# Three-dimensional Non-local Edge-preserving Regularization for PET Transmission Reconstruction

Daniel F. Yu and Jeffrey A. Fessler

Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122

## Abstract

Tomographic image reconstruction using statistical methods can provide more accurate system modeling, statistical models, and physical constraints than the conventional filtered backprojection (FBP) method. Because of the ill-posedness of the reconstruction problem, a roughness penalty is often imposed on the solution. To avoid smoothing of edges, which are important image attributes, various edge-preserving regularization schemes have been proposed. Most of these schemes rely on information from a local neighborhood to determine the presence of edges. In this paper, we propose an objective function that incorporates non-local boundary information into the 3-D regularization method. We use an alternating minimization algorithm with deterministic annealing to minimize the proposed objective function to jointly estimate region boundary surfaces and object pixel values. We apply variational techniques implemented using level sets to update the boundary estimates; then, using the most recent boundary information, we minimize a space-variant quadratic objective function to update the image estimate. We present three-dimensional reconstructions from real PET transmission data.

## I. Introduction

The problem of reconstructing an unknown image $f$ from a measurement vector $y$ is usually ill-posed. Knowledge of the direct model is rarely sufficient to determine a satisfactory solution. If we obtain the maximum likelihood estimate (MLE) of the image by maximizing the log-likelihood function $L(f; y)$, then the resulting image is very noisy. Thus it is necessary to regularize the solution by imposing a priori assumptions. One simple regularization method supposes that images are globally smooth, and enforces a roughness penalty on the solution by adding a quadratic function to the negative log-likelihood. Such a "penalized" likelihood objective function has the following form:

$$\Phi(f) = -L(f; y) + \beta V(f), \tag{1}$$

where

$$V(f) = \int |\nabla f(\vec{x})|^2 d\vec{x} \tag{2}$$

is a measure of image roughness. The image estimate is obtained by $\hat{f} = \arg\min_f \Phi(f)$, where the minimization with regard to $f$ is often restricted to the nonnegative values. The function given in (2) is often unsatisfactory, due to the fact that many images are not globally smooth. They have region boundaries across which the image values can vary rapidly. The quadratic regularization in (2) causes edges to become blurred.

A more "edge-preserving" penalty results if one replaces the quadratic penalty function in (2) with a nonquadratic function $\psi$ that increases less rapidly than the quadratic function for sufficiently large arguments, such as the Huber function [1]:

$$V(f) = \int \psi(|\nabla f(\vec{x})|) d\vec{x} \tag{3}$$

$$\text{or } V(f) = \int \psi\left(\left|\frac{df}{dx_1}\right|\right) d\vec{x} + \ldots + \int \psi\left(\left|\frac{df}{dx_n}\right|\right) d\vec{x},$$

where $\psi(t) = \begin{cases} t^2/2, & |t| \leq \delta \\ \delta|t| - \delta^2/2, & |t| > \delta. \end{cases}$

This function increases linearly, instead of quadratically, for arguments larger than $\delta$. Thus the objective function penalizes large differences between neighboring pixels less severely than the quadratic penalty, while maintaining the same level of penalty for small differences. This property allows sharper edges in the reconstructed image. However, this type of approach only relies on information from the *local neighborhood* to determine the presence of an edge locally.

## II. Three-Dimensional Non-local Edge-Preserving Regularization

We propose a *non-local penalty* that incorporates boundary and region information into the regularization. We assume that the actual object to be reconstructed is everywhere differentiable (and thus continuous). We also assume that the object consists of regions that are piecewise smooth everywhere except very close to the region boundaries where the object intensity changes rapidly but continuously to values in its neighboring region(s). Thus an edge-preserving penalty function should penalize local gradients that are within each region more than local gradients that are very close to a boundary. Furthermore, we assume that the boundary surfaces separating the regions are smooth.

The objective function we propose also consists of a data-fit term and a penalty term. However, our penalty considers not only the image values but also the characteristics of region boundaries within the image. Following the convention in PDE-based image analysis literature, we present a non-discretized formulation. Let $f$ denote the object, $\Gamma_k \in \mathcal{G}$ denote the $k$th boundary surface, and $\Omega$ denote the domain of the image. Let $R_m(\Gamma) \subset \mathbb{R}^3$ denote the $m$th region, where each $R_m$ does not include its boundary, hence they are open sets. The regions are separated by boundary surfaces $\Gamma$, where $\Gamma = \cup_{k=1}^{K} \Gamma_k$. We assume that the number of boundary surfaces $K$ is fixed (and known a priori) and that the boundary surfaces do not touch each other, i.e. $\Gamma_{k_1} \cap \Gamma_{k_2} = \emptyset$ if $k_1 \neq k_2$, which is reasonable for transmission tomography.

We propose the following objective function of the object $f$ and the boundaries $\Gamma$ [2]:

$$J(f, \Gamma) = -L(f; y) + V(f, \Gamma) \tag{4}$$

$$V(f, \Gamma) = \sum_{k=1}^{K} \left[ \beta J_2(f, k, \Gamma) + \mu J_1(\Gamma_k) \right] \tag{5}$$

$$J_1(f, \Gamma_k) = \int_{\Gamma_k} dS \tag{6}$$

$$J_2(f, k, \Gamma) = \int_{\Omega} h_k(\delta_k(\vec{x}, \Gamma)) |\nabla f(\vec{x})|^2 d\vec{x}. \tag{7}$$

The first term $-L(f; y)$ is the negative log-likelihood that measures the "faithfulness" of the reconstructed object to the measured data[1]. The term $\int_{\Gamma_k} dS$ penalizes the area of the boundary surface, so that the boundary surfaces remain smooth. The term $J_2(f, k, \Gamma)$, which is rotationally invariant, penalizes local gradients inside each region more than local gradients close to the boundary; $\delta_k : \mathbb{R}^3 \times \mathcal{G}^k \to \mathbb{R}$ is the signed distance of $\vec{x}$ to $\Gamma$ if the closest point on $\Gamma$ to $\vec{x}$ lies on $\Gamma_k$, otherwise, the function is zero (hence all locations $x$ where $\delta_k(x, \Gamma)$ is nonzero are necessarily in the two regions that are separated by $\Gamma_k$). The function $h_k : \mathbb{R} \to [0, 1]$ maps small arguments to values near zero and larger arguments to values near unity. For simplicity, we only use $h_k$'s that belong to $C^\infty(-\infty, +\infty)$. The $J_2$ term has a similar effect on the reconstructed image as the penalty described by Eqn (3); but in $J_2$, how much the local gradients at a specific location is penalized is decided by where this location is with regard to the boundary, hence the penalty is "nonlocal".

We use alternating minimization to jointly minimize the objective function given in (4) over $f$ and $\Gamma$, i.e., we first hold $f$ constant and minimize $\Phi$ with regard to $\Gamma$, then using the most recent estimate of $\Gamma$, we minimize $\Phi$ with regard to $f$; we alternate between these two steps until convergence. When $f$ is fixed, the second and third terms depend on $\Gamma$. We must minimize the following objective:

$$J_f(\Gamma) = \sum_{k=1}^{K} \left( \beta J_2(f, k, \Gamma) + \mu \int_{\Gamma_k} ds \right) \tag{8}$$

$$\Gamma^{n+1} = \arg\min_{\Gamma} J_{f^n}(\Gamma), \tag{9}$$

where $J_2$ was defined in (7). As is common in PDE-based image analysis, we perform steepest descent with respect to $\Gamma$. For any point $\vec{v} = (x, y)$ on the boundary $\Gamma$ we evolve that point according to the following differential equation:

$$\frac{d\vec{v}}{dt} = -\frac{\delta J_f(\Gamma)}{\delta \vec{v}}, \tag{10}$$

where the right-hand side is the negative functional derivative of the objective. Since it is difficult to analytically derive the functional derivatives of $J_2$, we evaluate its functional

derivatives numerically. The functional derivative of $J_2$ must point in the normal direction of the curve, as any movement in the tangential direction would not change the curve. We use a scheme similar to the central difference method for evaluating local derivatives. (Central differences are usually accurate to a higher order than one-sided differences.) For a given point $\vec{p}_0$ on the contour $\Gamma$, we define a function $z$ which is zero except in the neighborhood of $\vec{p}_0$ and for which $\Gamma + z$ differs from $\Gamma$ only in the normal direction. (We can imagine some force being exerted on the curve; this force is nonzero only in the neighborhood of $\vec{p}_0$; if we exert this force in the normal direction of the curve at $\vec{p}_0$ for an infinitesimal period of time, then we will have a small perturbation of the curve at $\vec{p}_0$ in the normal direction). Using this idea, we approximate the functional derivative of $J_2(\Gamma)$ at $\vec{v} = \vec{p}_0$ as follows:

$$\frac{\delta J_2}{\delta \vec{v}} \approx \frac{1}{2} \frac{J_2(\Gamma + z) - J_2(\Gamma - z)}{\Delta \sigma}$$
$$= \frac{1}{2} \left( \frac{J_2(\Gamma + z) - J_2(\Gamma)}{\Delta \sigma} - \frac{J_2(\Gamma - z) - J_2(\Gamma)}{\Delta \sigma} \right)$$
$$\triangleq \widehat{\frac{\delta J_2}{\delta \vec{v}}}, \tag{11}$$

where $\Delta \sigma$ is the area lying between the curve $\Gamma$ and the perturbed curve $\Gamma + z$ [3].

For the last term in $J_f(\Gamma)$, the direction in which the curve length decreases most rapidly is when [4]:

$$\frac{\partial \Gamma}{\partial t} = -\kappa \vec{\mathcal{N}}, \tag{12}$$

i.e., the speed of the evolution at any point is the curvature of the boundary at that point, and the curve evolves in the inward normal direction. Combining (10), (11), and (12), we evolve the boundary using the level set method [4–6] according to the following:

$$\frac{d\vec{v}}{dt} = -(\mu \kappa + \widehat{\frac{\delta J_2}{\delta \vec{v}}}) \vec{\mathcal{N}}. \tag{13}$$

Evolving the curve according to (13) yields a curve that approximately minimizes $J_f$; we call this step the "boundary estimation" step.

The force exerted by $J_2$ in (8) is nearly zero in smooth regions, and is only significant close to the actual boundary where local gradients are large. Fig. 1 illustrates this property in one dimension. Let $f'$ denote the derivative of $f$ in the $x$ direction; let $\Gamma_o$ denote the old boundary at 0.3 and $\Gamma_n$ denote the new boundary at 0.35. Moving the boundary $\Gamma$ from 0.3 to 0.35 would change $h(\delta(\cdot, \Gamma))$ from $h(\delta(\cdot, \Gamma_n))$ to $h(\delta(\cdot, \Gamma_n))$, i.e., the "valley" of $h(\delta(\cdot, \Gamma))$ is moved from 0.3 to 0.35, but the change in the roughness penalty, i.e., $\int (h(\delta(x, \Gamma_n)) - h(\delta(x, \Gamma_o))) |f'| dx$, would be very small. Thus evolution according to (13) alone would require a fairly close initialization to the actual boundary. We circumvent this problem by using the initialization procedure for the boundary that employs another force $J_3$ (from a global measure) which ensures that the boundary moves no matter where the contour

---

[1] The 3-D measurement can be a stack of 2-D measurements, i.e., there is no inter-slice rays as in the case of 2.5-D transmission scans; or it can be truly 3-D.

is:

$$J_3(f) \triangleq \sum_{m=1}^{M} \gamma \int_{R_m} \left| f(\vec{x}) - \frac{\int_{R_m} f(\vec{x}')d\vec{x}}{\int_{R_m} d\vec{x}} \right|^2 d\vec{x}.$$

$J_3$ penalizes the difference between every pixel value and the average pixel value of its region. This is a global measure which exerts a force on the curve no matter how close the boundary estimate is to the image gradients.
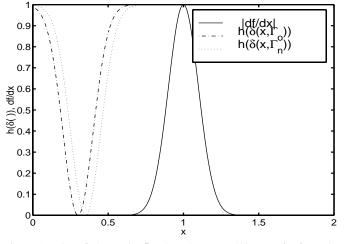


Figure 1: Plot of change in $J_2$ when the "curve" is very far from the actual boundary

For the second stage of the minimization, we hold $\Gamma$ fixed at its previous estimate $\Gamma^n$ and minimize with regard to $f$. When $\Gamma$ is held fixed, the relevant terms in the objective function (4) are the following:

$$J_{\Gamma}(f) = -L(f; y) + \beta \sum_{k=1}^{K} J_2(f, k, \Gamma) \quad (14)$$

$$J_2(f, k, \Gamma) = \int_{\Omega} h_k(\delta_k(\vec{x}, \Gamma)) |\nabla f(\vec{x})|^2 d\vec{x}.$$

Hence we minimize $J_{\Gamma}(f)$ with regard to $f$ as follows:

$$f^{n+1} = \arg \min_f J_{\Gamma^n}(f). \quad (15)$$

When updating the boundary using (9), the $h$ function in $J_2$ pushes the boundary toward image locations where the gradient is large; when updating the objective $f$ using (15), the $h$ function imposes a space-varying weighting of the penalty on local gradients. This weighting depends on the signed distance from each pixel to the nearest estimated boundary. Every term in (14) is quadratic in $f$, except possibly the log-likelihood term, which involves logarithms in the case of Poisson measurements. Therefore, the minimization problem (15) is a standard penalized likelihood problem, and we can minimize $J_{\Gamma}$ over $f$ using methods such as the conjugate gradient method [7,8] (if quadratic) or the paraboloid surrogates/coordinate ascent method (if not) [9].

We iteratively alternate between the two steps (9) and (15). Both these two steps will, under ideal circumstances[2],

---

[2]Under realistic circumstances, where $dt$ is taken to be finite, the

monotonically decrease the objective as defined in (4). In addition, the objective is bounded below, so the algorithm will presumably converge toward a local minimum.

The proposed regularization method allows edge-preserving reconstruction of piecewise smooth objects such as PET attenuation maps. It does not require any textbook values for the attenuation coefficients, and, as illustrated in the results below, it allows for nonuniform attenuation coefficients in regions such as the lungs, in contrast to some forms of segmentation methods. Qualitatively, the method produces good attenuation maps even from 3-minute transmission scans. Quantitatively, the bias-variance tradeoffs of the proposed method (in 2D results not shown) are superior to statistical reconstruction with the "conventional" local Huber penalty over the spectrum of regularization parameters. The principal disadvantages of the proposed method are implementation complexity and computation time, although with an optimized implementation the computation associated with the level sets should be less than the forward and backprojection computations.
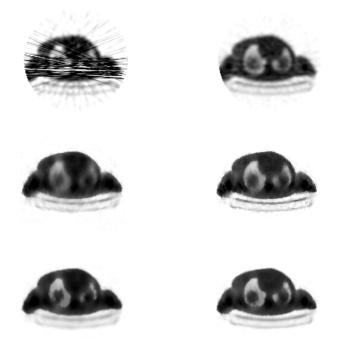


Figure 2: Slice No. 6: left column, 3-minute scan; right column, 10-minute scan; top row, FBP reconstruction; middle row, 3-D Huber penalty; bottom row, proposed penalty

## III. RESULTS

We obtained a 10-minute transmission scan of a real patient on a CTI 921 PET scanner, then we thinned this data to the equivalent of a 3-minute transmission scan. The image consists of $134 \times 134 \times 47$ pixels; the sinogram has 47 slices, each consisting of 192 radial samples and 160 angular samples. Figs. 4 and 5. show the reconstructions using the proposed

minimization of (8) according to a discretized version of (10) may not be exactly monotonic. Such effects are inevitable when continuous formulations are discretized.

penalty. Fig. 2 compares the 6th slice of the reconstruction using FBP, the Huber penalty, and the proposed penalty. Fig. 3 shows the boundary surfaces extracted using the proposed penalty.
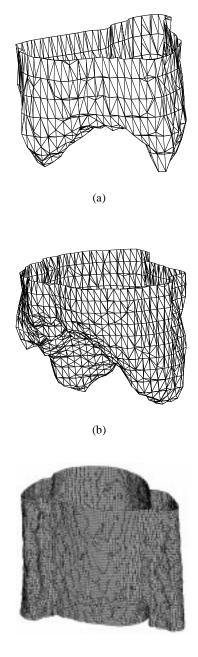
## IV. DISCUSSION AND CONCLUSION

We have presented a new regularization method for tomographic image reconstruction based on a nonlocal penalty function. Simulations show that the nonlocal penalty produces transmission reconstructions with better ROI bias/variance tradeoffs than a local Huber penalty; when these transmission reconstructions are applied to ideal emission data, the nonlocal penalty used for transmission reconstruction produces emission images with smaller variances (for a fixed spatial resolution) for most pixels in the image [10]. However, reconstruction using the proposed penalty is more time consuming than using "conventional" local penalties, *i.e.*, 7-10 times the time using local penalties.

Currently, the $h$ functions are chosen experimentally, *i.e.*, trial and error. A more systematic approach in choosing $h$ functions, so that the transition in pixel values between neighboring regions can be carefully controlled, will make this method much easier to use.

## V. REFERENCES

[1] H. Erdoğan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Tr. Med. Im.*, vol. 18, no. 9, pp. 801–14, September 1999.

[2] D. F. Yu and J. A. Fessler, "Edge-preserving tomographic reconstruction with nonlocal regularization," *IEEE Tr. Im. Proc.*, 1999. Submitted.

[3] I. M. Gelfand and S. V. Fomin, *Calculus of variations*, Prentice-Hall, NJ, 1963. Translation by R A Silverman.

[4] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, "A geometric snake model for segmentation of medical imagery," *IEEE Tr. Med. Im.*, vol. 16, no. 2, pp. 199–210, April 1997.

[5] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: a level set approach," *IEEE Tr. Patt. Anal. Mach. Int.*, vol. 17, no. 2, pp. 158–76, February 1995.

[6] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Conformal curvature flows: from phase transitions to active vision," *Archive for Rational Mechanics and Analysis*, vol. 134, no. 3, pp. 275–301, 1996.

[7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical recipes in C*, Cambridge Univ. Press, 1988.

[8] J. A. Fessler and S. D. Booth, "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction," *IEEE Tr. Im. Proc.*, vol. 8, no. 5, pp. 688–99, May 1999.

[9] H. Erdoğan and J. A. Fessler, "Accelerated monotonic algorithms for transmission tomography," in *Proc. IEEE*

(a)



(b)



(c)

Figure 3: Boundary surfaces extracted by the proposed penalty: the two lungs and the body. (Note: due to the 1Mbtye file-size limitation of this proceedings, this figure was rasterized, significantly degrading its appearance.)

*Intl. Conf. on Image Processing*, volume 2, pp. 680–4, 1998.

[10] F. Yu, *Statistical methods for transmission image reconstruction with nonlocal edge-preserving regularization*, PhD thesis, Univ. of Michigan, Ann Arbor, MI, 48109-2122, Ann Arbor, MI., 2000.
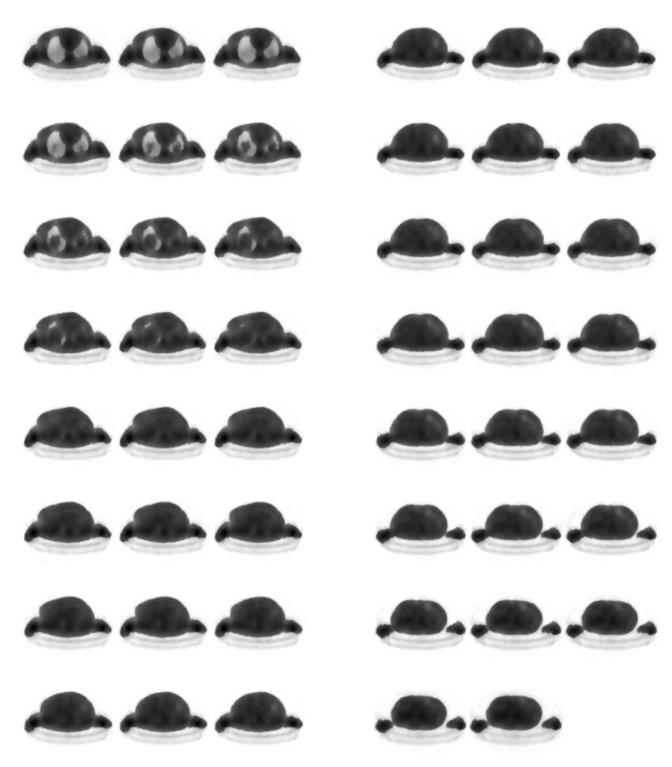
Figure 4: Reconstruction using proposed 3-D penalty: slices 1-24

Figure 5: Reconstruction using proposed 3-D penalty: slices 25-47