Parameter-Free Adaptive Total-Variation Based Noise Removal and Edge Strengthening for Mitochondrial Structure Extraction

P. Blomgren and C. Bazan

June 2007

Publication Number: CSRCR2007-16

Computational Science & Engineering Faculty and Students Research Articles

Database Powered by the Computational Science Research Center Computing Group

COMPUTATIONAL SCIENCE & ENGINEERING

SAN DIEGO STATE UNIVERSITY

Computational Science Research Center
College of Sciences
5500 Campanile Drive
San Diego, CA 92182-1245
(619) 594-3430

MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume α, Number **0**β, January **200**γ

http://www.mbejournal.org/

pp. **1–10**

PARAMETER-FREE ADAPTIVE TOTAL-VARIATION BASED NOISE REMOVAL AND EDGE STRENGTHENING FOR MITOCHONDRIAL STRUCTURE EXTRACTION

Abstract. We propose an iterative method that will allow noise removal and edge strengthening for mitochondrial images. The model is based on the well established total variation (TV) approach to image processing. The main objective of the method is to decompose the observed image as $u_0 = u + \eta$, whereby we can obtain a good approximation to the true image u, with minimum human intervention. The model can be used as a preprocessing tool that will allow better segmentation results.

Peter Blomgren

Department of Mathematics and Statistics,
Dynamical Systems Group & Computational Sciences Research Center,
San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-7720

Carlos Bazan

Computational Science Research Center, San Diego State University
5500 Campanile Drive, San Diego, CA 92182-1245

(Communicated by Peter Blomgren)

1. Motivation.

Segmentation is a classical image processing technique employed in electron tomography 3D structural reconstruction of biological systems. Segmentation allows a tomogram to be decomposed as a series of geometrical objects (contours) that can be rendered as a 3D model. This decomposition into structural components also facilitates interpretation, communication of results, and measurements [7]. In spite of the efforts that have been made to automate this process [2, 13, 19, 25], manual segmentation is still the tool of choice in most cases.

The 3D structural extraction from the tomographic volume relies on the processing of several hundred two-dimensional slices. When these images lack the required quality, the segmentation process can become a great burden. The model we are proposing aims at helping both the automated segmentation and the manual tracing processes by the removal of noise and the strengthening of edges of mitochondrion images. The model can be used as a preprocessing tool that will improve segmentation results.

The present paper was inspired by the challenges posed by processing and interpreting image data of mitochondria obtained by electron microscopy. Structures of interest include multicomponent structures [18], crista junctions [17], and membrane architecture [15]. Current methods in, and results of modern electron microscopy can be found in [7, 8, 16].

2000 *Mathematics Subject Classification.* 62H35.
Key words and phrases. total variation (TV), noise removal, image analysis, electron microscopy, mitochondria.

2. **Background.** Rudin, Osher and Fatemi [20] propose image noise removal by minimization of the total variation (TV) norm. The corresponding constrained minimization formulation for the restored image, u, is

$$\min_{u\in \mathrm{BV}(\Omega)} \mathrm{TV}\left(u\right) \stackrel{\text{def}}{=} \min_{u\in \mathrm{BV}(\Omega)} \int_\Omega \|\nabla u\|\, dxdy. \tag{1}$$

subject to a discrepancy constraint involving the original, measured, noisy, image u_0:

$$\frac{1}{2}\int_\Omega \left(u-u_0\right)^2\, dxdy = |\Omega|\sigma^2. \tag{2}$$

This constraint uses *a priori* information of the variance of the noise, σ^2, under the assumption the noise is normally distributed with mean zero. The quantity $|\Omega|$ measures the size of the image domain. In most practical application the noise intensity will not be known and the success of the method will require a good noise-estimate.

Common PDE-based solution strategies try to solve the corresponding Euler-Lagrange equation

$$\begin{cases} -\nabla\cdot\left(\dfrac{\nabla u}{\|\nabla u\|}\right)+\lambda\left(u-u_0\right)=0, \text{ in } \Omega \\[2ex] \dfrac{\partial u}{\partial \vec{n}}=0, \text{ on } \partial\Omega. \end{cases} \tag{3}$$

The solution procedure proposed in [20] uses a parabolic equation which introduces synthetic time, t, as an evolution parameter. Equivalently, this can be viewed as the gradient descent method applied to the minimization problem, *i.e.*

$$\begin{cases} u_t=\nabla\cdot\left(\dfrac{\nabla u}{\|\nabla u\|}\right)-\lambda\left(u-u_0\right), \\[2ex] \dfrac{\partial u}{\partial n}=0, \text{ on } \partial\Omega, \\[2ex] u(t=0,x,y)=u_0. \end{cases} \tag{4}$$

Here the original image, u_0, is used as initial condition for the PDE. The parameter λ measures the tradeoff between regularization, *i.e.* minimization of the TV-norm, and fidelity to the measured image. Since, typically, the noise level, and therefore the correct Lagrange multiplier λ are unknown, Rudin-Osher-Fatemi [20] suggest a dynamic value estimated by Rosen's gradient-projection method, which as $t\to\infty$ converges to

$$\lambda = -\frac{1}{2|\Omega|\sigma^2}\int_\Omega \left[\frac{\nabla u^T\left(\nabla u-\nabla u_0\right)}{\|\nabla u\|}\right] dxdy. \tag{5}$$

This evolution scheme is highly nonlinear and not well-posed in strong sense [21]. Numerically, difficulties arise as $\|\nabla u\|\to 0$. When the scheme converges, it does so at a linear rate. Further, direct application of classical schemes, *e.g.* the affine invariant form of the damped Newton method as described in Deuflhard [6] generically run into convergence problems due to the ill-conditioning of the problem introduced by the non-linearity. In practice, it is common to use a slightly modified version of the TV-norm [1]:

$$\int_{\Omega} \sqrt{\|\nabla u\|^2 + \beta}\, dxdy, \tag{6}$$

where β is a small positive number which smooths out the "corner" at $\|\nabla u\| = 0$.

Also, when β is very small, the Newton method does not work satisfactorily. To overcome the problems presented by the highly nonlinearity of the problem Vogel and Oman [24] propose a fixed point iteration scheme,

$$-\nabla \cdot \left(\frac{\nabla u^{k+1}}{\|\nabla u^k\|} \right) + \lambda \left(u^{k+1} - u_0 \right) = 0. \tag{7}$$

This is a robust scheme but it is only linearly convergent. Golub, Chan and Mulet [9] use interior-point primal-dual implicit method to solve the Euler-Lagrange equation by introducing a new variable

$$w = \frac{\nabla u}{\sqrt{\|\nabla u\|^2 + \beta}}, \tag{8}$$

and writing the problem as a system of nonlinear partial differential equations as follows

$$\begin{cases} -\nabla \cdot w + \lambda \left(u - u_0 \right) = 0 \\ \\ w\sqrt{\|\nabla u\|^2 + \beta} - \nabla u = 0. \end{cases} \tag{9}$$

All these approaches, which attempt to solve the original TV-minimization problem, lead to solutions which exhibit the "staircase effect," *i.e.* a strong preference for piecewise constant patches.

Marquina and Osher [11] propose a different version of the transient parabolic equation that helps speed up the convergence of the scheme. The modified evolution equation is

$$u_t = \|\nabla u\| \nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} \right) - \|\nabla u\| \lambda G_\sigma * \left(G_\sigma * u - u_0 \right), \tag{10}$$

in Ω, for $t > 0$, where $G_\sigma(x, y)$ is the heat kernel. The well-posedness of this equation in the sense that there is a maximum principle that determines the solution is shown by Osher and Sethian [14]. This approach fixes the staircase problem of the original scheme and is used for removal of both blur and noise.

Strong and Chan [23] introduce the weighted total variation functional for spatially adaptive image restoration:

$$TV_\alpha = \int_{\Omega} \alpha(x) \|\nabla u\|\, dx. \tag{11}$$

Blomgren, Chan and Mulet [3] propose a new approach considering regularizing functionals of the type

$$R(u) = \int_{\Omega} \Phi\left(\|\nabla u\| \right) dx, \tag{12}$$

for suitable real functions Φ. They consider the functional in (12) for $\Phi(x) = x^p$, for $p \in [1, 2]$

$$R(u) = \int_\Omega \|\nabla u\|^p \, dx. \tag{13}$$

For the exponent $p = 1$, one has the TV-norm and when $p = 2$, one would be using the L^2-norm of the image gradient. Song [21], in his dissertation, pursues this approach further and renames it "Adaptive TV Model." The model considers

$$\min_u \frac{1}{p} \int_\Omega \|\nabla u\|^p \, dxdy, \quad 1 < p < 2 \tag{14}$$

subject to $\frac{1}{2} \int (u - u_0)^2 \, dxdy = |\Omega|\sigma^2$. The Euler-Lagrange equation for this model is

$$u_t - \nabla \cdot \left(\|\nabla u\|^{p-2} \nabla u \right) + \lambda (u - u_0) = 0 \tag{15}$$

The proof of the uniqueness of the solution is given in [10].

Levine, Chen, Stanich and Rao [5, 12] propose a variant to that of Blomgren *et al*, where they define the exponent p based on the observed data u_0 only, their model is:

$$J(u) = \int_\Omega \phi(x, \nabla u) + \frac{\lambda}{2} \int_\Omega |u_0 - u|^2,$$

where

$$\phi(x, r) = \begin{cases} \frac{1}{p(x)} |r|^{p(x)} & \text{if } |r| < \varepsilon \\ |r| - \frac{\varepsilon p(x) - \varepsilon^{p(x)}}{p(x)} & \text{if } |r| \geq \varepsilon. \end{cases}$$

Here, $\varepsilon > 0$ is fixed, and $p(x)$ is based on a smoothed version of the observed image u_0,

$$p(x) = \frac{1}{1 + k |\nabla G_\sigma * u_0(x)|^2}$$

where k and σ are adjustable parameters, and

$$G_\sigma(x) = \frac{1}{\sigma} e^{\frac{-|x|^2}{4\sigma^2}}$$

is a Gaussian smoothing kernel. The authors show existence and uniqueness of minimizers for this functional, and develop a numerical method for computing them based on gradient descent.

Chambolle [4] also touches upon this subject where he combines two functionals $\int |\nabla u|$ and $\int |\nabla u|^2$ as

$$F(u) = \frac{1}{2\varepsilon} \int_{|\nabla u| < \varepsilon} |\nabla u|^2 + \int_{|\nabla u| \geq \varepsilon} \left(|\nabla u| - \frac{\varepsilon}{2} \right) + \int_\Omega |u - u_0|^2$$

where ε is an adjustable parameter to be chosen. The Euler-Lagrange equation for this functional resembles that of the models discussed in this section.

Schults, Bollt, Chartrand, Esedoglu and Vixie [22] have recently revisited the subject and they suggest to minimize the following functional,

$$\min_u J(u) = \int_\Omega |\nabla u|^{p(|\nabla u|)} + \frac{\lambda}{2} \int_\Omega |u_0 - u|^q, \ q = 1 \text{ or } 2$$

For two cases:

Case 1: $p(x) = P(|\nabla (G_\sigma * u_0)(x)|)$; and

Case 2: $p(x) = P(|\nabla (G_\sigma * u)(x)|)$.

They prove existence in both cases, and uniqueness in the case of $q = 2$.

3. **Dynamic Implementation.** We implement a variation of Blomgren *et al*'s version of the fully nonlinear Euler-Lagrange equation (10),

$$u_t - \|\nabla u\| \nabla \cdot \left(L\left(\|\nabla u\|^{p-2} \right) \nabla u \right) + \Lambda (u - u_0) = 0 \tag{16}$$

defined in the domain Ω with boundary conditions $u_{\vec{n}} = 0$ on $\partial\Omega$ (where $\vec{n}$ is the outward unit normal vector to the boundary of the domain Ω). The Neumann boundary conditions should guarantee that the filtering does not significantly affect the average gray value of the image. The initial condition is the original image $u(0, x, y) = u_0(x, y)$ in Ω.

The model (16) can be regarded as an "adaptive TV model with morphological convection and anisotropic diffusion." As opposed to the approach of Marquina and Osher [11], we implement a user-independent choice of all the parameters in the model. We start by estimating the standard deviation of the noise, *i.e.* the parameter σ. Since we consider that the image has been perturbed by additive Gaussian "white" noise, $u_0 = u + \eta$, the variance of the noisy image equals the sum of the variance of the original image and the variance of the noise, $\sigma^2_{u_0} = \sigma^2_{G*u_0} + \sigma^2_\eta$. Here, the variance of the (unknown) original image is approximated by the variance of the convolved noisy image. This parameter will be updated iteratively as we will see below.

For the parameter λ, we implement a variation of the method suggested in [20]. Instead of integrating over the domain Ω, we implement a pixel-wise $\Lambda \equiv \|\nabla u\| \lambda$ as

$$\Lambda = -\frac{1}{2|\Omega|\sigma^2} \left[\nabla u^T (\nabla u - \nabla u_0) \right]. \tag{17}$$

The dynamic parameter Λ has the following attributes:

1. The smaller the value of Λ, the more the diffusion contributed by the forcing term. Analogously, the larger the value of Λ, the lesser the diffusion contributed by the forcing term.
2. At the beginning of the time-marching iterations the gradients $\nabla u \approx \nabla u_0$, therefore the gradient discrepancies $(\nabla u - \nabla u_0)$ are very small and the forcing term tends to contribute more to the diffusion process. In areas of large gradients (*i.e.* near edges), these values compensate for the small terms $(\nabla u - \nabla u_0)$.
3. As the evolution progresses the discrepancies $(\nabla u - \nabla u_0)$ get larger. Near edges, the forcing term prevents diffusion and helps reach convergence.

We can also get an *a posteriori* estimate to the variance of the noise σ^2 by integrating over the domain after convergence,

$$\sigma^2 = -\frac{1}{2|\Omega|} \int_\Omega \frac{1}{\Lambda} \left[\nabla u^T (\nabla u - \nabla u_0) \right] dxdy. \tag{18}$$

This will be an improved value that can be used to run the model with a better estimate of the unknown parameter σ.

The diffusion tensor $L\left(\|\nabla u\|^{p-2}\right)$ incorporates the parameter $1 \leq p \leq 2$, as suggested in [3]. The diffusion tensor becomes

$$L\left(\|\nabla u\|^{p-2} \right) = \begin{bmatrix} \|\nabla u\|^{p^x - 2} & -\|\nabla u\|^{p^{xy} - 2} \\ -\|\nabla u\|^{p^{xy} - 2} & \|\nabla u\|^{p^y - 2} \end{bmatrix}, \tag{19}$$

where p^x, p^y, p^{xy}, are the following unnormalized Gaussians:

$$\begin{aligned} p^x &= 1 + e^{-\tilde{u}_x^2/4\sigma} \\ p^y &= 1 + e^{-\tilde{u}_y^2/4\sigma} \\ p^{xy} &= 1 + e^{-\left(\tilde{u}_x^2+\tilde{u}_y^2\right)/4\sigma} \end{aligned} \tag{20}$$

In equation (20) above, $\tilde{u}_x$ and $\tilde{u}_y$ are the gradient-components of the convolved noisy image $G * u_0$ used to estimate the unknown parameter σ.[1] The dynamic parameters p^x, p^y, p^{xy}, have the following attributes:

1. For every pixel in the image, the parameters take values $1 \leq p^x \leq 2$, $1 \leq p^y \leq 2$, and $1 \leq p^{xy} \leq 2$.
2. When $p^x = 1$, $p^y = 1$, or $p^{xy} = 1$ the model uses the TV-norm in the corresponding direction, and when $p^x = 2$, $p^y = 2$ or $p^{xy} = 2$, the model uses the L^2- norm in the corresponding direction.
3. When the parameters $1 < p^x < 2$, $1 < p^y < 2$ and $1 < p^{xy} < 2$, the model interpolates between both norms.

4. **Experimental Results.**

The mitochondrial images produced by the electron microscope are of extremely low contrast (see Figure 1). If we plot the distribution of intensities of the image we observe that the intensity range is very narrow. It does not cover the potential range of gray tones [0, 255], and is missing the high and low values that would result in good contrast. To improve the contrast in the image we spread the intensity values over the full range of the image by histogram equalization. This process notably improves the contrast of the image which becomes more suitable for the application of our method.

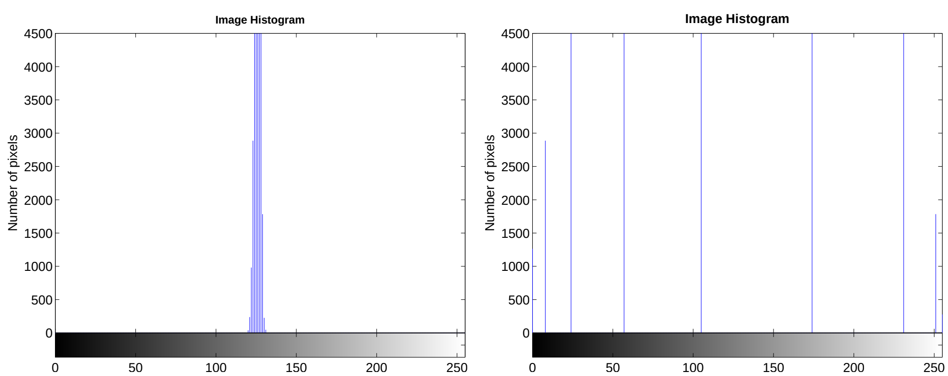

FIGURE 1. (Left) Image histogram: the intensity range is very narrow in the original electron microscope image, and when the data is visualized it looks “flat gray” to the human eye; (Right) Image histogram: after histogram equalization, the intensity range is spread out over the range [0, 255].

After the histogram equalization we estimate the value of the variance of the noise as described above, $\sigma_\eta^2 = \sigma_{u_0}^2 - \sigma_{G*u_0}^2$. This value will be dynamically updated each time the model reaches convergence using (18). Figure 2 shows the

[1] Alternatively, at a higher computational cost, the current iterate u can be used for an updated estimate, $\sigma = G * u$

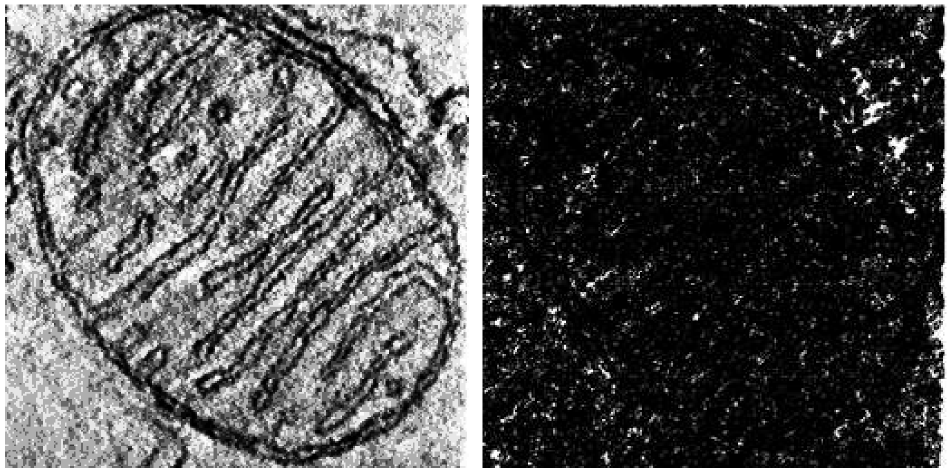

FIGURE 2. (Left) Mitochondrion image after histogram equalization. (Right) Image contours: although the human eye can extract the structure in the left panel, any gradient-driven segmentation algorithm has to contend with noise-induced false edges.

mitochondrion image after the histogram equalization and its corresponding image contours. Figure 3 illustrate the processed image and its contours. We observe that the treated image presents better characteristics for either automated or manual segmentation. Figures 4 and 5 show the final values of the adaptive parameters p^x, p^y, p^{xy} and Λ.

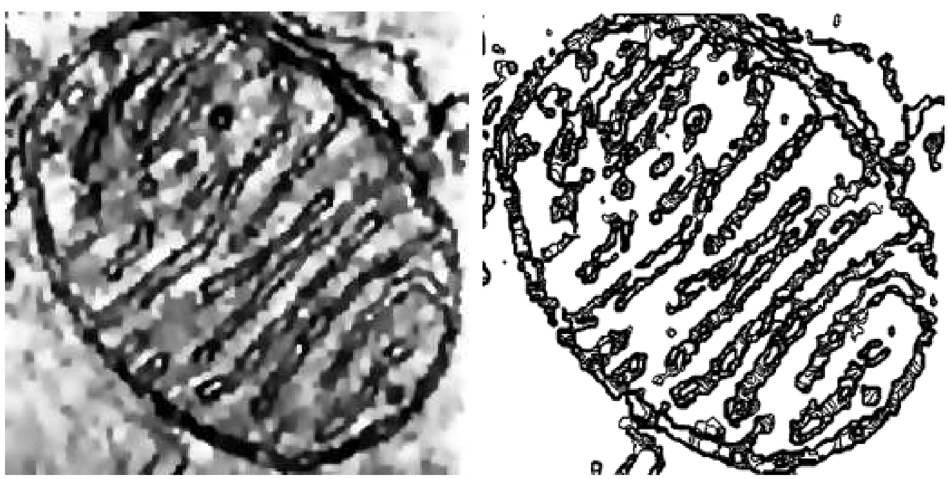

FIGURE 3. (Left) Mitochondrion image after processing by the proposed dynamic model. (Right) Image contours of the image: here, the contours clearly outline the structures of interest.

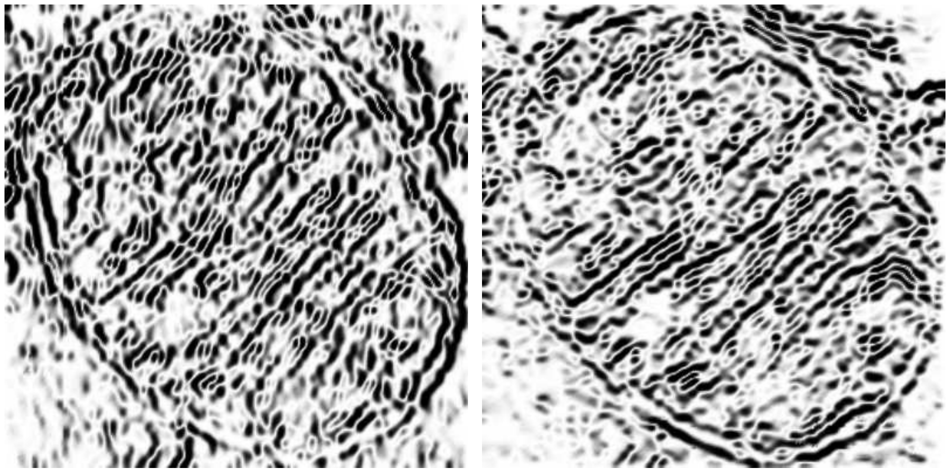

FIGURE 4. (Left) The adaptive parameter p^x, final value. (Right) The adaptive parameter p^y, final value.

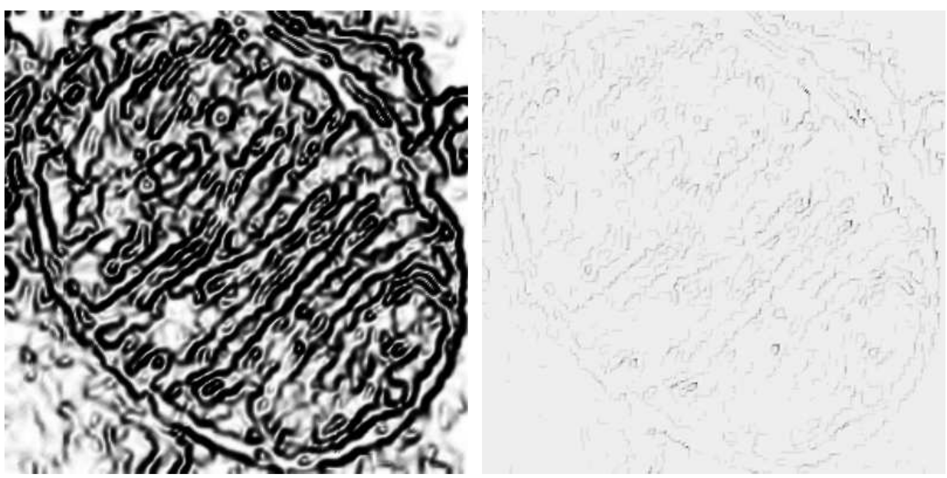

FIGURE 5. (Left) The adaptive parameter p^{xy}, final value. (Right) The adaptive parameter Λ, final value.

5. **Discussion.** Good three-dimensional reconstruction of the structure of mitochondria depends upon good segmentation of either the full three-dimensional tomograms, or hundreds of thin two-dimensional sections of it. This segmentation process usually involves tracing membrane profiles in each of the many parallel image slices either manually or via an automated process. In both cases the quality of the images prior to segmentation plays a fundamental role in the results. Preprocessing the images before segmentation can help obtain better results without losing much information. Our proposed model is based on the well established TV approach for noise removal. The implementation is such that almost no human intervention is necessary making it very attractive for its incorporation into automated systems.

6. **Acknowledgments.** The authors would like to thank Terry Frey and Peter Salamon for providing the electron microscopy images and for their helpful comments. This work has been supported in part by NIH RoadMap Initiative award R90 DK07015.

REFERENCES

[1] R. Acar and C.R. Vogel, Analysis of total variation penalty methods for ill-posed problems. Inverse Problems 10(1994) 1217-1229.

[2] C. Bajaj, Z. Yu and M. Auer, Volumetric feature extraction and visualization of tomographic molecular imaging. J. Struct. Biol. 144(2003) 132143.

[3] P. Blomgren, T. Chan and P. Mulet, Extensions to Total Variation Denoising. Proceedings of SPIE 3162(1997).

[4] A. Chambolle, An Algorithm for Total Variation Minimization and Applications. J. Math. Imaging in Vision 20(2004) 89-97.

[5] Y. Chen, S. Levine and M. Rao, Variable Exponent, Linear Growth Functions in Image Restoration. SIAM J. Appl. Math. (2005).

[6] P. Deuflhard, A modified Newton method for the solution of ill-conditioned systems of nonlinear equations with application to multiple shooting. Numer. Math. 22(1974) 289-315.

[7] T. Frey, G. Perkins, M. Ellisman, Electron Tomography of Membrane-Bound Cellular Organelles. Annu. Rev. Biophys. Biomol. Struct. 35(2006) 199-224.

[8] T.G. Frey, C.W. Renken, and G.A. Perkins, Insight into Mitochondrial Structure and Function from Electron Tomograpgy, Biochimica et Biophysica Acta, 155(1–3):196–203, 2002.

[9] G. Golub, T. Chan and P. Mulet, A nonlinear primal-dual method for total variation-based image restoration. SIAM J. Sci. Comput. 20(1999) 1964-1977.

[10] B. Kawohl, Remarks on the operator $\operatorname{div}(\nabla u/|\nabla u|)$. Contemporary Mathematics 127(1992) 69-83.

[11] A. Marquina and S. Osher, Explicit algorithms for a new time dependent model based on level set motion for a nonlinear debluring and noise removal. SIAM J. Sci. Comput. 22(2000) 387-405.

[12] S. Levine, Y. Chen and J. Stanich, Image Restoration via Nonstandard Diffusion. Duquesne University, Department of Mathematics and Computer Science Technical Report 04-01 (2004).

[13] V. Lucic, F. Foster and W. Baumeister, Structural studies by electron tomography: from cells to molecules. Annu. Rev. Biochem. 74(2005) 833-865.

[14] S. Osher and J.A. Sethian, Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulation. SIAM J. Sci. Comput. 79(1988) 12-49.

[15] G.A. Perkins, C.W. Renken, T.G. Frey, and M.H. Ellisman, Membrane Architecture of Mitochondria in Neurons and the Central Nervous System, Journal of Neuroscience Research, 66(5):857–865, 2001.

[16] G.A. Perkins and T.G. Frey, Recent Structural Insight into Mitochondria Gained by Microscopy, Micron, 31(1):97–111, 2000

[17] G.A. Perkins, J.Y. Song, L. Tarsa, T.H. Deerinck, M.H. Ellisman, and T.G. Frey, Electron Tomography of Mitochondria from Brown Adipocytes Reveals Crista Junctions, Journal of Bioenergetics and Biomembranes, 30(5):431–442, 1998.

[18] G.A. Perkins, C.W. Renken, J.Y. Song, T.G. Frey, S.J. Young, S. Lamont, M.E. Martone, S. Lindsey, and M.H. Ellisman, Electron Tomography of Large, Multicomponent Structures, Journal of Structural Biology, 120(3):219–227, 1997.

[19] D. Ress, M. Harlow, R. Marshall and U. McMahan, Methods for generating high-resolution structural models from electron microscope tomography data. Structure 12(2004) 17631774.

[20] L. Rudin, S. Osher, and E. Fatemi, Nonlinear Total Variation based Noise Removal Algorithm. Physica D 60(1992) 259-268.

[21] B. Song, Topics in Variational PDE Image Segmentation, Inpainting and Denoising. PhD Thesis, University of California Los Angeles (2003).

[22] P. Schults, E. Bollt, R. Chartrand, S. Esedoglu and K. Vixie, Graduated Adaptive Image Denoising: Local Compromise Between Total Variation and Isotropic Diffusion. Preprint, (2005).

[23] D. Strong, Adaptive total variation minimization image restoration. PhD Thesis, University of California Los Angeles (1997).

[24] C.R. Vogel and M.E. Oman, Iterative methods for total variation denoising. SIAM J. Sci. Comput. 17(1996) 227-238.

[25] N. Volkmann, A novel three-dimensional variant of the watershed transform for segmentation of electron density maps. SIAM J. Sci. Comput. J. Struct. Biol. 138(2002) 123129.

Received on June 13, 2007. Revised on June 13, 2007.

E-mail address: `Blomgren@terminus.sdsu.edu`
E-mail address: `Carlos.Bazan@sdsu.edu`