

POISSON IMAGE RECONSTRUCTION WITH TOTAL VARIATION REGULARIZATION

Rebecca M. Willett¹, Zachary T. Harmany¹, and Roummel F. Marcia²

¹Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA

²School of Natural Sciences, University of California, Merced, Merced, CA 95348 USA

ABSTRACT

This paper describes an optimization framework for reconstructing nonnegative image intensities from linear projections contaminated with Poisson noise. Such Poisson inverse problems arise in a variety of applications, ranging from medical imaging to astronomy. A total variation regularization term is used to counter the ill-posedness of the inverse problem and results in reconstructions that are piecewise smooth. The proposed algorithm sequentially approximates the objective function with a regularized quadratic surrogate which can easily be minimized. Unlike alternative methods, this approach ensures that the natural nonnegativity constraints are satisfied without placing prohibitive restrictions on the nature of the linear projections to ensure computational tractability. The resulting algorithm is computationally efficient and outperforms similar methods using wavelet-sparsity or partition-based regularization.

Index Terms— Photon-limited imaging, Poisson noise, total variation, convex optimization, sparse approximation

1. INTRODUCTION

In applications such as nuclear medicine imaging, night vision, astronomy, and hyperspectral imaging, data are collected by counting a series of discrete events, namely photons hitting a detector within a certain time period. These counts reflect an underlying rate whose features we wish to reconstruct as accurately as possible. However the measurements are often inherently noisy when the count levels are low. In such contexts, the inhomogeneous Poisson process model [1, 2] has been used to effectively describe such phenomenon.

In many of these settings, the reconstruction problem from Poisson counts is complicated by the indirect nature of the measurements. Specifically, instead of observing the image of interest directly, we collect Poisson measurements of linear projections of the image [2]. The number of linear projection measurements may be much smaller than the number of pixel or voxel intensities to be estimated, resulting in a very ill-posed inverse problem. Regularization techniques are often employed to compensate for the ill-posedness of the estimation problem. Outside the Poisson context, for example in the presence of additive white Gaussian noise, regularization methods based on wavelet or curvelet sparsity [3], models of wavelets' clustering and persistence properties [4], and a variety of other penalties (cf. [5]) have proven successful.

Regularization based on a *total variation (TV)* seminorm has also garnered significant recent attention (cf. [6, 7]). This seminorm is described in detail below; in general, it measures how much an image varies across pixels, so that a highly textured or noisy image will have a large TV seminorm, while a smooth or piecewise

constant image would have a relatively small TV seminorm. This is often a useful alternative to wavelet-based regularizers, which are also designed to be small for piecewise smooth images but can result in spurious large, isolated wavelet coefficients and related image artifacts.

In the context of Poisson inverse problems, however, adaptation of these regularization methods can be challenging for two main reasons. First, the negative Poisson log likelihood used in the formulation of an objective function often requires the application of relatively sophisticated optimization theory principles. Second, because Poisson intensities are inherently nonnegative, the resulting optimization problem must be solved over a constrained feasible set, increasing the complexity of most algorithms. Some recent headway has been made using multiscale or smoothness-based penalties [1, 8, 9]. One recent work [10] bypasses the Poisson statistical model in favor of an additive Gaussian noise model through the use of the Anscombe variance stabilizing transform. This statistical simplification is not without cost, as the linear projections of the scene must now be characterized as nonlinear observations. Another recent effort solves Poisson image reconstruction problems with TV seminorm regularization, but the method involves a matrix inverse operation which can be extremely difficult to compute for large problems outside of deconvolution settings [11].

In this paper, we show that recent research into computationally efficient TV denoising can be effectively leveraged for TV-regularized Poisson image reconstruction, even in inverse problem settings. The resulting algorithm has several compelling features:

- the solution is guaranteed to satisfy the nonnegativity constraints, even if stopped early (before convergence);
- the solution does not exhibit spurious wavelet artifacts associated with wavelet sparsity regularization; and
- the method exhibits important convergence properties.

2. PROBLEM FORMULATION

Under the inhomogeneous Poisson process model, the true scene intensity $f^* \in \mathbb{R}_+^m$ is passed through an intensity preserving passive imaging system described by the matrix $A \in \mathbb{R}_+^{N \times m}$, yielding a detector photon intensity Af^* . We then observe a Poisson realization $y \in \mathbb{Z}_+^N = \{0, 1, 2, \dots\}^N$, that is

$$y \sim \text{Poisson}(Af^*), \quad (1)$$

Specifically, under the model in (1), the negative log likelihood of observing a particular vector of counts y is given by

$$-\log p(y|Af^*) = \sum_{j=1}^N (Af^*)_j - y_j \log(Af^*)_j - \log y_j!,$$

This work was supported by NSF CAREER Award No. CCF-06-43947, DARPA Grant No. HR0011-07-1-003, and NSF Grant DMS-08-11062.

where $(Af^*)_j$ is the j^{th} component of Af^* .

We propose estimating f^* from y using a TV-regularized Poisson log-likelihood objective function. Our algorithm reconstructs an estimate \hat{f} of the true intensity f^* by solving the following constrained minimization problem:

$$\begin{aligned} \hat{f} = \arg \min_f \quad & \Phi(f) \triangleq \phi(f) + \tau \|f\|_{\text{TV}} \\ \text{subject to} \quad & f \geq 0, \end{aligned} \quad (2)$$

where

$$\phi(f) \triangleq \sum_{j=1}^N [(Af)_j - y_j \log(Af)_j], \quad (3)$$

$$\begin{aligned} \|f\|_{\text{TV}} \triangleq & \sum_{k=1}^{\sqrt{m}-1} \sum_{l=1}^{\sqrt{m}} |f_{k,l} - f_{k+1,l}| \\ & + \sum_{k=1}^{\sqrt{m}} \sum_{l=1}^{\sqrt{m}-1} |f_{k,l} - f_{k,l+1}|, \end{aligned} \quad (4)$$

$\tau > 0$ is a regularization parameter, and the standard notation $f \geq 0$ means that the components of f are nonnegative. In (4), we have used 2-d pixel indices instead of vector indices in a small abuse of notation. We have also assumed that $f \in \mathbb{R}^m$ is a vector corresponding to a square $\sqrt{m} \times \sqrt{m}$ image for simplicity of presentation, but this assumption is not necessary for the algorithm. Here we assume the regularization parameter τ , which balances between the two terms in (2), is specified by the user. However in practice this parameter can be chosen via a cross-validation procedure.

3. TOTAL VARIATION REGULARIZATION

Recent work by Beck and Teboulle [7] presents a fast computational method for solving the TV-regularized problem

$$\begin{aligned} \tilde{f} = \arg \min_f \quad & \|\tilde{A}f - b\|_2^2 + \lambda \|f\|_{\text{TV}} \\ \text{subject to} \quad & f \in C, \end{aligned} \quad (5)$$

where $\lambda > 0$ is a tuning parameter, C is a closed convex set and \tilde{A} is a linear, *spatially invariant* blur operator. This method utilizes a gradient-based optimization approach founded on a monotone iterative shrinkage and thresholding algorithm. When $\tilde{A} = I$ (the identity matrix), then (5) reduces to least-squares “denoising” of b with a total variation regularizer. While (5) is certainly different from the objective upon which this paper is focused, namely (2), we will see in the sequel that methods for solving (5) can be leveraged effectively within an optimization framework for solving (2).

4. SPIRAL ALGORITHM

In previous work [9], we proposed optimization methods called *Sparse Poisson Intensity Reconstruction Algorithms (SPIRAL)* for solving (2) for various types of penalties. Specifically, we considered sparsity-based (ℓ_1) and structure-based (recursive dyadic partition [8]) penalties. In this section, we describe the application of SPIRAL for solving the TV-penalized optimization problem (2).

The SPIRAL approach sequentially approximates the objective function in (2) by regularized quadratic functions that are easier to minimize. In particular, if we denote the current estimate of f^* at the k^{th} iteration by f^k and consider the second-order Taylor expansion

of $\phi(f)$ around f^k where the Hessian $\nabla^2 \phi(f^k)$ is approximated by a positive scalar multiple of the identity matrix (i.e., $\nabla^2 \phi(f^k) \approx \eta_k I$ for $\eta_k > 0$), we obtain the following quadratic approximation $\phi_k(f)$ to $\phi(f)$ near f^k :

$$\phi_k(f) \triangleq \phi(f^k) + (f - f^k)^T \nabla \phi(f^k) + \frac{\eta_k}{2} \|f - f^k\|_2^2.$$

Replacing $\phi(f)$ in (2) by this quadratic function $\phi_k(f)$ leads to a separable quadratic minimization problem, which can be reformulated as

$$\begin{aligned} f^{k+1} = \arg \min_f \quad & \|f - s^k\|_2^2 + \frac{2\tau}{\eta_k} \|f\|_{\text{TV}} \\ \text{subject to} \quad & f \geq 0, \end{aligned} \quad (6)$$

where $s^k = f^k - \frac{1}{\eta_k} \nabla \phi(f^k)$. We choose the scalar η_k using Barzilai-Borwein (spectral) methods (see [9] for details).

The minimization (6) can be viewed as a denoising subproblem applied to s_k , the next gradient descent iterate. In fact, this denoising subproblem is of the form (5), where

$$\tilde{A} = I, \quad b = s^k, \quad \lambda = \frac{2\tau}{\eta_k}, \quad \text{and} \quad C = \{f \in \mathbb{R}^m: f \geq 0\}.$$

With this identification, we can solve this optimization subproblem with the method described in [7]. The process of approximating (2) by the denoising subproblem (6) is iterated until suitable convergence criteria are satisfied. We call this sequential quadratic programming approach with a total variation penalty SPIRAL-TV.

5. CONVERGENCE

In [9], we proved the following convergence result for the SPIRAL algorithm when applied to a generalized version of the objective function in (2).

Theorem 1. *Given $\epsilon > 0$, there exists some $K_\epsilon > 0$ such that for all $k > K_\epsilon$, the primal-dual pair (f^k, λ^k) nearly satisfies the Karush-Kuhn-Tucker (KKT) optimality conditions associated with (2); that is, for all k sufficiently large, we have*

$$\begin{aligned} \|\nabla \Phi(f^k) - \lambda^k\| &< \epsilon \\ \lambda^k &\geq 0 \\ f^k &\geq 0 \\ (\lambda^k)^T f^k &= 0. \end{aligned}$$

This theorem implies that for any arbitrarily small tolerance level ϵ , we can satisfy the KKT conditions to within that tolerance if the number of SPIRAL iterations, k , is sufficiently large. Note, as claimed in the introduction, that the iterates are always feasible with respect to the nonnegativity constraint $f^k \geq 0$.

6. SIMULATIONS

Here we consider the application of our reconstruction method to emission computed tomography (ECT). In medical ECT, a human subject is injected with a radioactive pharmaceutical used to tag certain tissues or tumors. To obtain a mapping of the radiopharmaceutical uptake, we detect photons emitted as the radiopharmaceutical decays. From these *projection* data (the indirect projections y in our problem), we wish to estimate the underlying radiopharmaceutical distribution (the intensity f^*). The probability transition matrix A is

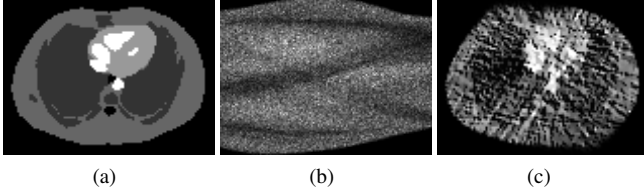


Fig. 1. Simulated emission tomography experimental setup: (a) true emission image, (b) noisy projection data, (c) estimate used as initialization.

derived from the physics and geometry of the detection device and data collection process, and also incorporates the attenuation correction needed to account for the attenuation effects introduced by bodily tissue [12].

The underlying 2-d intensity in our simulation is the 128×128 square image shown in Figure 1(a)); it is a standard test image included in the Image Reconstruction Toolbox (IRT) by Fessler [13]. We consider the case where the tomographic projection corresponds to a limited-angle parallel strip-integral geometry with 128 radial samples and 128 angular samples spaced uniformly over 135 degrees. With these parameters, the resulting transition probability matrix A was computed using the IRT software [13]. Ten realizations of the noisy tomographic data y are then simulated according to the inhomogeneous Poisson process (1). We only show images reconstructed using a particular realization of the data shown in Figure 1(b), the other nine realizations are used to examine the ten-trial average performance of the reconstruction methods considered. In the particular case shown, the noisy sinogram observations have a total photon count of 2.0×10^5 , a mean count over the support of the tomographic projections of 18.08, and a maximum count of 44.

6.1. Comparison with other methods

To show the effectiveness of the TV regularization, we compare the results of our SPIRAL-TV approach with the SPIRAL framework applied to objectives resulting from two different regularization methods. The first, a wavelet ℓ_1 -norm regularizer, reconstructs an estimate \hat{f}^{wavelet} according to the following objective:

$$\hat{f}^{\text{wavelet}} = \underset{f}{\arg \min} \quad \phi(f) + \tau \|W^T f\|_1 \quad (7)$$

subject to $f \succeq 0$,

where $W^T f$ is the discrete wavelet transform of f . Methods for solving this optimization are described in [9]. We denote this method as SPIRAL- ℓ_1 . As seen below, this regularization can result in significant artifacts corresponding to spurious non-zero wavelet coefficients.

The second regularization scheme is built on the framework of *recursive dyadic partitions* (RDPs), which we summarize here and are described in detail in [9]. It can be shown that partition-based methods are closely related to Haar wavelet denoising with an important hereditary constraint placed on the thresholded coefficients—if a parent coefficient is thresholded, then its children coefficients must also be thresholded [8]. This constraint is akin to wavelet-tree ideas which exploit persistence of significant wavelet coefficients across scales and have recently been shown highly useful in compressed sensing settings [4]. The hereditary constraint is also useful computationally since it admits an efficient dynamic programming implementation. Within this approach, it is straightforward to introduce

an efficient cycle-spinning strategy that yields a translationally invariant algorithm, boosting performance in practice. For brevity, we only show results for this translationally-invariant method, denoted SPIRAL-RDP-TI. While partition-based regularization yields accurate results in many settings, it also results in a non-convex optimization problem. The regularizer penalizes the number of cells in the partition-based estimate, essentially an ℓ_0 measure of sparsity. Although RDP-based regularization is computationally tractable (unlike ℓ_0 regularization), the non-convexity of the problem causes the solution to depend heavily on the initialization and convergence to a global minimizer cannot be guaranteed.

We also compare our SPIRAL approaches with two competing Poisson reconstruction methods. The first, denoted SPS-OS, uses a separable paraboloidal surrogate with ordered subsets algorithm [1]. The second, denoted EPL-INC-3, employs an incremental penalized Poisson likelihood EM algorithm and was suggested by Prof. Fessler as representative of the current state-of-the-art in emission tomographic reconstruction. In our experiments a roughness penalty based on the Huber potential function yielded the best results for both of these methods and are shown below. Both of these methods are available as part of the IRT [13]; specifically, we used the `pwls_sps_os` and `epl_inc` functions from the toolbox. In addition to these Poisson methods, we also compare to the SpaRSA algorithm [5] which solves the traditional CS ℓ_1 -regularized least-squares (ℓ_2 - ℓ_1) problem. Both wavelet-based approaches (SPIRAL- ℓ_1 and SpaRSA) use the Daubechies-6 wavelet basis for W . As the solution provided by SpaRSA is not constrained to be nonnegative, we threshold the result to obtain a feasible – and therefore more accurate – solution. Including this result allows us to demonstrate the effectiveness of solving the optimization formulation that utilizes the Poisson likelihood.

6.2. Algorithm setup

All of the methods considered here were initialized with the estimate shown in Fig. 1(c). This initialization results from two iterations of a non-convergent version of the EPL-INC-3 algorithm, itself initialized by a filtered back-projection estimate. All algorithms executed for a minimum of 50 iterations, and global convergence was declared when the relative change in the iterates, $\|f^{k+1} - f^k\|_2 / \|f^k\|_2$, fell below a tolerance $\text{tolP} = 5 \times 10^{-4}$. The advantage of this criteria is that it applies to general penalty functions. The disadvantage, however, is that it is possible that the change between two consecutive iterates may be small even though the iterates are still far from the true minimizer. However, forcing the algorithms to perform a suitable minimum number of iterations typically avoids any issues with premature termination.

Lastly, in all of the experiments presented in this paper, we chose any parameters associated with each algorithm (such as τ) to minimize the root-mean-square error (RMSE (%)) = $100 \cdot \|\hat{f} - f^* \|_2 / \|f^* \|_2$ of the reconstruction. While this would not be possible in practice, it does allow us to compare the *best-case* performance of various algorithms and penalization methods. In practical settings, regularization parameters can be chosen via cross-validation. This is particularly well-suited to many photon-limited imaging applications in which each detected photon has a time stamp associated with it; this timing information can be used to construct multiple independent and identically distributed realizations of the underlying Poisson process in software. The details of this procedure are a significant component of our ongoing research.

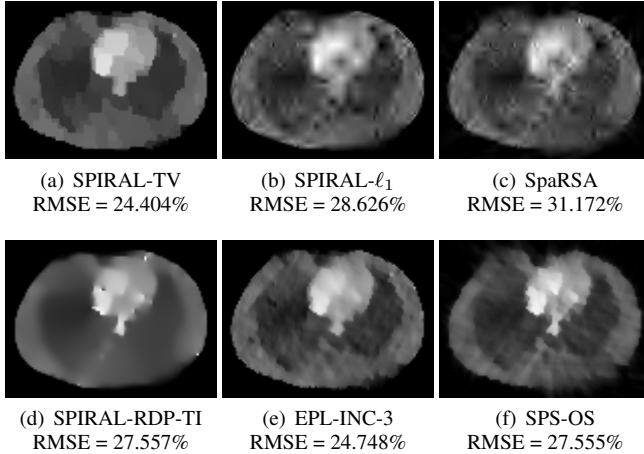


Fig. 2. Single-trial reconstructed images with RMSE values.

6.3. Results

The result of the ECT simulation described above are presented in Figure 2. The TV-regularized result is highly accurate and has the lowest RMSE among all the methods we considered. The wavelet-regularized methods (SPIRAL- ℓ_1 and SpaRSA) exhibit significantly more spurious artifacts and have higher RMSE. The performance of SpaRSA is particularly poor since it solves the conventional ℓ_2 - ℓ_1 minimization problem, which is unsuitable for Poisson image reconstruction. The partition-based result fares better than the ℓ_1 -based approaches; however it is oversmoothed and loses edge detail captured in the SPIRAL-TV result. The EPL-INC-3 method offers the toughest competition to the RMSE achieved by the SPIRAL-TV method, but there is a rough noise-like texture throughout the image, causing poor reconstruction in homogeneous regions. The SPS-OS method is ultimately limited by many streaking artifacts that cross through the image. Both the single-trial RMSE, and a ten-trial average RMSE are reported in Table 1, showing the claims above are robust with respect to different realizations of the data.

Method	Single-Trial RMSE (%)	Ten-Trial Average RMSE (%)
SPS-OS	27.555	27.057
EPL-INC-3	24.748	24.462
SpaRSA	31.172	29.987
SPIRAL- ℓ_1	28.626	28.050
SPIRAL-TV	24.404	24.270
SPIRAL-RDP-TI	27.557	27.669

Table 1. Reconstruction RMSE for single-trial results, and results averaged over ten trials. $\text{RMSE}(\%) = 100 \cdot \|\hat{f} - f^*\|_2 / \|f^*\|_2$.

7. CONCLUSIONS

Using total variation to regularize solutions to Poisson inverse problems yields highly accurate estimates. The proposed approach outperforms the current state-of-the-art approaches developed specifically for emission tomography. Particularly, it results in fewer spurious artifacts than wavelet-regularized methods, and unlike partition-regularized methods, is the result of a convex optimization procedure with provable convergence properties.

8. REFERENCES

- [1] S. Ahn and J. Fessler, “Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms,” *IEEE Trans. Med. Imaging*, vol. 22, no. 5, pp. 613–626, May 2003.
- [2] R. Nowak and E. Kolaczyk, “A multiscale statistical framework for Poisson inverse problems,” *IEEE Trans. Info. Theory*, vol. 46, pp. 1811–1825, 2000.
- [3] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: universal encoding strategies?,” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [4] R.G. Baraniuk, V. Cevher, M.F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. Inform. Theory*, 2008, Submitted.
- [5] S. Wright, R. Nowak, and M. Figueiredo, “Sparse reconstruction by separable approximation,” *IEEE Trans. Signal Process.*, vol. 57, pp. 2479–2493, 2009.
- [6] T. Chan and J. Shen, *Image Processing And Analysis: Variational, PDE, Wavelet, And Stochastic Methods*, Society for Industrial and Applied Mathematics, 2005.
- [7] A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems,” *IEEE Trans. Image Processing*, vol. 18, no. 11, pp. 2419–34, 2009.
- [8] R. Willett and R. Nowak, “Multiscale Poisson Intensity and Density Estimation,” *IEEE Trans. Inform. Theory*, vol. 53, no. 9, pp. 3171–3187, 2007.
- [9] Z. T. Harmany, R. F. Marcia, and R. M. Willett, “SPIRAL out of convexity: Sparsity-regularized algorithms for photon-limited imaging,” in *SPIE Computational Imaging*, 2010.
- [10] F.-X. Dupe, J. M. Fadili, and J.-L. Starck, “A proximal iteration for deconvolving poisson noisy images using sparse representations,” *IEEE Trans. Image Processing*, vol. 18, no. 2, pp. 310–321, 2009.
- [11] M. A. T. Figueiredo and J. M. Bioucas-Dias, “Restoration of Poissonian images using alternating direction optimization,” *IEEE Trans. Image Process.*, 2010, Accepted.
- [12] Y. Vardi, L. A. Shepp, and L. Kaufman, “A statistical model for positron emission tomography,” *J. Amer. Statist. Assoc.*, vol. 80, pp. 8–37, 1985.
- [13] J. Fessler *et al.*, “Image reconstruction toolbox,” <http://www.eecs.umich.edu/~fessler/code>.