norm was proposed.[13,14] This technique aims to preserve edges, but the images resulting from its application in the presence of noise are often piecewise constant; thus the finer details in the original object may not be recovered satisfactorily, and ramp structures give staircase effects that can destroy the structures present in a complex heterogeneous object.[15] A solution that allows smooth transitions without penalizing edges and ramp structures could be achieved by using high-order TV-based methods[15] or by using statistical methods such as Markov random fields[16] or histogram-based methods.[17]

In this Letter we suggest the use of the Hubert function.[18] This function acts as a local diffusion regularization process on the contrast of the optical parameters of the object. It has characteristics as follows: (1) it is isotropic (as is the Tikhonov method) in homogeneous regions and where the solution gradient $|\nabla x|$ is small; (2) in the vicinity of an edge, it acts to smooth parallel to the edge in the direction of $(\nabla x/|\nabla x|)^{\perp}$ and not across it, thus preserving the integrity of discontinuities. This function is given by

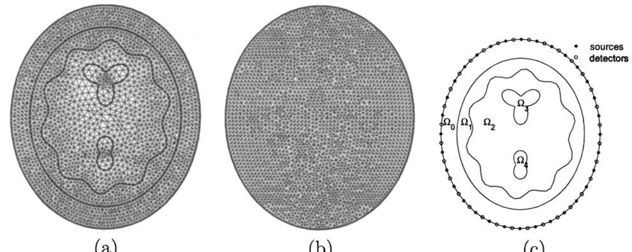
$$\psi(|\nabla x|)\begin{cases} \dfrac{|\nabla x|^2}{2} & \text{if } |\nabla x| \leqslant \sigma \\[2ex] \sigma|\nabla x| - \dfrac{\sigma^2}{2} & \text{otherwise}, \end{cases} \tag{8}$$

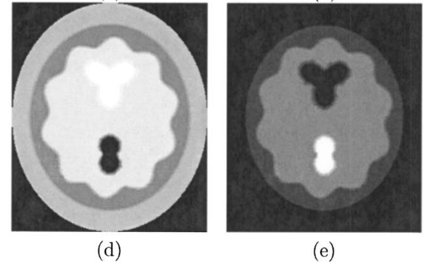
where σ is a scale that is adjusted in each iteration. An automatic choice of this parameter that can make a decision between flat or edge regions could be carried out by using a method based on the cumulated histogram[19] or on the local geometry[20] of the object at each iteration. The first part of Eq. (8) is tantamount to the assumption that the spatial gradient of the object in homogeneous regions is drawn from a zero-mean Gaussian probability density function with a small variance. However, this is not true in the neighborhood of the edges where $|\nabla x|$ contains large jumps. Thus in the neighborhood of the boundaries $|\nabla x|$ can be viewed as an outlier because it does not conform to the statistical assumptions. Thus we can use robust statistics tools to automatically select the robust estimator σ to distinguish the boundaries (outliers) between piecewise constant regions (inliers) in the object[21,22]:

$$\sigma = 1.4826\ \mathrm{MAD}(\nabla x) \tag{9}$$

where $\mathrm{MAD}(\nabla x) = \mathrm{median}[|\nabla x - \mathrm{median}|\nabla x||]$ denotes the median absolute deviation and the constant is the ratio of the standard deviation of a normal random variable to its median absolute deviation. The Hubert regularizer can be seen as a robust estimation procedure that estimates a piecewise smooth object from noisy input data. Figure 1 illustrates the choice of σ. The local diffusion function that ensures the isotropy–anisotropy criterion is closely related to the error norm of ψ, and the influence function is proportional to the derivative ψ' in robust statistical analysis. In our experience, this method seems to act as a compromise between Tikhonov and TV regularization with fast convergence, as the objective function is strictly convex. It is simple to implement and gives reconstructed images of high quality without destroying any structural information.

We now present some results on the comparison of the isotropic (Tikhonov) regularization, the anisotropic (TV) regularization, and the Hubert regularization. In order to show the advantage of the pro-





	μ_a	μ'_s
Ω_0	0.0149	0.80
Ω_1	0.0100	1.00
Ω_2	0.0178	1.25
Ω_3	0.0200	0.90
Ω_4	0.0012	1.75

(f)

Fig. 2. 2D phantom: (a) Adapted forward mesh, 2572 nodes; (b) regular backward mesh, 3062 nodes; (c) schematic diagram; (d) target μ'_a; (e) target μ'_s; (f) optical properties.

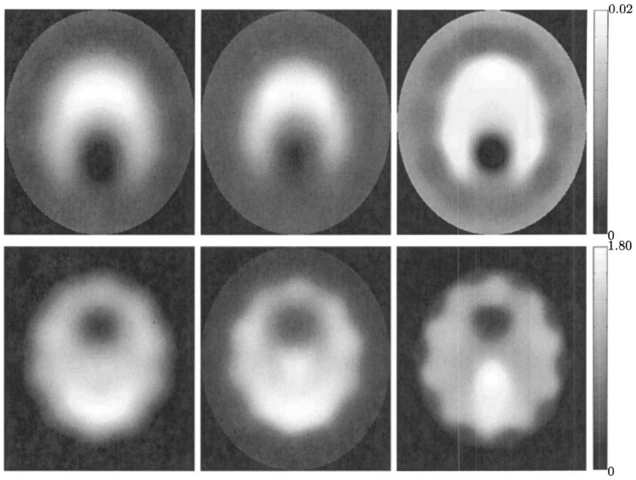


Fig. 3. Comparison of the reconstructions of μ_a (top) and μ_s' (bottom): Tikhonov, NHE=0.012 (left), TV, NHE=0.011 (middle), and Hubert, NHE=0.008 (right).

posed method compared with Tikhonov and TV methods, particularly the enhancement of edge and structure information, we consider a 2D phantom object with scattering and absorption coefficients distributions μ_a and μ_s' varying relative to each region; 32 sources are placed at equidistant spacing along the surface, and 32 detectors are located so that each detector is equidistant between two source sites. Forward data, g, is generated in terms of log amplitude and phase by using a finite-element method implementation of Eqs. (1) and (2).[23] Figure 2 illustrates the meshes for the forward and inverse solvers together with the optical properties of the phantom used for the simulations. The modulation frequency was 100 MHz. All the data for these simulations were corrupted with 1% Gaussian noise, and we used a regularization parameter of $\tau=10^{-5}$ determined according to the L-curve method.[24]

Figure 3 shows the reconstructions of the phantom with Tikhonov regularization, TV regularization, and the Hubert regularization method and their normalized H^1 error $(\mathrm{NHE}=\|x-x_{true}\|^2_{H^1}/\|x_{true}\|^2_{H^1}$, where $\|x\|^2_{H^1}=\|x\|^2_2+\|\nabla x\|^2_2)$. These images clearly show the advantage of the proposed regularization method and demonstrate the shortcomings of the Tikhonov and TV methods, which adversely affect shape as shown in the reconstruction of μ_s' in Fig. 3, particularly the small region of object Ω_4. This regularization method reduces high-frequency noise in homogeneous regions and results in a notable enhancement in recovering the structural information present in the object.

These improvements in the reconstruction of the optical properties with DOT could contribute to a better diagnosis in medical optical tomography in particular and in other nonlinear ill-posed imaging in general.

This work was supported by Engineering and Physical Scienes Research Council grant GR/R86201/01.

References

1. F. F. Jöbsis, Science **198**, 1264 (1977).
2. S. Fantini, S. A. Walker, M. A. Franceschini, M. Kaschke, P. M. Schlag, and K. T. Moesta, Appl. Opt. **37**, 1982 (1998).
3. B. W. Pogue, K. D. Paulsen, C. Abele, and H. Kaufman, J. Biomed. Opt. **5**, 185 (2000).
4. M. Cope and D. T. Delpy, Med. Biol. Eng. Comput. **26**, 289 (1988).
5. S. R. Arridge, Inverse Probl. **15**, R41 (1999).
6. S. Geman and D. Geman, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-6**, 721 (1984).
7. J. Weickert, in *Scale-Space Theory in Computer Vision*, Vol. 1252 of Lecture Notes in Comp. Science (Springer-Verlag, 1997), p. 3.
8. C. R. Vogel, *Computational Methods for Inverse Problems* (SIAM, 2002).
9. M. Schweiger, S. R. Arridge, and I. Nissila, Phys. Med. Biol. (to be published).
10. A. Tikhonov, Sov. Math. Dokl. **4**, 1624 (1963).
11. S. R. Arridge and M. Schweiger, in *Information Processing in Medical Imaging (IPMI'93 Proceedings)*, Vol. 687 of Lecture Notes in Computer Science (Springer-Verlag, 1993), p. 259.
12. B. W. Pogue, T. O. McBride, J. Prewitt, U. L. Osterberg, and K. D. Paulsen, Appl. Opt. **38**, 2950 (1999).
13. L. I. Rudin, S. Osher, and E. Fatemi, Physica D **60**, 259 (1992).
14. K. D. Paulsen and H. Jiang, Appl. Opt. **35**, 3447 (1996).
15. T. Chan, A. Marquina, and P. Mulet, SIAM J. Sci. Comput. (USA) **22**, 503 (2000).
16. C. A. Bouman and K. Sauer, IEEE Trans. Image Process. **2**, 296 (1993).
17. A. H. Hielscher and S. Bartel, J. Biomed. Opt. **6**, 183 (2001).
18. P. J. Hubert, *Robust Statistics* (Wiley, 1981).
19. P. Perona and J. Malik, IEEE Trans. Pattern Anal. Mach. Intell. **12**, 629 (1990).
20. D. S. Luo, M. A. King, and S. Glick, IEEE Trans. Nucl. Sci. **41**, 2800 (1994).
21. F. R. Hampel, E. M. Ronchetti, and P. J. Rousseeuw, *Robust Statistics: the Approach Based on Influence Functions* (Wiley, 1986).
22. M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, IEEE Trans. Image Process. **7**, 421 (1998).
23. S. R. Arridge, M. Schweiger, M. Hiraoka, and D. T. Delpy, Med. Phys. **20**, 299 (1993).
24. P. C. Hansen and D. P. O'Leary, SIAM J. Sci. Comput. (USA) **14**, 1487 (1993).