Histogram derived penalty functions in gradient-based optical tomography

A.H. Hielscher and S. Bartel
State University of New York - Downstate Medical Center, Department of Pathology, 450 Clarkson Avenue, Brooklyn, New York 11203, USA
ahielscher@netmail.hscbklyn.edu

Abstract: It is well known that the image reconstruction problem in optical tomography is ill-posed. In this work we approach the problem within a gradient-based image iterative reconstruction (GIIR) scheme. The reconstruction is considered as a minimization of an objective function. This function can be separated into a least-square-error term, which compares predicted and actual detector readings, and additional penalty terms that contain *a priori* information about the system. In this work penalty functions are considered that are derived from full or partial knowledge of the histogram of the image to be reconstructed.

OCIS codes: (170.6960) Tomography; (170.3010) Image reconstruction techniques

1. Introduction:

A major difficulty in OT remains that the image reconstruction problem is ill-posed or underdetermined. In othe words, there are many different distribution of optical properties inside the medium under investigation that lead to th same set of detector readings on the surface of the medium. In the general, the image reconstruction problem can b formulated as an optimization problem in which the objective function, also called performance measure or erro function,

$$\Phi(\zeta) = \sum_s \sum_d (M_{s,d} - P_{s,d}(\zeta(r)))^2 / 2\sigma_{s,d} \qquad (1)$$

is minimized. In this equation the parameter $\zeta(r)$ is a vector that contains the optical properties at all position r in th medium. If the image is discretized into n pixels, and each pixel can vary in both absorption and scattering coefficien the vector ζ is of length N = 2n. $P_{s,d}$ (ζ(r)) is the predicted detector reading for detector d and source s, and M is th measured value for the same detector and source. The parameter $\sigma_{s,d}$ is a normalization constant. While there ar different ways to optimize the objective function in (1), we concentrate in this work on so-called gradient based iterativ reconstruction schemes [1-4].

In the GIIR scheme ill-posedness is identical to fact that a global minimum is not well defined. In general, th reconstruction result will strongly depend on initial guess ζ_o. These phenomena are well known in the field of genera optimization theory. A typical way to overcome this problem is to add penalty terms that provide additional constrain on the solution space. While the use of penalty terms has been studied for a variety of problems [5-9], there use in GIII schemes for optical tomography has not been explored.

The goal of this paper is to introduce and study the effects of penalty functions that are derived from *a prioi* knowledge about the systems under investigation. *A priori* knowledge is here defined as information other than th difference between predicted and measured data. We focus in this work on information regarding the composition of given tissue volume. A typical situation is that the types of tissue in the system are known. For example, in brai imaging it is known that one encounters, white matter, gray matter, skin, skull and cerebrospinal fluid. The optica properties of these tissues are known with certain degree of accuracy. We will derive the appropriate penalty terms, an show examples that demonstrate the effect of these penalty terms on the quality of the reconstructed images.

2. Methods

For the derivation of useful penalty functions we considered two cases. First, we assumed that *n* tissues with *n* differer optical properties are present in our tissue, however we know neither their location nor their respective volume fractio in the tissue sample (*tissue-type* penalty function). Second, we assumed to have prior knowledge of all tissue type present *and* the volume percentage they occupy but again, not knowing how they are distributed inside the mediur under investigation. This corresponds to knowing the histogram of the medium to be reconstructed and the resultin term is called "histogram penalty function." Since the penalty functions are used in a gradient based reconstructio scheme, it is desirable that the penalty functions are differentiable with respect to the optical properties ζ(r). Therefoi they should not contain discontinuities.

The *tissue-type* penalty function is defined as

$$P^{tt}(\zeta) = \lambda \sum_{x \in S} \sum_{k}^{K} \left(1 - \exp\left(-\frac{(a_k - \zeta(x))^2}{2\sigma_k^2} \right) \right) \quad (2)$$

Here K is the number of different tissue types in the system and N is the number of pixels in the image. The parameter a_k is the most likely optical property of tissue k, and ζ_n is the reconstructed optical property at pixel n. The parameter σ_k can be interpreted as the confidence that a_k is the exact value. For σ_k->0 the exponent becomes a delta function at $\zeta_n = a_k$. For any non-zero value of σ_k there will be a range of ζ_n for which the sum is smaller than 1. Furthermore, for any non-zero value the derivative of this penalty function with respect to a pixel value $\zeta(x_i)$ is easily found and can be used in our gradient based minimization scheme.

A suitable histogram penalty function is more difficult to find. We start by defining the histogram H for a tissue with K discrete tissue types and arbitrary spatial composition $\zeta(\mathbf{x})$ as'

$$H(\zeta_k) \equiv \sum_{x \in S} \delta(\zeta_k, \zeta(\mathbf{x})), \quad \text{with} \quad \delta(\zeta_k, \zeta(\mathbf{x})) \equiv \begin{cases} 1 & for\ \zeta(\mathbf{x}) \in [\zeta_k, \zeta_{k+1}] \\ 0 & otherwise \end{cases} \quad (3)$$

The above sum over all pixels x of an image S, sorts all pixel values $\zeta(\mathbf{x})$ into k bins of the histogram, since the δ-function only contributes to the sum $H(\zeta_\kappa)$, if a pixels lies within the corresponding interval. We can define a penalty function that evaluates the histogram $H(\zeta_\kappa)$ of a reconstructed image relative to the expected histogram $H^0(\zeta_\kappa)$. A straightforward approach is to choose the χ^2-error norm of both functions and define the penalty term P^{Hist} as

$$P^{hist} = \chi^2_{H^0,H} = \sum_{k \in \{bins\}} \frac{(H^0(\zeta_k) - H(\zeta_k))^2}{H^0(\zeta_k)} \quad (4)$$

However, if the definition of the delta function in Eq. (3) is considered, one can see that the derivative of Eq.(4) is not continuous, which makes it unsuitable for gradient based schemes. To assure differentiability we introduce a slightly different definition of the δ-function. Instead of sampling the histogram with a sharply peaked function, we chose to employ a Gaussian of finite width, hence

$$\delta(\zeta_k, \zeta(x_i)) \equiv N \cdot \exp\left(-\frac{(\zeta - \zeta(x_i))^2}{2\sigma^2} \right) \quad (5)$$

Using Eq. (5) to generate H, results in smoothing the histogram and we can now easily calculate the derivative of P^{Hist}.

Each of the penalty functions has to be coupled to the error norm, with a coupling or hyperparamter λ. This parameter fixes the relative strength of the penalty function in the minimization scheme. The exact choice of λ has to be investigated for each case. A good starting point [5] is to define λ in a way that insures the gradients of both, the error norm and the penalty function are of similar magnitude:

$$\lambda = \left| \frac{\partial \Phi^2}{\partial \zeta(x_i)} \right| \Big/ \left| \frac{\partial P}{\partial \zeta(x_i)} \right|^{-1} \approx 1 \quad (14)$$

In this work we always normalize both gradients to magnitude 1, so that λ = 1 in eq. 14 is always fulfilled. We then introduce a second hyper-parameter ω to have well defined control over the percentage relative strength of the prior.

3. Results:

To test how some the penalty terms can improve reconstruction results we consider the following example. Given is a 4x4cm epoxy resin medium that consist of two objects in a background medium (Fig. 1a). The optical properties of the background medium are given by μ_a= 0.1 cm^{-1}, μ_s'= 8 cm^{-1}, (D = 0.905 cm^{-1}). The inclusions have optical properties of μ_a= 0.1 cm^{-1}, μ_s'= 14.0 cm^{-1}, (D = 0.520 cm^{2}ns^{-1}), and μ_a= 0.1 cm^{-1}, μ_s'= 5.0 cm^{-1}, (D = 1.438 cm^{2}ns^{-1}). For the reconstruction the medium is discretized into a 40x40 mesh with Δx = 0.1 cm. Four sources, one in the center of each side, surround the medium. For each source 20 detector readings are available, 4 on each side and 1 on each corner. Therefore we have 4x20x=80 data points. As the initially guess for the GIIR reconstruction scheme we chose D = 1 cm^{2}ns^{-1} and μ_a= 0.1 cm^{-1} for all points in the medium.

Fig. 1b shows the resulting reconstruction after 30 iterations when no penalty terms are used. The minimal and maximal values of D in the heterogeneities are indicated in the figures. The general features of the medium are recovered, while the absolute values are and sharp edges are not recovered. Fig 1c shows the result when the tissue-type penalty term P^{tt} is added with ω = 1. The reconstruction is not successful. Similar images are found with all ω >0.3. For values ω<0.3 the penalty term has almost no effect and the reconstruction is almost identical to the one shown in Fig. 1b. Better results can be achieved with a dynamical ω as shown in Fig. 1d. In this reconstruction ω was changed with each iteration (ω= 0.005•i^2). However, even this approach strongly depends on the appropriate choice of the initial value of ω and the rate at which it is increased. In the case of tissue-type penalty term we did not find a general solution and we were forced to find the optimal coupling parameter empirically.

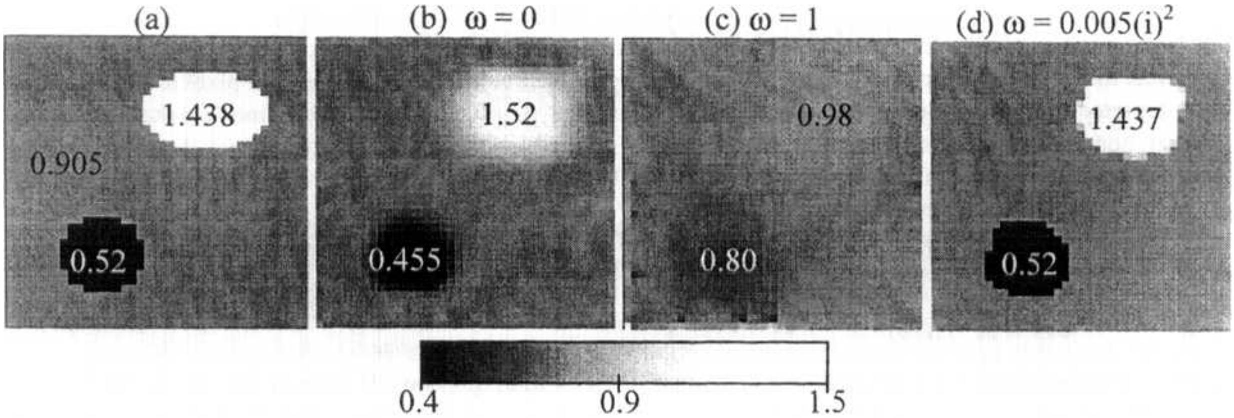

Fig. 1: a) Geometry of original 4x4cm tissue phantom made from epoxy resin with TiO_2 particles. b) Reconstruction result without penalty function, c) with tissue-type penalty function P^{tt} with coupling factor $\omega = 1$ and d) with dynamic coupling factor $\omega = 0.005(i)^2$.

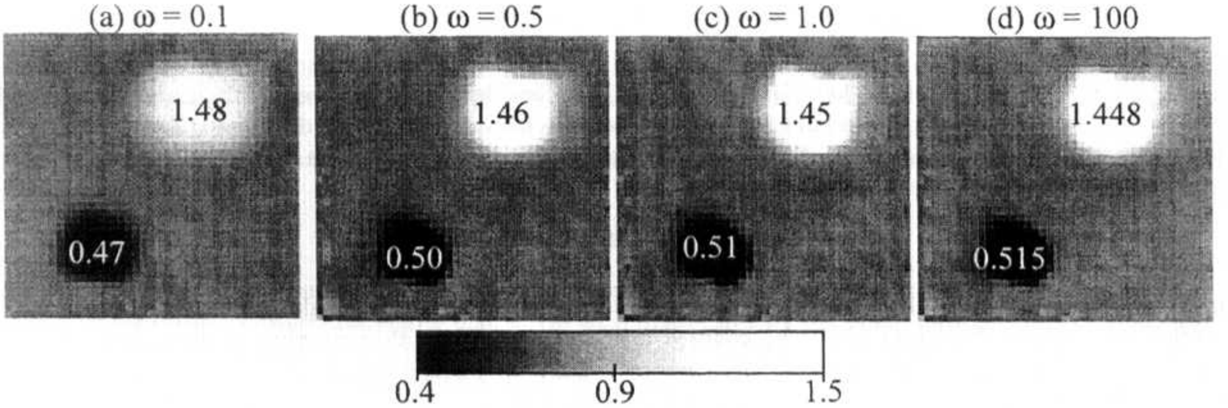

Fig. 2: Reconstruction results with histogram penalty function P^{hist} for different coupling factor ω.

Figs.2a,b, and c displays results when the histogram penalty term P^{hist} is employed for 4 different values of $\omega = 0.5$ 1 and 100. Choosing $\omega < 0.2$ does not result in a significant improvement of the reconstruction and yields images similar to the unbiased case (Fig. 1b). For all $\omega \geq 0.5$ the two objects become more localized and show plateaus of constant D-values. Although their shape does not exactly match the original objects they reflect the correct size or rather volume fraction of these objects, as enforced by the histogram penalty function. Furthermore, the reconstructed D-values remain constant across most of the areas covered by the objects. In Fig. 2d the penalty term is weighted extremely strong relative to the error norm. Nevertheless we still obtain reasonably good agreement with the original image. Therefore the advantages of histogram penalty functions lies in the fact that they appear to be very insensitive to the exact choice of ω. This makes them much more robust and reliable than the tissue-type penalty function.*

4. References

[1] R. Roy and E.M. Sevick-Muraca,"Truncated Newton's optimization scheme for absorption and fluorescence optical tomography: Part I Theory and formulation, "*Optics Express* **10**, 353-371 (1999).

[2] A.D. Klose and A.H. Hielscher, "Iterative reconstruction scheme for optical tomography based on the equation of radiative transfer," *Medical Physics* **26**, 1698-1707 (1999).

[3] A.H. Hielscher, A.D. Klose, K.M. Hanson, "Gradient-based iterative image reconstruction scheme for time-resolved optical tomography," *IEEE Tran.s. Med. Imag* **18**, 262-271 (1999).

[4] S. R. Arridge and M. Schweiger, "A gradient-based optimisation scheme for optical tomography," Optics Express **2** (6), 213-226 (1998).

[5] D. A. Pierre, *Optimization Theory with Applications* (Dover Publications Inc., New york, NY 1986)

[6] Xu, H, "Stochastic Penalty Function Methods for Nonsmooth Constrained Minimization," *J of Opt Theo. and App*, **88**, 709-717 (1996)

[7] Fabien, B. C., "An Extended Penalty Function Approach to the Numerical Solution of Constrained Optimal Control Problems," *Optimal Control Applications & Methods* **17**, 341-348 (1996).

[8] Snyman, J. A.; Stander, N.; Roux, W. J., "A dynamic penalty function method for the solution of structural optimization problems, *Applied Mathematical Modelling*, **18**, 453-460 (1994).

[9] Toma, K., "Protein three-dimensional structure generation with an empirical hydrophobic penalty function," *Journal of Molecular Graphics* **11**, 222-232 (1993).

* This work was supported in part by The Whitaker Foundation (grant # 98-0244), the City of New York Council Speaker's Fund for Biomedical Research: Toward the Science of Patient Care., and the National Institute of Arthritis and Musculoskeletal and Skin Diseases, a part of the National Institute of Health (grant # R01 AR46255-01).