Published in final edited form as:

IEEE Trans Med Imaging. 2012 December ; 31(12): 2241–2252. doi:10.1109/TMI.2012.2214229.

Shape-based Regularization of Electron Tomographic Reconstruction

Ajay Gopinath,

Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712, USA

Guoliang Xu,

Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100190, China

David Ress,

Neurobiology and Imaging Research Center, University of Texas at Austin, Austin, TX 78712, USA

Ozan Öktem,

The Center for Industrial and Applied Mathematics, Department of Mathematics, KTH - Royal Institute of Technology SE-100 44 Stockholm, Sweden

Sriram Subramaniam, and

Laboratory of Cell Biology, National Cancer Institute, NIH, Bethesda, MD 20892, USA

Chandrajit Bajaj

Department of Computer Science and the Institute of Computational Engineering and Sciences, University of Texas at Austin, Austin, TX 78712, USA

Ajay Gopinath: ajay.gopinath@utexas.edu

Abstract

We introduce a tomographic reconstruction method implemented using a shape based regularization technique. Spatial models of known features in the structure being reconstructed are integrated into the reconstruction process as regularizers. Our regularization scheme is driven locally through shape information obtained from segmentation and compared with a known spatial model. We demonstrated our method on tomography data from digital phantoms, simulated data and experimental Electron Tomography (ET) data of virus complexes. Our reconstruction showed reduced blurring and an improvement in the resolution of the reconstructed volume was also measured. This method also produced improved demarcation of spike boundaries in viral membranes when compared with popular techniques like Weighted Back Projection and the Algebraic Reconstruction Technique. Improved ET reconstructions will provide better structure elucidation and improved feature visualization, which can aid in solving key biological issues. Our method can also be generalized to other tomographic modalities.

Copyright (c) 2010 IEEE.

Correspondence to: Ajay Gopinath, ajay.gopinath@utexas.edu.

Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Index Terms

Electron Microscopy; Reconstruction; Tomography; Shape Based Regularization; Bayesian methods

I. Introduction

Electron microscopy is the preferred tool for structural biologists to visualize three-dimensional structures of molecular and cellular complexes *in-situ*. Electron tomography (ET) involves acquiring planar Transmission Electron Microscope (TEM) images of the biological sample from different projection angles (tilt series). Simulating this procedure is called the *forward problem*, and reconstructing a 3D image (or map) from such data using the principles of tomography is called the *inverse problem*. ET is currently the *only* approach that allows one to reconstruct the 3D structure of individual biological complexes in their native state.

The reconstruction from ET data must overcome several challenges. First, TEM images suffer from limited contrast to noise ratio. A second major challenge is that the angle of rotation (tilt angle) cannot exceed $\pm 70^\circ$ (with respect to the horizontal plane) because the beam's path-length through the sample and its supporting structure becomes inappropriate to form projections at higher tilt angles. Hence projection images are available only for a limited range of tilt angles. In Fourier space, the missing tilt images manifests themselves as a missing wedge (Fig. 1), leading to severe instability in the reconstruction problem. Due to the missing data, the problem is ill posed and there is no unique solution to the reconstruction problem. Thus, generation and visualization of models from ET data is a huge challenge.

Our objective is to develop mathematical and computational tools to solve the *inverse problem* in ET through a shape regularization method derived from a Bayesian formulation. We formulated likelihood probability distributions that accurately model the forward problem in ET as well as the main noise characteristic in the measured data. Structural shape information about the specimen being imaged was included into a shape based regularization by means of an energy functional. Reconstruction was performed by calculating the maximum *a posteriori* (MAP) estimate through a modified Expectation Maximization (EM) optimization scheme. In the example presented, we used structural information of spike features of virus complexes that are available through sources like X-ray crystallography, Protein Data Bank (PDB) and Electron Microscopy Data Bank (EMDB). Our reconstruction scheme provides a novel framework for incorporating such *a priori* shape information about local structures into the reconstruction process. The algorithm was tested on a single axis tilt series images of the Simian Immunodeficiency Virus (HIV like retrovirus that causes AIDS in monkeys) interacting with a neutralizing molecule D1D2-IgP (data courtesy of NIH) [1].

The paper is organized as follows: A preliminary discussion about image formation in Transmission Electron Microscopy is presented in Section II. An overview of regularization theory, including methods like Algebraic Reconstruction Technique (ART) and Weighted Back Projection (WBP), which are popular in ET is also provided. Section III discusses the statistical regularization framework as applicable to our method. In Section IV, the shape based regularization scheme is introduced where segmentation is performed to compare local features with a known model. The models used specifically in our reconstruction analysis are described in Section IV-E. Finally, our reconstruction results and analysis are presented in Section

II. Preliminaries

When a high energy electron wave interacts with a biological specimen, it is scattered due to the electrostatic field of the specimen's constituent atoms. The output wave after passing through the specimen contains information about the electrostatic field of the specimen which is captured as image contrast. The main contrast mechanism in TEM imaging of unstained biological samples is phase contrast (interference contrast), which results from the quantum superposition (interference) of the crests of the incident wave and the scattered wave. This can often be approximately described by projecting the potential of the specimen. Typically, thinner specimens with higher accelerating voltage and lighter atoms will satisfy this projection approximation [2]. The emergent electron wave interacts with the microscope optics to generate a 2D projection image. The specimen is tilted and several such 2D projection images are acquired.

The 2D projection images in TEM typically suffer from a number of degrading factors. The images acquired have spatial resolution on the order of Angstroms (Å), but slight mechanical noise during tilting and acquisition can cause large misalignment in the images. The projection assumption used above for the electron beam-specimen interaction only holds for weakly scattering specimens. These factors combined with the missing wedge problem described earlier degrade the accuracy of 3D reconstructions in ET.

The inverse problem in ET is to reconstruct a 3D image based on images acquired within a tilt series. Mathematically this can roughly be stated as the problem of recovering an estimate of a signal f (a function representing the 3D image) given measured data g (the vector representing data in the tilt series) obtained from indirect observations of the signal. The precise mathematical formulation depends on whether one takes the statistical or classical approach for regularization. All regularization methods, both classical and statistical, are based on the principle of replacing the original ill-posed inverse problem with a neighboring well-posed problem. This is done by making use of a prior model and a data model.

The projection images collected are denoted by

$$I_{\theta}(x, y): 0 \leq x < X, 0 \leq y < Y, -70^{\circ} \leq \theta \leq +70^{\circ}.$$

The size of each projection image X, Y is determined by the microscope hardware and the range of the tilt angle θ is set by the biologist. The forward model for ET is derived in [3] where it is shown that under certain conditions, one can assume that the tilt series images correspond to a simple forward projection transform (x-ray transform) of the specimen. Some of these conditions are: validity of the first order Born approximation (for weakly scattering specimens), thin specimen width-allowing electrons to pass through and a phase contrast imaging model. Assuming the forward projection transform, if the function f representing the 3D image is discretized and represented by a vector, typically by taking voxel values, then the forward model can be written as a projection matrix H . For noise-free data

$$Hf = g \quad (1)$$

where g is the data from the projection images I_{θ} . The data model for ET is derived in [4] and is quite complex. But the dominating contribution to the noise is from the shot noise which is Poisson distributed.

In statistical regularization, and in particular Bayesian inference, the philosophy is to recast the inverse problem in the form of a statistical quest for information. While classical regularization seeks to recover a single estimate of the unknown signal, statistical regularization is based on recovering the probability distribution of the signal.

A. Classical regularization methods

These are commonly subdivided into *analytical*, *iterative*, and *variational* methods.

1) Analytical reconstruction methods—The starting point is to consider the operator equation

$$H(f)=g$$

where the data is considered as continuum. The idea is to directly approximate H^{-1} and appropriately discretize it. Doing this in a manner that preserves stability depends highly on the specific properties of the operator H and the data collection geometry (the interpretation of the discretization of g). For the case when H represents a simple forward projection transform, we have the well-known Filtered Back Projection (FBP) method. Weighted Back Projection (WBP) reconstruction is another analytic method that is popular in ET [5], [6].

2) Iterative methods—Here the idea is to generate an iterative sequence that converges to a least-squares solution of $H(f) = g$ [7]. The most popular iterative method in ET is the Kaczmarz method, also called Algebraic Reconstruction Technique (ART) [8]. In this technique each tomogram image g_i is treated like a hyperplane in a N dimensional space. A solution to the entire equation system for all tomograms g is a point in the N dimensional space where all the hyperplanes intersect [8].

3) Variational methods—The idea here is to reformulate the reconstruction problem as an optimization problem:

$$\arg \min_f \lambda S(f) + D(H(f), g) \quad (2)$$

In the above, the functional $f \mapsto S(f)$ is the regularization functional that encodes a priori information about the true unknown signal, D is the data discrepancy functional, and $\lambda > 0$ is a weighting parameter.

A common approach in many imaging modalities is to select $S(f) := \|\nabla f\|_2$, i.e. as the L_2 -norm of the gradient magnitude of f (Dirichlet energy).

Total Variation (TV) and related methods are applied in [9] to the reconstruction problem in ET with promising results. Higher order regularizers have also been applied to ET. Dey et al. [10] have used a TV based formulation to perform confocal microscope deconvolution.

A recent method introduced by Xu et al. [11] [12] reconstructs f by solving (2) with

$$S(f) := \lambda \int u(\sigma(f)(x), \kappa(f)(x)) |\nabla f(x)| dx,$$

and

$$D(H(f), g) := \sum_i \int (h_i f(x) - g_i)^2 dx$$

where $\sigma(f)(x)$ and $\kappa(f)(x)$ are the mean and Gaussian curvatures of the level-set surface of f at a point x , and u is a user specified function. The second term is the data fidelity term while the first term is a geometric regularizer. In [11], the solution of (2) with S and D as above is obtained by solving an L^2 -gradient flow (L2GF) using a calculus of variations optimization method.

B. Statistical regularization methods

A statistical approach for solving inverse problems using a Bayesian framework has been used before for a variety of reconstruction problems. Under a Bayesian framework, the prior information is encoded by a probability distribution and is called the *a priori* (or simply, prior) [13].

Bayes rule gives us the relation

$$p(f|g) = \frac{p(g|f)p(f)}{p(g)}, \quad (3)$$

where $p(f)$ is the prior about f , $p(g|f)$ is the likelihood term that models the physics of image acquisition. $p(g)$ is the marginal probability distribution that is typically ignored when using estimators since it does not vary with f [14] [13].

Various priors have been used by different groups depending on the imaging model and application. In medical imaging, Gindi et al. [14] have used edge maps of anatomical images as priors while performing reconstruction of functional images. Their goal was to induce edges and hence better contrast in functional images (PET, SPECT). The anatomical image (typically CT, MRI) is registered with the functional image (typically PET, SPECT). An edge map is created using the anatomical image which is used as a prior while performing reconstruction of the functional image. Gindi et al. model both the likelihood and prior as Gibbs distribution. A predetermined edge map is built from a CT or MRI image and is incorporated in the *a priori* as an energy term. Anatomical information from MRI and CT have similarly been used in Hsiao et al. [15] and Rangarajan et al. [16].

Methods that use global smoothness priors do not incorporate structure information in the reconstruction process. The assumption of smoothness can be arbitrary and can affect edges. Methods like Gindi et al. [14] and Hsiao et al. [15] that use CT/MRI image as priors for PET reconstruction have issues where edges seen in a CT/MRI image are not always seen in a PET image leading to inconsistent features. Yoon et al. [17] have introduced a joint segmentation-reconstruction method. A level set segmentation and reconstruction is shown with a piecewise constant intensity function for PET and CT. Unlike ours, no explicit spatial model is used, instead a segmentation algorithm partitions the image into piecewise continuous regions during reconstruction. A major limitation of Yoon et al. [17] compared to our method is that their algorithm has not been tested on real data and has been demonstrated only on phantoms, on which segmentation is easier to perform. Our method is validated on highly noisy ET data and a quantitative improvement in reconstruction has been shown for real data.

III. Statistical Regularization

The statistical regularization framework allows us to use shape based priors explicitly in the reconstruction process. Also, the physical process of data collection can be modeled through the likelihood term.

A. Likelihood probability distribution

A Poisson distribution is used in a system that has a high number of events of which only a small percentage are counted. In the context of ET, a large number of electrons are bombarded on the specimen and a large fraction is scattered from the small collection aperture of the TEM creating contrast. As already mentioned, the dominating contribution to the noise is shot noise which is Poisson distributed. Other modalities like SPECT and PET also use a Poisson distribution to model noise arising during the image acquisition [18], [19], [14]. Under this assumption, the likelihood becomes

$$p(g|f) = \prod_m^M \frac{(Hf)_m^{g_m} \exp(-(Hf)_m)}{g_m!}, \quad (4)$$

where M is the total number of projection measurements, i.e. size of a 2D projection image times the number of projections, g_m is the m^{th} projection measurement and $(Hf)_m$ is the corresponding projection of the current reconstruction estimate f using the forward operator H .

B. Prior probability distribution

We model the prior $p(f)$ as a Gibbs distribution:

$$p(f) := \exp\left(\frac{-\lambda S(f)}{Z}\right), \quad (5)$$

where $S(f)$ is a shape based energy functional, Z is a normalization factor and λ is a weighting parameter. Gibbs distributions are widely used in image restoration problems and have a number of advantages. They can enforce smoothness in images as well as model images containing piecewise homogeneous regions [20]. They also satisfy the Markov property, i.e. only neighboring voxels have direct interactions with each other (Markov-Gibbs equivalence).

C. MAP estimate and optimization

Maximum-a-posteriori (MAP) is one of the most popular estimators used in statistical regularization [13] [14]. The MAP estimator is defined as the element f that maximizes the posteriori $p(f/g)$, alternatively minimizes the negative log of the *a posteriori*. An optimization step is involved in calculating the MAP estimate. Gindi et al. [14] use a EM-ML optimization step to solve the objective function resulting from their MAP estimate. Hsiao et al. [15] use deterministic annealing for optimizing the MAP. Their resulting objective function is not convex since they use a Poisson likelihood probability distribution and a Gamma prior probability distribution.

Here, we combine a Poisson likelihood assumption with a Gibbs prior to perform a segmentation-based optimization with an EM algorithm. The MAP estimator is given as the solution to:

$$f_{\text{MAP}} := \arg \max_f p(f|g), \quad (6)$$

We define a cost function $c(f)$ as:

$$c(f) := \log(p(g|f)) + \log(p(f)), \quad (7)$$

where $\log(p(g|f))$ is the log likelihood and $\log(p(f))$ acts as a regularizer. $\log(p(f))$ is minimum when the energy term $S(f)$ is minimum. Using Bayes theorem with the prior in (5), the MAP estimator given in (6) can be written as (2).

It can be shown that under certain circumstances, (2) can be solved using iterates from a penalized EM method [21] [22].

$$f^{(i+1)} := \frac{f^{(i)}}{1 + \lambda |\nabla S(f^{(i)})|} H^T \frac{g}{H f^{(i)}}, \quad (8)$$

where $S(f^{(i)})$ acts as the regularizer, λ is the regularization parameter, and H^T is the adjoint of the linear forward operator, which in tomography is the back-projection operator. Note that the division above between functions is taken point wise, and if functions are represented by vectors (discretized), it is taken component wise.

IV. Shape Based Regularization

We introduce a regularization approach in the framework of statistical regularization, where the energy term $S(f)$ encodes local information about the structure being reconstructed f . Spatial models M of local structures (virus spikes in our example that follows) are built using PDB and EMDB information (see Section IV-E). Models can also be built using a composite segmentation technique [23] where virus features segmented from different reconstructions are aligned together and combined through a voting scheme to provide a high confidence model. In our reconstruction method, at each iteration virus spikes of the current reconstruction are segmented creating a binary volume (0 or 1) that is compared with the known spatial model of the structure. The regularizer applied at a voxel in f_j is chosen to be either a penalizing or an increasing term depending on whether the voxel neighborhood is determined to be over-segmented or under-segmented, respectively.

A. Local comparison of sets and under- & over-segmentation

For a given measurable domain $U \subset \mathbb{R}^3$, let $\chi_U: \mathbb{R}^3 \rightarrow \{0, 1\}$ denote an indicator function of U , i.e

$$\chi_U(x) := 1 \text{ if } x \in U \text{ and } 0 \text{ otherwise.}$$

Let $|U|$ denote the Lebesgue measure of U , so

$$|U| := \int_{\mathbb{R}^3} \chi_U(x) dx.$$

Finally, let $B_r(x)$ denote the ball in $\mathbb{R}^3$ with center at x and radius $r > 0$.

Consider two domains $M, U \subset \mathbb{R}^3$ and let $r > 0$ be fixed. For $x \in \mathbb{R}^3$, define

$$\Delta_r(U, M)(x) := \frac{|\{y \in B_r(x) : \chi_M(y) \neq \chi_U(y)\}|}{|B_r(x)|} \quad (9)$$

Hence,

$$0 \leq \Delta_r(U, M)(x) \leq 1 \quad \text{for all } x \in \mathbb{R}^3 \text{ and all } r \geq 0$$

which measures the relative portion of a local r -neighborhood of x where the sets U and M agree. To introduce the concept of under- & over-segmentation, consider $x \in \mathbb{R}^3$ with $r > 0$ fixed and $\varepsilon > 0$. One of the sets, M , is considered as the given *shape model*. Then, U is *under-segmented* at x compared to M if

$$\chi_M(x) > \chi_U(x) \text{ and } |U \cap B_r(x)| > \varepsilon.$$

Correspondingly, U is *over-segmented* at x compared to M if

$$\chi_M(x) < \chi_U(x) \text{ and } |U \cap B_r(x)| > \varepsilon.$$

The parameter $\varepsilon > 0$ is a tolerance level. We define

$$s_{r,\varepsilon}(U, M): \mathbb{R}^3 \rightarrow \{-1, 1\}$$

as

$$s_{r,\varepsilon}(U, M)(x) := \begin{cases} 1 & \text{if } U \text{ is over-segmented at } x \text{ compared to } M \\ -1 & \text{if } U \text{ is under-segmented at } x \text{ compared to } M. \end{cases}$$

In our discretized implementation, images are represented by voxels where each voxel has an associated center point denoted by x_j . $B_r(x_j)$ is given by a $3 \times 3 \times 3$ voxel neighborhood around x_j and denoted by $N(j)$. We also choose $\varepsilon > 0$ so that it equals the volume of a single voxel. Over-segmentation at a voxel is detected when the current voxel of the segmented region U has a 1 while the model M has a 0 and at least one neighbor of the current voxel in the segmented region has a 1. Over segmentation generally occurs because of insufficient gradient at a nearby feature boundary, often caused by blurring or distortion of features that is typical of limited-angle tomography reconstruction. Under-segmentation is detected when the current voxel of the segmented region has a 0 while the model has a 1 and at least one neighbor of the current voxel in the segmented region has a 1. Under segmentation occurs due to low intensity value of the voxel. Figs. 2 and 3 illustrates over- and under-segmentation in the 2D case.

B. Energy functional for local segmentation information

Given a model M and a segmentation region U (both domains of $\mathbb{R}^3$), the energy functional is defined as:

$$S(f, U) := \int_{\mathbb{R}^3} (1 + \gamma_{r,e}(U, M)(x)) \|\nabla f(x)\|_2 dx \quad (10)$$

where $\gamma_{r,e}(U, M)(x) := s_{r,e}(U, M)\Delta_r(U, M)(x)$.

C. Reconstruction scheme

Our reconstruction method is an intertwined iterative scheme given as:

$$\begin{cases} U^{(i)} \subset \mathbb{R}^3 \text{ is segmented based on } f^{(i)} \\ f^{(i+1)} := \arg \min_f \lambda S(f, U^{(i)}) + D(H(f), g). \end{cases} \quad (11)$$

The segmentation method used to get $U^{(i)}$ from $f^{(i)}$ is described in Section IV-D. This is an iterated regularization method (i.e. solving a sequence of regularizations) where the iterate updating f provides the MAP estimate using the prior probability distribution in (5) with S as in (10). The segmentation region $U^{(i)}$ is updated while the model M is fixed. When convergent, the set $U^{(i)}$ will not change and the entire iterative sequence will tend to a MAP estimator.

The inner iterative scheme based on (8) is discretized to calculate the MAP estimate $f^{(i+1)}$. The reconstruction update at each voxel is given as:

$$f_j^{(i+1)} := \frac{f_j^{(i)}}{1 + \lambda |\partial_j S_j(f_j^{(i)}, U^{(i)})|} \left[H^T \frac{g}{H f_j^{(i)}} \right]_j \quad (12)$$

where

$$S_j(f_j, U) := (1 + \gamma_j(U, M)) \sum_{k \in N(j)} \frac{|f_j - f_k|}{d_{j,k}} \quad (13)$$

is the resulting energy functional at voxel j . Here, $d_{j,k}$ denotes the Euclidian distance between the midpoints x_j and x_k of the j th and k th voxels, and $\gamma_j(U, M)$ is the discretization of $\gamma_{r,e}(U, M)(x_j)$ where the continuum neighborhood $B_r(x_j)$ is replaced by the voxel neighborhood $N(j)$ and β is chosen so that it corresponds to the volume of one voxel. The term $\Delta_r(U, M)(x)$ depends on local neighborhood matches between U and M ; more matches imply higher confidence. Depending on whether over or under segmentation is detected, the term $s_{r,e}(U, M)(x)$ controls the effect of $\gamma_{r,e}(U, M)(x_j)$. This attempts to reduce blurring or distortion by controlling the effect of the regularization locally in the reconstruction. Fig. 5 shows a flow diagram of the overall reconstruction algorithm.

The penalized EM optimization method was found to be superior when compared to a Conjugate Gradient optimization approach; in that good reconstructions were obtained in fewer iterations. For a stopping criterion we calculated the root mean square (RMS) difference between the current reconstruction and the previous iteration reconstruction, normalized over an intensity range between 0 and 1. If the overall RMS difference between iterations is lower than 0.07 for two consecutive iterations, the iterations were terminated. The reconstruction was executed for at least 10 iterations before this check was performed.

The RMS differences between reconstructed volumes of consecutive iterations were examined for our datasets over several reconstruction runs. Most reconstructions, after an initial set of iterations, would enter a plateau with the reconstructed volume being unchanged over several iterations. At this stage there was only a small variation in RMS value between iterations and further iterations did not provide any additional improvement in the reconstruction volume. This was an indication of convergence and the RMS threshold number (0.07) that best captured this was used for our reconstructions. The cost function $\alpha(f)$ (equation 7) was plotted at each iteration along with the RMS difference (see Fig. 4). The cost function decreases until it reaches a plateau around iteration 20. This has been shown for the Shepp Logan phantom (see Section V). The RMS difference also reaches a plateau around the same iteration. The decreasing RMS and cost function plots indicate that the algorithm is converging. The value was determined empirically using phantom reconstructions (range of 0.0001 to 0.0003). If the location and shape of the model in the reconstruction has a low confidence, a low value of lambda should be used.

D. Segmentation

The segmentation method described in [24] was used to segment a region of interest U from the reconstructed 3D volume f . The method is a variant of the fast marching algorithm developed by Sethian [25]. The level set evolution of the advancing wave front is given by:

$$\varphi_t + \mathbf{F}|\nabla\varphi| = 0 \quad (14)$$

φ is the level set function and $\mathbf{F}$ is a numerical force field. Two point classes are seeded on f , one for the foreground feature and another for the background. These seeds are treated as initial snakes or contours whose movements are directed by an external force field determined by the normalized gradient vector field [24], [26]. When the snakes stop evolving, a segmented volume U is obtained. Using an initial back projection reconstruction, without any regularizers, the seed points for features corresponding to the known model are provided by the user. These points are used at every iteration of the reconstruction to obtain a new segmentation. We find that this algorithm is adequate for our objective of segmenting features like virus spike regions robustly with acceptable variability. We also obtained similar segmentations through a combination of thresholding and connected component analysis. Since the segmentation is performed on small local regions of features, the reconstruction process was fairly invariant to the choice of segmentation technique.

E. Biological spatial models

We used spatial models of the AIDS virus spike in our regularization step. The spike is a mushroom-shaped structure with tapering stem that is attached to the outer envelope of the virus. Using X-ray crystallography, biochemists have been able to obtain accurate atomic level information of the spike proteins of the AIDS virus. TEM based volumetric models of the spikes have been generated by averaging several thousand images of the virus spike [27] (Fig. 6). This data is available in the Protein Data Bank (PDB) and the Electron Microscopy Data Bank (EMDB). The spike is the primary target for drug design as it is responsible for infecting the host cell. Accurate reconstructions of the regions surrounding the spikes can aid in understanding the process with which the virus infects the host cells.

V. Results and Evaluation

Our reconstruction method was evaluated on projection simulations from a 3D Shepp Logan digital phantom (Fig. 7), a TEM simulator and experimental tilt series data from a TEM. Our shape based regularization (Shape Reg) reconstruction output was compared with reconstructions from four different methods, WBP, ART, a L^2 gradient flow based classical

regularization method (L2GF) (described in Section II-A3 [11]), a Maximum Likelihood approach, which is our current EM implementation without the prior term (ML-EM) and a Total Variation regularization (TV Reg) method [10].

The spike model is aligned to match the spikes visible in the initial tomographic reconstruction volume (a back projection reconstruction, without any regularizers). The alignment, currently done through user supervision, needs to be done only once. Manual alignment of the Shepp-Logan phantom was based on the center of the spherical structure used as a model for the reconstruction. For the spike, the junction of the spike head and the stem and the junction of the spike stem with the spherical envelope, were used for the alignment. Gopinath (author on this paper) and Bovik [28] have recently implemented a spike detection method that extracts the location and orientation of spikes automatically. This can be used for aligning the spike model with the spike location during reconstruction. The regularization is performed over subsequent iterations in a sub-volume region containing the spike and compared with the aligned model.

A. Digital phantom

2D projection images of a $128 \times 128 \times 128$ 3D Shepp Logan phantom (Fig. 8) were simulated between tilt angles $\pm 69^\circ$ in steps of 2° . The spherical region (region ① in Fig. 8) in the phantom was used as the known spatial model M in our Shape Reg reconstruction method. 2D central cross sections were used to compare Shape Reg with WBP, ART, L2GF, TV Reg and ML-EM reconstructions (Fig. 8).

1) Spatial error—1D intensity profiles of the reconstructions are plotted in Fig. 8. This 1D profile is across the Y axis through the central X and Z coordinate as illustrated in the line through the corresponding 2D central slice shown in Fig. 8. Three regions of interest are marked as ①, ② and ③. Regularization for the Shape Reg method was provided for region ①. Shape Reg reconstruction intensity profile shows the best match to the profile of the phantom. In the region ① where the regularization was applied, the Shape Reg expectedly performs well (Fig. 8). It shows sharp edges without loss of structure fidelity. WBP's intensity profile suffers from noise and artifacts in all the regions. ART and ML-EM show a higher intensity curve that smooths the edges. L2GF performs best at region ② but both L2GF and ART show reduced intensity in the first peak in region ③. It is interesting to note that regions where regularization was *not* applied (region ③ and region ② fig 8) also show improved reconstruction in Shape Reg when compared with ML-EM. This implies that providing a local regularization has a global effect. The optimization step has a forward projection Hf and a back-projection H^T operator. Corrections that are made in local regions of f will affect the entire volume f due to these operators.

Mean Squared Error (MSE) was measured through a central region (2048 voxels) in each of the reconstructed volumes. The intensity range for each of the reconstruction volumes was normalized to 0 – 255. The error for Shape Reg was the lowest at 489.7, while that of L2GF, TV Reg and ML-EM were at 905.8, 1684 and 746 respectively. WBP and ART showed MSE values of 2302 and 3150.

2) Fourier spectrum plots—To analyze the effect in frequency (resolution) space, we compared the power at various resolutions. The reconstructed volumes were Fourier transformed in 3D, and the radial power spectrum was calculated. The power spectrum for the Shape Reg was compared with that of the original Shepp Logan phantom, see Fig. 9. All methods showed some amount of blurring compared to the original spectrum, but the Shape Reg's Fourier spectrum best matched the original Shepp Logan's spectrum. L2GF and ART exhibited the greatest amount of blurring.

3) Robustness: Reconstruction with misaligned model—We translated the region where regularization is applied while maintaining the same spherical model (figure 10). The underlying region on which regularization is performed does not match the spherical model anymore. Also, the weight, λ , was higher than what was used in section V-A at 0.0004 (also see Section IV-C). The resulting reconstructed volume did not have any prominent artifacts. In figure 10(b), a feeble edge was seen where the boundary of the sphere model overlaps the actual sphere in the reconstructed volume. At this region, the regularization process detects over-segmentation and compensates by trying to reduce the intensity resulting in a false edge. The overall reconstruction quality appears to be good with no major feature artifacts. While accurate model alignment is necessary for good reconstructions, misalignments do not result in large artifacts.

B. Electron Tomography Simulation

We used a TEM simulator [4] to simulate tilt series projections of a group RNA polymerase particles. Projections were simulated with and without TEM noise between $\pm 60^\circ$ in steps of 2° . Shape Reg reconstruction was performed and the output volume was compared with the original RNA polymerase phantom volume. Reconstruction with a misaligned prior model was also done to evaluate robustness.

1) Without Noise—Reconstructions were performed without TEM noise using the Shape Reg method, where a part of the original RNA polymerase was used as a regularizer. The reconstructions were compared with the original phantom and the mean squared error (MSE) measured for the central RNA polymerase particle was 695.61.

2) With Noise—The Shape Reg reconstructions performed above were repeated with TEM noise in the tilt series projection simulations. The MSE for the reconstructed volume was 2101.77. The images were severely corrupted with noise but the structure was still discernible (Fig. 11(e)).

3) Rotated Spatial Model—To evaluate the robustness of the Shape Reg method to misalignments, we rotated the prior model by 15° about its center. The Shape Reg reconstruction was then compared with the Shape Reg reconstruction obtained previously without misalignment. Visually the Shape Reg reconstruction with misalignment appears good with no artifacts that can be attributed to the misalignment. The MSE for Shape Reg with misalignment was only slightly higher at 704.15 compared with the MSE for Shape Reg without misalignment at 695.61. This indicates that Shape Reg reconstruction is robust under small misalignments.

C. Electron tomography data

Tilt series projection images of the Simian immunodeficiency virus (similar to the HIV, AIDS virus) particles interacting with D1D2IgP neutralizing drug were acquired between angular tilts $\pm 69^\circ$ [1] using a TEM. Shape Reg reconstruction was performed using spatial models of the spike region as regularizers (see Section IV-E). The model was applied to eight spike regions whose locations were identified by a knowledgeable user in an initial reconstruction performed by a simple back-projection. The resulting Shape Reg reconstruction, shown in Figs. 12 and 13, was compared with WBP, ART, L2GF and reconstruction with no regularization (ML-EM). The spikes looked visually more distinct with better spatial differentiation between neighboring spikes. Blurring, a typical ailment of limited angle tomography, was substantially reduced in the Shape Reg method. In our regularization process we used a segmentation step where blur in these regions resulted in over segmentation that was detected and corrected.

The neighboring regions of the spike (D1D2IgP in Fig. 12) also showed an improvement though no regularization was explicitly provided here. A spike region in a neighboring virus also showed improvement in its overall structure when compared to WBP, ART, ML-EM and L2GF. As we saw in the digital phantom experiment, regularization at one region improved the reconstruction in nearby regions.

1) Resolution through Fourier Shell Correlation (FSC)—The FSC method is a common approach to estimate the resolution and quality of reconstructed volumes in ET [2] [29]. To calculate FSC, the projection images were split into two sets of alternating projection images. Reconstruction was performed on the two sets independently. The resulting reconstruction outputs were then correlated in Fourier space. The spatial frequency at which the correlation value drops to 0.5 of its maximum was estimated as the resolution (Fig. 14). The resolution for ML-EM, ART, WBP, and L2GF was measured to be around 31Å, 28 Å, 30 Å and 28Å respectively, while the Shape Reg method measured at 22Å.

2) Contrast to Noise Ratio (CNR)—The CNR was measured at various locations (Table I) of the reconstruction volumes.

$$\text{CNR} = \frac{\mu_{\text{foreground}} - \mu_{\text{background}}}{\sigma_{\text{background}}}, \quad (15)$$

where $\mu_{\text{foreground}}$ is the average intensity on a feature and $\mu_{\text{background}}$ is the average intensity in the adjoining background region. $\sigma_{\text{background}}$ is the standard deviation in the background region. The CNR was measured at features like virus envelope region and spikes including those on which Shape based regularization was applied. Overall, Shape Reg gave higher CNR values. Locations 5, 6, 7 (Table I) are spike regions where regularization was applied in Shape Reg reconstruction. The regularization decreased blurring in Shape Reg, especially between neighboring spikes leading to significantly higher CNR values. Other methods blur the region between spikes resulting in decreased CNR. We performed a paired difference test of the CNR values and computed the p value. Since the CNR values follows a non-normal distribution we performed two sided Wilcoxon T tests (also called signed-rank test) between CNR values obtained for Shape Reg and other reconstructions method's CNR values. A $p < 0.02$ was obtained for all the comparisons indicating that the CNR improvement seen in Shape Reg is statistically significant.

VI. Conclusion

We presented a method to include known spatial models of critical small structures in a tomographic reconstruction process. This is done through a regularization step where spatial models are introduced through a prior term. Our technique does not impose the spatial model directly but influences the shape of the features by varying the prior function through an intermediate segmentation step. Typical regularization approaches assume a global prior characteristic like smoothness, which may not be valid in regions with very small features. Our regularization, driven locally by determining local shape information through segmentation, ensures accurate reconstruction and reduces blurring that is typical in limited angle tomography. The local regularizers also improves the reconstruction in neighboring regions that do not have regularization applied, as seen in our Shepp Logan phantom experiments, and in the improved resolution of the virus data. The application of local regularizers appears to provide a global improvement of spatial resolution, as illustrated by the Fourier power spectrum and confirmed by Fourier Shell Correlation analysis. It also yields statistically significant improvement in contrast to noise ratios when compared with other reconstruction methods. Our method also leads to improved visualization of features,

which is important in the context of sub-volume averaging [30]. Commonly used methods in ET, such as WBP, can lead to erroneous selection of sub-volume features causing poor or even flawed results. Accurate reconstructions from our method can further aid in solving biological issues involving drug discovery and design.

Future work will include a detailed convergence analysis of the reconstruction method. An additional line search step that can lead to better convergence will be considered. We will explore extensions such as De Pierros method [31] for our reconstruction. A limitation of the current method is that the user needs to provide the location of the region on which to apply the model. To overcome this for the AIDS virus data, Gopinath and Bovik [28] have a fully automated method that detects the location of each spike and identifies a sub-volume around a virus spike. This method was developed to automatically generate shape models of the spike by extracting hundreds of spike volumes automatically. Used in conjunction with this reconstruction method, it can create a powerful reconstruction-segmentation-model generation framework. However, other kinds of tomographic data may require user intervention or a feature detection algorithm specific to that data to apply the shape model. Another limitation is that we do not account for scaling and non-rigid distortions of the model. Robustness tests have been performed by applying a translated shape model, a more extensive analysis will be part of our future work.

Our method can further be useful in the reconstruction of tilt series of stained tissue samples that utilize small gold beads to facilitate alignment. The spherical shape of these beads can provide a model to facilitate global improvement of the reconstruction. While imaging a biological complex, neighboring structures or certain substructures whose models are known can be used as local regularizers. This will reduce the blurring effects in those local regions and improve the overall reconstruction of the biological complex. Our approach can be further improved by incorporating probabilistic shapes and flexible fitting models as priors.

WBP and ART reconstruction was performed using the software package XMIPP [32]. The remaining reconstruction and visualization methods used to produce the figures are encapsulated in our freely available tool Volume Rover 2.0 [33] <http://cvcweb.ices.utexas.edu/cvcwp/>

Acknowledgments

A.G. and C.B. were supported in part by NIH contracts R01-EB00487, R01-GM074258, and a grant from the UT-Portugal CoLab project.

References

1. Bennett A, Liu J, Ryk DV, Bliss D, Arthos J, Henderson RM, Subramaniam S. Cryoelectron tomographic analysis of an HIV-neutralizing protein and its complex with native viral gp120. *The Journal of Biological Chemistry*. 2007; 282:27754–27759. [PubMed: 17599917]
2. Frank, J. *Electron tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*. Vol. ch 3. Springer; 2006. p. 10
3. Fanelli D, Öktem O. Electron tomography: a short overview with an emphasis on the absorption potential model for the forward problem. *Inverse Problems*. 2008; 24(013001)
4. Rullgård H, Öfverstedt LG, Masich S, Daneholt B, Öktem O. Simulation of transmission electron microscope images of biological specimens. *Journal of Microscopy*. 2011; 243(3):243–256.
5. Herman, G. *Image Reconstructions from Projections: The Fundamentals of Computerized Tomography*. Vol. ch 8. Computer Science and Applied Mathematics; 1980.
6. Radermacher, M. *Electron tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*. Vol. ch 8. Springer; 2006.
7. Engl, HW.; Hanke, M.; Neubauer, A. *Regularization of Inverse Problems*. Springer; 2000.

8. Kak, A.; Slaney, M. Principles of Computerized Tomographic Imaging. Vol. ch 7. IEEE Press; 1988.
9. Rullgård H. A new principle for choosing regularization parameter in certain inverse problems. Apr. 2008 arXiv:0803.3713v2 [math.NA].
10. Dey N, Blanc-Feraud L, Zimmer C, Roux P, Kam Z, Olivo-Marin JC, Zerubia J. Richardson–Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microscopy Research and Technique*. 2006; 69:260–266. [PubMed: 16586486]
11. Xu G, Li M, Gopinath A, Bajaj C. Computational inversion of electron tomography images using L^2 -gradient flows. *Journal of Computational Mathematics*. 2011; 29(5):501–525.
12. Li M, Xu G, Sorzano CO, Sun F, Bajaj CL. Single-particle reconstruction using L^2 -gradient flow. *Journal of Structural Biology*. 2011; 176(3):259–267. [PubMed: 21864687]
13. Kaipio, J.; Somersalo, E. Statistical and Computational Inverse Problems. Vol. ch 3. Springer-Verlag; 2005.
14. Gindi G, Lee M, Rangarajan A, Zubal IG. Bayesian reconstruction of functional images using anatomical information as priors. *IEEE Transactions on Medical Imaging*. 1993; 12(4):670–680. [PubMed: 18218461]
15. Hsiao I-T, Rangarajan A, Gindi G. Bayesian image reconstruction for transmission tomography using deterministic annealing. *Journal of Electronic Imaging*. 2003:7–16.
16. Rangarajan A, Hsiao IT, Gindi G. A Bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction. *Journal of Mathematical Imaging and Vision*. 2000; 12:199–217.
17. Yoon S, Pineda AR, Fahrig R. Simultaneous segmentation and reconstruction: A level set method approach for limited view computed tomography. *Medical Physics*. 2010; 37(2329)
18. Gourion D, Noll D. The inverse problem of emission tomography. *Inverse Problems*. 2002; 18:1435–1460.
19. Zanella R, Boccacci P, Zanni L, Bertero M. Efficient gradient projection methods for edge-preserving removal of Poisson noise. *Inverse Problems*. 2009; 25(045010)
20. Levitan E, Chan M, Herman GT. Image-modeling Gibbs priors. *Graphical Models and Image Processing*. 1995; 57(2):117–130.
21. Natterer, F.; Wübbeling, F. Mathematical Methods in Image Reconstruction. Vol. ch 5. SIAM; 2001.
22. Green PJ. On use of the EM algorithm for penalized likelihood estimation. *Journal of Royal Statistical Society*. 1990; 52(3):443–452.
23. Gopinath, A. Reconstruction comparison and a composite segmentation method for electron tomography. *IEEE International Conference on Image Processing*; 2010. p. 585–588.
24. Yu Z, Bajaj C. Computational approaches for automatic structural analysis of large bio-molecular complexes. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2007; 5(4): 568–582. [PubMed: 18989044]
25. Sethian JA. A marching level-set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences USA*. 1996; 93(4):1591–1595.
26. Bajaj C, Yu Z, Auer M. Volumetric feature extraction and visualization of tomographic molecular imaging. *Journal of Structural Biology*. 2003; 144:132–143. [PubMed: 14643216]
27. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S. Molecular architecture of native HIV-1 gp120 trimers. *Nature*. Sep.2008 455:109–113. [PubMed: 18668044]
28. Gopinath A, Bovik AC. Automatic feature extraction and statistical shape model of the AIDS virus spike. *IEEE Transactions on Biomedical Engineering*. 2012
29. van Heel M, Schatz M. Fourier shell correlation threshold criteria. *Journal of Structural Biology*. 2005; 151(3):250–262. [PubMed: 16125414]
30. Yu Z, Frangakis AS. Classification of electron sub-tomograms with neural networks and its application to template-matching. *Journal of Structural Biology*. 2011; 174(3):494–504. [PubMed: 21382496]

31. Pierro ARD. On the convergence of an EM-type algorithm for penalized likelihood estimation in emission tomography. *IEEE Transactions on Medical Imaging*. 1995; 14(4):762–765. [PubMed: 18215882]
32. Sorzano COS, Marabini R, Velazquez-Muriel J, Bilbao-Castro JR, Scheres SHW, Carazo JM, Pascual-Montano A. XMIPP: a new generation of an open-source image processing package for electron microscopy. *Journal of Structural Biology*. Nov.2004 148
33. Zhang Q, Bettadapura R, Bajaj C. Macromolecular structure modeling from 3D EM using VOLROVER 2.01. *Biopolymers*. 2012

\$watermark-text

\$watermark-text

\$watermark-text

Fig. 1. (a) The Electron Tomography (ET) image acquisition process. (b) Projection-slice theorem: The missing wedge region in Fourier space is due to limited angle projections in the spatial domain.

Fig. 2. A pixel to pixel comparison between the model and segmented reconstruction over a given neighborhood detects over-segmentation at j th pixel x_j . The two tables below show the values of the indicator function for a neighborhood region of the model and corresponding segmented region.

Fig. 3.

A pixel to pixel comparison between the model and segmented reconstruction over a given neighborhood $\mathcal{N}(j)$ detects under-segmentation at j th pixel x_j . The two tables show the values of the indicator functions for a neighborhood region of the model and corresponding segmented region.

Fig. 4. (a) The cost function $c(f)$ (eqn.7) at each iteration. (b) RMS difference in the reconstructed volumes (normalized intensity) between consecutive iterations.

Fig. 5.

Flow diagram of the reconstruction algorithm. An initial reconstruction was used to locate regions having features that match a known spatial model. Regularization was performed in this region by performing segmentation of the updated reconstruction at each iteration and comparing the segmented region with the known spatial model.

Fig. 6.

(a) A schematic of HIV AIDS virus. The mushroom shaped virus spikes are the gp120 and the gp41 regions that protrude out of the virus envelope. (source: www.niaid.nih.gov) (b) Surface Rendering of the AIDS virus spike in un-bound state [27]. The virus spike is used as our model M in the regularization process.

Fig. 7.

Original 3D Shepp Logan Phantom: Volume rendering with a clipping plane and a 2D slice of the phantom. We simulate projections of this phantom and perform reconstruction comparison and evaluation.

Fig. 8. Reconstruction comparison of Shepp Logan phantom

2D central slices of Shape Reg, ML-EM, L2GF, WBP and ART reconstructions and 1D intensity profiles corresponding to the line drawn in the 2D image is shown along with the original Shepp Logan phantom's 1D profile. Regularization was applied at region ① in the Shape Reg reconstruction. WBP has sharp edges but suffers from high noise and ringing artifacts while ART appears blurred. Shape Reg performs best especially in the region ① where regularization is provided. In region ③ and the adjoining peaks Shape Reg performs best. Though no regularization was provided in regions ② and ③, Shape Reg performs better than ML-EM. However, Shape Reg has more non-smooth regions than ART and TV Reg.

Fig. 9.

Radial Fourier power spectrum plot. The spectrum of Shape Reg is closest to that of the original phantom's spectrum. At all spatial frequencies the Shape Reg spectrum closely follows that of the original.

Fig. 10. Misaligned model location: (a) 2D slice of the reconstruction output where regularization was performed in the marked region using a sphere model. (b) Arrows indicate edge artifacts that appear on the sphere due to misalignment.

Fig. 11.

RNA polymerase phantom for TEM simulation (a) 3D Volume rendering (VR) of all the particles, (b) 3D VR of the central RNA particle, (c) 2D slice of RNA particle.

Reconstruction of TEM simulations: 2D slice view of Shape Reg reconstruction of central RNA particle (d) with no TEM noise (e) with TEM noise and (f) misaligned prior model without TEM noise.

Fig. 12.

2D slices of the reconstructed volumes showing the virus spike regions. (Shape Reg)

Regularization through a prior model was applied *only* to the two spike regions shown. The regions between spikes show blurring in the other reconstruction methods. Our Shape Reg method shows reduced blurring overall. An improvement can also be seen in the neighboring regions where no regularization was applied.

Fig. 13.

2D slices of reconstructed volumes. Regularization (Shape Reg) was applied to three spike regions. Blurring between two of the spikes was severe and there was no reduction even with regularization. However, blurring was substantially reduced around one of the spikes. In other reconstruction methods, the regions between all the spikes show severe blurring and the spikes are not distinct.

Fig. 14.

Measuring resolution through Fourier Shell Correlation (FSC): The 0.5 FSC drop off for ML-EM was 31Å, for WBP was 30 Å, ART and L2GF were at 28Å while Shape Reg showed an improved resolution of 22Å

\$watermark-text

\$watermark-text

\$watermark-text

TABLE I

CNR at different locations of the reconstructed volumes. The average CNR value (μ) and the p value computed using Wilcoxon paired T test between Shape Reg and other reconstruction method's CNR is shown.

Location	Shape Reg	ML-EM	WBP	L2GF	ART
1	5.8	5.3	4.6	4.7	4.9
2	5.1	4.2	2.1	3.6	3.1
3	4.0	3.3	2.5	3.7	2.8
4	11.2	9.8	6.7	7.3	4.2
5	16.1	9.9	4.7	5.4	6.6
6	17.7	10.4	5.2	5.6	7.3
7	18.5	13.8	5.4	3.5	5.7
8	11.6	10.0	4.6	6.9	6.3
9	10.0	7.8	8.8	8.7	6.8
10	5.0	4.2	3.3	7.4	3.9
μ (p)	10.5	7.8 (0.002)	4.8 (0.002)	5.7 (0.0195)	5.1 (0.002)