

# A variational approach to reconstructing images corrupted by Poisson noise

UCLA CAM Report 05-49

Triet Le\*      Rick Chartrand<sup>†</sup>      Thomas J. Asaki<sup>‡</sup>

September 21, 2005

## Abstract

We propose a new variational model to denoise an image corrupted by Poisson noise. Like the ROF model described in [8] and [9], the new model uses total-variation regularization. Unlike the ROF model, our model uses a data-fidelity term that is suitable for Poisson noise. The result is that the strength of the regularization is signal dependent, precisely like Poisson noise.

## 1 Introduction

An important task of mathematical image processing is image denoising. The general idea is to regard a noisy image  $f$  as being obtained by corrupting a noiseless image  $u$ ; given a model for the noise corruption, the desired image  $u$  is a solution of the corresponding inverse problem.

---

\*University of California, Los Angeles, Mathematics Department, Box 951555 Los Angeles, CA 90095-1555. [tle@math.ucla.edu](mailto:tle@math.ucla.edu)

<sup>†</sup>Los Alamos National Laboratory, Theoretical Division, MS B284, Los Alamos, NM 87545. [rickc@lanl.gov](mailto:rickc@lanl.gov) Funded by the Department of Energy under contract W-7405-ENG-36.

<sup>‡</sup>Los Alamos National Laboratory, Computer and Computational Sciences Division, MS D413, Los Alamos, NM 87545. [asaki@lanl.gov](mailto:asaki@lanl.gov) Funded by the Department of Energy under contract W-7405-ENG-36.

Many algorithms are in use for reconstructing  $u$  from  $f$ . Since the inverse problem is generally ill-posed, most denoising procedures employ some sort of regularization. A very successful algorithm is that of Rudin, Osher, and Fatemi [8], which uses total-variation regularization. The ROF model regards  $u$  as the solution to a variational problem, to minimize the functional

$$F(u) := \int_{\Omega} |\nabla u| + \frac{\lambda}{2} \int_{\Omega} |f - u|^2, \quad (1)$$

where  $\Omega$  is the image domain and  $\lambda$  is a parameter to be chosen. The first term of (1) is a regularization term, the second a data-fidelity term. Minimizing  $F(u)$  has the effect of diminishing variation in  $u$ , while keeping  $u$  close to the data  $f$ . The size of the parameter  $\lambda$  determines the relative importance of the two terms.

Like many denoising models, the ROF model is most appropriate for signal independent, additive Gaussian noise. See [5] for an explanation of this in the context of Bayesian statistics. However, many important data contain noise that is signal dependent, and obeys a Poisson distribution. A familiar example is that of radiography. The signal in a radiograph is determined by photon counting statistics and is often described as particle-limited, emphasizing the quantized and non-Gaussian nature of the signal. Removing noise of this type is a more difficult problem. Besbeas et al. [1] review and demonstrate wavelet shrinkage methods from the now classical method of Donoho [3] to Bayesian methods of Kolaczyk [7] and Timmermann and Novak [10]. These methods rely on the assumption that the underlying intensity function is accurately described by relatively few wavelet expansion coefficients. Kervrann and Trubuil [6] employ an adaptive windowing approach that assumes locally piecewise constant intensity of constant noise variance. The method also performs well at discontinuity preservation.

In this paper, we propose a variational, total-variation regularized denoising model along the lines of ROF, but modified for use with Poisson noise.

## 2 Description of the Proposed Model

In what follows, we assume that  $f$  is a given grayscale image defined on  $\Omega$ , an bounded, open subset of  $\mathbb{R}^2$ , with Lipschitz boundary  $\partial\Omega$ . Usually,  $\Omega$  is a rectangle in the plane. We assume  $f$  is bounded and positive. Where

convenient below, we regard  $f$  as integer valued, but this will ultimately be unnecessary.

Recall the Poisson distribution with mean and standard deviation  $\mu$ :

$$P_\mu(n) = \frac{e^{-\mu} \mu^n}{n!}, \quad n \geq 0. \quad (2)$$

Our discussion follows well-known lines for formulating variational problems using Bayes' Law. See [5] for an example with the ROF model.

We wish to determine the image  $u$  that is most likely given the observed image  $f$ . Bayes' Law says that

$$P(u|f) = \frac{P(f|u)P(u)}{P(f)}. \quad (3)$$

Thus, we wish to maximize  $P(f|u)P(u)$ . Assuming Poisson noise, for each  $x \in \Omega$  we have

$$P(f(x)|u) = P_{u(x)}(f(x)) = \frac{e^{-u(x)} u(x)^{f(x)}}{f(x)!} \quad (4)$$

Now we assume that the region  $\Omega$  has been pixellated, and that the values of  $f$  at the pixels  $\{x_i\}$  are independent. Then

$$P(f|u) = \prod_i \frac{e^{-u(x_i)} u(x_i)^{f(x_i)}}{f(x_i)!}. \quad (5)$$

The total-variation regularization comes from our choice of prior distribution:

$$P(u) = \exp\left(-\beta \int_\Omega |\nabla u|\right), \quad (6)$$

where  $\beta$  is a regularization parameter.

Instead of maximizing  $P(f|u)P(u)$ , we minimize  $-\log(P(f|u)P(u))$ . The result is that we seek a minimizer of

$$\sum_i (u(x_i) - f(x_i) \log u(x_i)) + \beta \int_\Omega |\nabla u|. \quad (7)$$

We regard this as a discrete approximation of the functional

$$E(u) := \int_\Omega (u - f \log u) + \beta \int_\Omega |\nabla u|. \quad (8)$$

The functional  $E$  is defined on the set of  $u \in BV(\Omega)$  such that  $\log u \in L^1(\Omega)$ ; in particular,  $u$  must be positive almost everywhere.

The Euler-Lagrange equation for minimizing  $E(u)$  is

$$0 = \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right) + \frac{1}{\beta u} (f - u), \text{ with } \frac{\partial u}{\partial \vec{n}} = 0 \text{ on } \partial\Omega. \quad (9)$$

Compare this with the Euler-Lagrange equation for minimizing the ROF functional (1),

$$0 = \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda (f - u), \text{ with } \frac{\partial u}{\partial \vec{n}} = 0 \text{ on } \partial\Omega. \quad (10)$$

Notice that equation (9) is similar to equation (10), but with a variable  $\lambda = \frac{1}{\beta u}$ , which depends on the reconstructed image  $u$ . This local variation of the regularization parameter is better suited for Poisson noise because the expected noise increases with image intensity. Decreasing the value of the regularization parameter increases the denoising effect of the regularization term in the functional. We thus have a model that is similar to ROF but with a self-adjusting parameter.

### 3 Existence and Uniqueness

Next, we show existence and uniqueness of the minimizer for the model (8).

**Theorem 1.** *Let  $\Omega$  be a bounded, open subset of  $\mathbb{R}^2$  with Lipschitz boundary. Let  $f$  be a positive, bounded function. For  $u \in BV(\Omega)$  such that  $\log u \in L^1(\Omega)$ , let  $J(u) = \int_{\Omega} (u - f \log(u))$ ,  $TV(u) = \int_{\Omega} |\nabla u|$ ,  $E = TV + J$ . Then  $E(u)$  has a unique minimizer.*

*Proof.* First,  $J$  is bounded below by  $J(f)$ , so  $E$  is bounded below. Thus we can choose a minimizing sequence  $\{u_n\}$  for  $E$ . Then  $TV(u_n)$  is bounded, as is  $J(u_n)$ . By Jensen's inequality,

$$J(u_n) \geq \|u_n\|_1 - \|f\|_{\infty} \log \|u_n\|_1, \quad (11)$$

so  $\|u_n\|_1$  is bounded as well. This and the boundedness of  $TV(u_n)$  mean that  $\{u_n\}$  is a bounded sequence in the space  $BV(\Omega)$ . By the compactness of  $L^1$  in  $BV$  [4, p.176], there is  $u \in BV$  such that a subsequence  $\{u_{n_k}\}$

converges to  $u$  in  $L^1$ ; without loss of generality, we may assume that  $u_{n_k} \rightarrow u$  pointwise almost everywhere. By the lower semicontinuity of the  $BV$  norm [4, p.172],  $TV(u) \leq \liminf TV(u_{n_k})$ . Since  $u_{n_k} - f \log(u_{n_k})$  is bounded below (by  $-\|f - f \log f\|_\infty$ ), we may use Fatou's Lemma to conclude that  $J(u) \leq \liminf J(u_{n_k})$ . Thus  $E(u) \leq \liminf E(u_{n_k})$ , and  $u$  minimizes  $E$ .

Clearly  $TV$  is a convex function. Since the logarithm is a strictly concave function and  $f$  is positive,  $J$  is strictly convex. Hence  $E$  is strictly convex. Therefore, the minimizer  $u$  is unique.  $\square$

**Remark 1.** In (6) we use total variation as our choice of prior. However, we can replace  $\int |\nabla u| dx$  by  $\int |\nabla u|^2 dx$  or other convex functionals [2], and the existence and uniqueness still apply.

## 4 Numerical Results

We use gradient descent to solve (9). We implement a straightforward, discretized version of the following PDE:

$$u_t = \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right) + \frac{1}{\beta u} (f - u), \quad \text{with } \frac{\partial u}{\partial \vec{n}} = 0 \text{ on } \partial\Omega. \quad (12)$$

Derivatives are computed with standard centered-difference approximations. The quantity  $|\nabla u|$  is replaced with  $\sqrt{|\nabla u|^2 + \epsilon}$  for a small, positive  $\epsilon$ . The time evolution is done with fixed timesteps, until the change in  $u$  is sufficiently small. A similar procedure is used to implement the ROF model (1), which we use for comparison with our proposed model.

The example in Figure 1 consists of circles with intensities 70, 135, and 200, enclosed by a square frame of intensity 10, all on a background of intensity 5. Poisson noise is then added; see Figure 2(a). Note that there is no parameter associated with Poisson noise, but the noise magnitude depends on the absolute image intensities. The amount of noise in a region of the image increases with the intensity of the image there.

In Figures 2(b) and 2(c), the image has been denoised with the ROF (total variation) model (1). The result depends on the parameter  $\lambda$ . We choose  $\lambda$  according to the *discrepancy principle*, which says that the reconstruction should have a mean-squared difference from the noisy data that is equal to the variance of the noise. This is equivalent to the idea that of all the possible reconstructed images that are consistent with the noisy data, the image that

should be chosen is the one that is most regular. In our example, we used a noise variance of the mean-squared difference between the noised image and the original image. (In cases where there is no original image available, the noise variance would have to be estimated.) The resulting  $\lambda$  of 0.04 gives the image in Figure 2(b). The frame is almost completely washed out, as it differs from the background by less than the average noise standard deviation of 7.33. The frame can be preserved by increasing  $\lambda$  to 0.4, which has the effect of decreasing the strength of the regularization. The result, in Figure 2(c), is that the noise in the higher-intensity regions of the image is not removed.

For our Poisson-modified total variation model, we chose the parameter  $\beta$  according to a suitably modified discrepancy principle: the value of the data fidelity term  $\int u - f \log u$  for the reconstructed image should match that of the original image. In the example, this resulted in a  $\beta$  of 0.25. As noted above, the model behaves locally like ROF with a signal-dependent  $\lambda$  equal to  $1/\beta u$ . We thus have an effective  $\lambda$  of 0.8 for the background, 0.4 for the frame, and a smallest value of 0.02 in the center. Note that 0.4 was a value for  $\lambda$  for which the ROF model preserved the frame, while 0.02 gives a stronger regularization than that of ROF from the discrepancy principle. Therefore, it is not surprising that in Figure 2(d), the frame is preserved as well as in Figure 2(c), while the large-magnitude noise in the center is removed as well as in Figure 2(b). Also see Figure 3 for lineouts from the middle of the images, in which the qualitative properties of the results can be more clearly seen.

We can also compare our model with the ROF model by measuring the mean-squared difference between the reconstructed images and the original, noise-free image. This was 4.40 for our model, and 5.10 for the ROF model.

**Acknowledgement:** The first author would like to thank Professor Luminata Vese for many wonderful discussions and for introducing him to the great paper [5] by Professor Mark Green.

## References

- [1] P. BESBEAS, I. D. FIES, AND T. SAPATINAS, *A comparative simulation study of wavelet shrinkage estimators for Poisson counts*, International Statistical Review, 72 (2004), pp. 209–237.

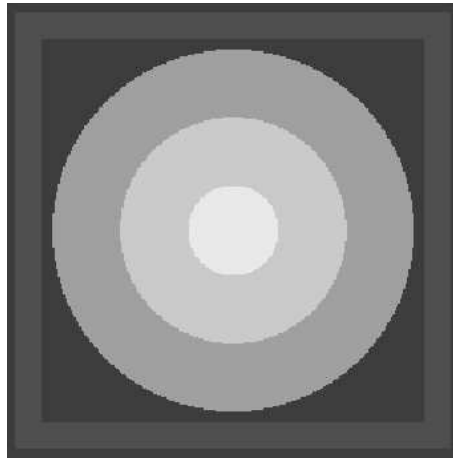


Figure 1: Circles image with frame. Image brightness has been adjusted for display, to allow the frame to be visible.

- [2] F. DEMENGEL AND R. TEMEM, *Convex functions of a measure and application*, Indiana Univ. Math. J., 33 (1984), pp. 673–709.
- [3] D. DONOHO, *Nonlinear wavelet methods for recovery of signals, densities and spectra from indirect and noisy data*, in Proceedings of symposia in applied mathematics: Different perspectives on wavelets, American Mathematical Society, 1993, pp. 173–205.
- [4] L. C. EVANS AND R. F. GARIEPY, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, 1992.
- [5] M. GREEN, *Statistics of images, the TV algorithm of Rudin-Osher-Fatemi for image denoising and an improved denoising algorithm*, CAM Report 02-55, UCLA, October 2002.
- [6] C. KERVRANN AND A. TRUBUIL, *An adaptive window approach for poisson noise reduction and structure preserving in confocal microscopy*, in International Symposium on Biomedical Imaging (ISBI'04), Arlington, VA, April 2004.
- [7] E. KOLACZYK, *Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds*, Statist. Sinica, 9 (1999), pp. 119–135.

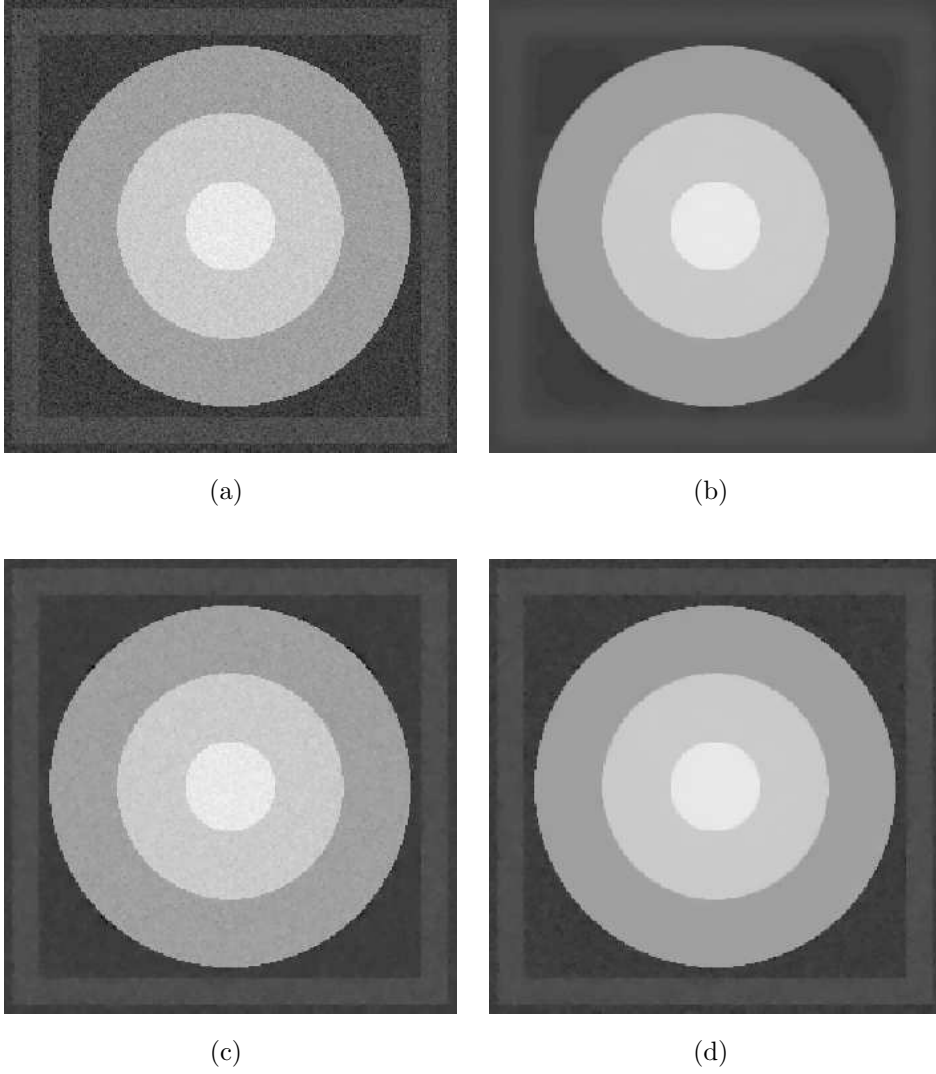


Figure 2: (a) Circles image with Poisson noise. (b) ROF denoised image. The frame is not well preserved. (c) ROF denoised image with decreased regularization strength. The frame is preserved, but the noise in the higher-intensity regions remains. (d) Image denoised with Poisson-modified total variation. Noise is removed at all scales, while preserving the frame.



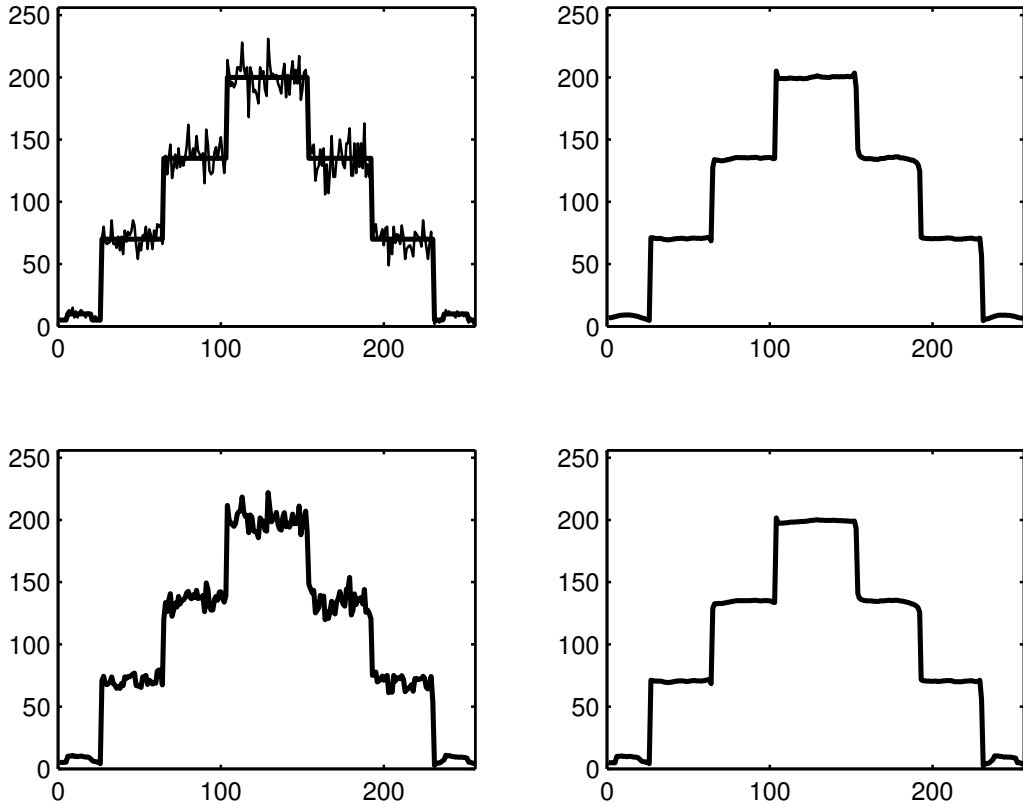


Figure 3: Upper left: Lineout of circles image with frame (heavy line) and Poisson-noised image (lighter line). Upper right: Lineout of ROF denoised image. The frame is not well preserved. Lower left: Lineout of ROF denoised image with decreased regularization strength. The frame is preserved, but the noise in the higher-intensity regions remains. Lower right: Lineout of image denoised with Poisson-modified total variation. Noise is removed at all scales, while preserving the frame.

- [8] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, *Physica D*, 60 (1992), pp. 259–268.
- [9] L. I. RUDIN AND S. OSHER, *Total variation based image restoration with free local constraints.*, in *ICIP (1)*, 1994, pp. 31–35.
- [10] K. TIMMERMANN AND R. NOVAK, *Multiscale modeling and estimation of Poisson processes with applications to photon-limited imaging*, *IEEE Trans. Inf. Theor.*, 45 (1999), pp. 846–852.