A Fast Algorithm for Deconvolution and Poisson Noise Removal

Xiongjun Zhang

School of Mathematics and Statistics
Central China Normal University

Joint work with Michael K. Ng (HKBU) and Minru Bai (HNU)

Outline

Outline

Degradation model

We focus on the following image reconstruction model:

$$g = \text{Poisson}(Af + c\mathbf{e}) \tag{1.1}$$

- $g \in \mathbb{R}^n$ is the observed image
- $\mathbf{e} := (1, \ldots, 1)^T \in \mathbb{R}^n$ is a vector of all ones
- $c \geq 0$ is a fixed background
- $A \in \mathbb{R}^{n \times n}$ is a spatial-invariant blur matrix
- $f \in \mathbb{R}^n$ is a column vector concatenated from the original image with size $m_1 \times m_2$ $(n = m_1 m_2)$
- Poisson$(\cdot)$ denotes the degradation by Poisson noise

Maximum a posterior (MAP) estimate

- Assume that the observed data g_i are independent random variables
- The probability distribution of g is the product of the probability distributions of all pixels[1]

$$P(g|f) = \prod_{i=1}^{n} \frac{[(Af + c\mathbf{e})_i]^{g_i} e^{-(Af+c\mathbf{e})_i}}{g_i!},$$

where $(Af + c\mathbf{e})_i$ and g_i denote the ith components of $Af + c\mathbf{e}$ and g, respectively.

- By taking the negative logarithm of the probability distribution, the Kullback-Leibler (KL) divergence $D_{KL}(Af + c\mathbf{e}, g)$ of $Af + c\mathbf{e}$ from g is given by[2]

$$D_{KL}(Af + c\mathbf{e}, g) := \sum_{i=1}^{n} (Af + c\mathbf{e})_i - g_i \log((Af + c\mathbf{e})_i) + \log(g_i!)$$

[1] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. IEEE Trans. Med. Imaging, 1(2):113-122, 1982.

[2] I. Csiszar. Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. Ann. Stat., 19(4):2032-2066, 1991.

Maximum a posterior (MAP) estimate

- The data fidelity term

$$\Phi(f) := \mathbf{e}^T(Af + c\mathbf{e}) - g^T \log(Af + c\mathbf{e}),$$

where the notation ^{T} denotes the transpose operator and $\log(\cdot)$ means componentwise.

TV regularization with Poisson noise removal

- Based on the prior information of an image, we use the TV as the regularization term. Then, a discrete version of the TV deblurring problem under Poisson noise is given by[3]

$$\min_{f \geq 0} \; \alpha(\mathbf{e}^T(Af + c\mathbf{e}) - g^T \log(Af + c\mathbf{e})) + \sum_{i=1}^{n} \|D_i f\|, \qquad (1.2)$$

  - α is a regularization parameter to balance the regularization term and the KL divergence term
  - the nonnegative constraint $f \geq 0$ guarantees that no negative intensities occur in the restored images
  - $\|\cdot\|$ denotes the Euclidean norm of a vector, and $D_i f \in \mathbb{R}^2$ denotes the first-order finite difference of f at pixel i in both horizontal and vertical directions

[3] T. Le, R. Chartrand, and T. J. Asaki. A variational approach to reconstructing images corrupted by Poisson noise. J. Math. Imaging Vis., 27(3):257-263, 2007.

Existing algorithms

- A gradient descent method (Le et al., 2007[4])
- Augmented Lagrangian (PIDAL) method (Figueiredo et al., 2010[5])
- An alternating split Bregman algorithm (PIDSplit+) (Setzer et al., 2010[6])
- Primal dual (PD) algorithm (Wen et al., 2016[7])

Disadvantage: slow for moderate accuracy

[4] T. Le, R. Chartrand, and T. J. Asaki. A variational approach to reconstructing images corrupted by Poisson noise. J. Math. Imaging Vis., 27(3):257-263, 2007.

[5] M. A. T. Figueiredo and J. M. Bioucas-Dias. Restoration of Poissonian images using alternating direction optimization. IEEE Trans. Image Process., 19(12):3133-3145, 2010.

[6] S. Setzer, G. Steild, and T. Teuber. Deblurring Poissonian images by split Bregman techniques. J. Vis. Commun. Image Rep., 21(3):193-199, 2010.

[7] Y.-W. Wen, R. H. Chan, and T. Zeng. Primal-dual algorithms for total variation based image restoration under Poisson noise. Sci. China Math., 59(1):141-160, 2016.

Outline

- Introduce $w_i \in \mathbb{R}^2$ to approximate $D_i f$ and z to approximate f
- Add the quadratic terms to penalize the differences between every pair of original and auxiliary variables

$$\min_{u,z,w,f} Q(u,z,w,f) := \alpha(\mathbf{e}^T u - g^T \log(u)) + \delta_{\mathbb{R}^n_+}(z) + \beta_1 \|u - (Af + c\mathbf{e})\|^2 + \sum_{i=1}^{n} \|w_i\| + \beta_2 \sum_{i=1}^{n} \|D_i f - w_i\|^2 + \beta_3 \|z - f\|^2, \tag{2.1}$$

  - $\beta_1, \beta_2, \beta_3$ are positive penalty parameters, respectively
  - $\delta_{\mathbb{R}^n_+}(z)$is the indicator function over $\mathbb{R}^n_+$

An alternating minimization algorithm

Algorithm 1: An alternating minimization algorithm for solving problem (2.1).
Input: $\alpha, \beta_1, \beta_2, \beta_3,\ u^0 > 0, z^0, w^0$. For $k = 0, 1, \ldots$, perform the following steps:
Step 1. Compute

$$f^{k+1} := \arg\min_f \left\{ \beta_1 \|u^k - (Af + c\mathbf{e})\|^2 + \beta_2 \sum_{i=1}^{n} \|D_i f - w_i^k\|^2 + \beta_3 \|f - z^k\|^2 \right\}. \tag{2.2}$$

Step 2. Compute

$$u^{k+1} := \arg\min_u \left\{ \alpha(\mathbf{e}^T u - g^T \log(u)) + \beta_1 \|u - (Af^{k+1} + c\mathbf{e})\|^2 \right\}. \tag{2.3}$$

Step 3. Compute

$$z^{k+1} := \arg\min_z \left\{ \delta_{\mathbb{R}^n_+}(z) + \beta_3 \|z - f^{k+1}\|^2 \right\}. \tag{2.4}$$

Step 4. For $i = 1, \ldots, n$, compute w_i

$$w_i^{k+1} := \arg\min_{w_i} \left\{ \|w_i\| + \beta_2 \|D_i f^{k+1} - w_i\|^2 \right\}. \tag{2.5}$$

An alternating minimization algorithm

- Algorithm 1 could be applied to both the isotropic and anisotropic TV. For simplicity, we will only discuss the isotropic case in detail since it is similar to deal with the anisotropic case.
- Compared with PIDAL, the subproblems of the alternating minimization algorithm can be solved exactly in each iteration.

Outline

Basic results

- Fixed point iteration by Browder et al.[8] and Opial[9]

Theorem

Let $\mathcal{X}$ be a Hilbert space, and let $\mathcal{T}$ be a nonexpansive self-mapping of $\mathcal{X}$ such that T is asymptotically regular, i.e., for any x in $\mathcal{X}$, $\|\mathcal{T}^{n+1}x - \mathcal{T}^n x\|$ tends to 0 as n tends to $+\infty$. Assume that $\mathcal{T}$ has at least one fixed point. Then, for any x in $\mathcal{X}$, a weak limit of a weakly convergent subsequence of the sequence of successive approximations $\{\mathcal{T}^n x\}$ is a fixed point of $\mathcal{T}$.

- The weak and strong convergence coincide for the finite dimension case
- When $\mathcal{X}$ is $\mathbb{R}^n$, the sequence of successive approximations can converge to a fixed point.
- Only need to prove that the operator $\mathcal{T}$ about the iteration sequence is nonexpansive, asymptotically regular, and has a fixed point

[8] F. E. Browder and W. V. Petryshyn. The solution by iteration of nonlinear functional equations in Banach spaces. Bull. Amer. Math. Soc., 72(3):571-575, 1966.

[9] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Amer. Math. Soc., 73(4):591-597, 1967.

Construction of the iteration

- We construct the iteration sequence of Algorithm 1. Let

$$H := \begin{pmatrix} \sqrt{\beta_1}A \\ \sqrt{\beta_3}I \\ \sqrt{\beta_2}D \end{pmatrix} \text{ and } v := \begin{pmatrix} u \\ z \\ w \end{pmatrix}$$

  and $M := \beta_1 A^T A + \beta_2 D^T D + \beta_3 I$, then $M = H^T H$ by the definition of M.

- $$\begin{aligned} Mf^{k+1} &= \beta_1 A^T(u^k - c\mathbf{e}) + \beta_2 D^T w^k + \beta_3 z^k \\ &= (BH)^T v^k - \beta_1 c A^T \mathbf{e}, \end{aligned}$$

  where

$$B := \begin{pmatrix} \sqrt{\beta_1}I & 0 & 0 \\ 0 & \sqrt{\beta_3}I & 0 \\ 0 & 0 & \sqrt{\beta_2}I \end{pmatrix}.$$

Construction of the iteration

- $$f^{k+1} = M^{-1}((BH)^T v^k - \beta_1 c A^T \mathbf{e}) := \mathcal{T}_f(v^k). \tag{3.1}$$
- Let $\varphi(u) := \mathbf{e}^T u - g^T \log(u)$. Then
$$u^{k+1} = \mathrm{Prox}_{\frac{\alpha}{2\beta_1}\varphi}(Af^{k+1} + c\mathbf{e}) := \mathcal{T}_1(f^{k+1}) = \mathcal{T}_1(\mathcal{T}_f(v^k)), \tag{3.2}$$
where $\mathrm{Prox}_\varphi : x \mapsto \arg\min_{y \in \mathcal{X}} \left\{\varphi(y) + \frac{1}{2}\|x - y\|^2\right\}$.
- $$z^{k+1} = \mathcal{P}_{\mathbb{R}^n_+}(f^{k+1}) := \mathcal{T}_2(f^{k+1}) = \mathcal{T}_2(\mathcal{T}_f(v^k)). \tag{3.3}$$
- Let $g(x) = \|x\|$, we obtain that
$$w^{k+1} := \begin{pmatrix} w_1^{k+1} \\ \vdots \\ w_n^{k+1} \end{pmatrix} = \begin{pmatrix} \mathrm{Prox}_{\frac{1}{2\beta_2}g}(D_1 f^{k+1}) \\ \vdots \\ \mathrm{Prox}_{\frac{1}{2\beta_2}g}(D_n f^{k+1}) \end{pmatrix} := \mathcal{T}_3(f^{k+1}) = \mathcal{T}_3(\mathcal{T}_f(v^k)). \tag{3.4}$$

- Combining (3.2), (3.3), and (3.4), we get

$$v^{k+1} = \begin{pmatrix} \mathcal{T}_1(\mathcal{T}_f(v^k)) \\ \mathcal{T}_2(\mathcal{T}_f(v^k)) \\ \mathcal{T}_3(\mathcal{T}_f(v^k)) \end{pmatrix}.$$

- Let

$$\mathcal{T} := \begin{pmatrix} \mathcal{T}_1(\mathcal{T}_f) \\ \mathcal{T}_2(\mathcal{T}_f) \\ \mathcal{T}_3(\mathcal{T}_f) \end{pmatrix}, \tag{3.5}$$

then $v^{k+1} = \mathcal{T}(v^k)$, which is the iteration of Algorithm 1.

Some properties of $\mathcal{T}$

Lemma

The operator $\mathcal{T}$ is nonexpansive.

Lemma

Let $v^k := ((u^k)^T, (z^k)^T, (w^k)^T)^T$ be generated by Algorithm 1, then $\sum_{k=1}^{+\infty} \|v^k - v^{k-1}\|^2$ converges and $\mathcal{T}$ is asymptotically regular.

Lemma

Suppose that $\mathrm{Ker}(A) \cap \mathrm{Ker}(D) = \{0\}$. Then the set of fixed points of $\mathcal{T}$ is nonempty.

Convergence result

Theorem

Suppose that $\mathrm{Ker}(A) \cap \mathrm{Ker}(D) = \{0\}$. For any initial point $u^0 \geq 0, z^0, w^0$, the sequence $\{(f^k, u^k, z^k, w^k)\}$ generated by Algorithm 1 converges to a minimizer (f^*, u^*, z^*, w^*) of Q.

- The condition $\mathrm{Ker}(A) \cap \mathrm{Ker}(D) = \{0\}$ is easy to satisfy. If A is a blurring matrix, then their rows sum up to one, while the differences matrix D (under appropriate boundary conditions) has kernel consisting of constant vectors so that their joint kernel must be the zero vector
- β_i increase gradually, $i = 1, 2, 3$

Outline

1. Introduction
2. An alternating minimization algorithm
3. Convergence analysis
4. **Experimental results**
5. Conclusions

Generated images

- The original image f is first scaled with the peak intensities M
- The scaled image is convolved with the blur kernel A and the background is added
- The Poisson noise is added in Matlab using the function **poissrnd**
- Different noise levels generated by different M and c

Average blur

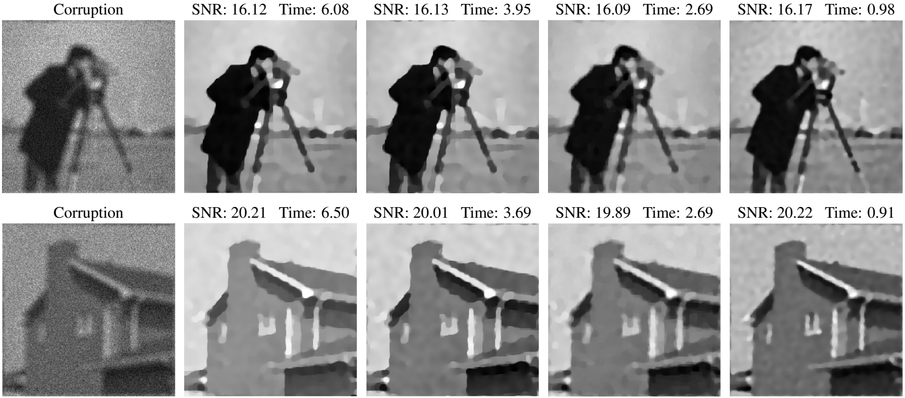

Figure: Recovered images (with SNR(dB) and CPU time(s)) of different algorithms on the Cameraman and House images corrupted by Average blur and Poisson noise with peak intensity $M = 80$ and $c = 10$. First column: Noisy and blurry images. Second column: Images recovered by PIDAL. Third column: Images recovered by PIDSplit+. Fourth column: Images recovered by PD. Fifth column: Images recovered by Algorithm 1.

Table: SNR(dB) values and CPU time(s) of different algorithms for the Cameraman and House images corrupted by Average blur and Poisson noise.

Image	c	M	PIDAL		PIDSplit+		PD		Algorithm 1	
			SNR	Time	SNR	Time	SNR	Time	SNR	Time
Cameraman	1	20	15.26	7.58	15.15	6.72	15.10	2.67	15.34	**0.64**
		40	15.57	6.21	15.47	5.73	15.08	2.21	15.79	**0.78**
		60	15.78	6.14	15.62	4.56	15.00	2.87	16.11	**0.82**
		80	15.99	5.55	16.01	3.72	14.63	2.39	16.30	**0.91**
	10	20	15.13	7.77	15.05	7.36	14.70	3.13	15.08	**0.75**
		40	15.75	7.47	15.64	5.29	14.60	3.50	15.55	**0.83**
		60	16.03	7.06	16.00	5.84	13.94	3.47	15.96	**1.00**
		80	16.12	6.08	16.13	3.95	16.09	2.69	16.17	**0.98**
House	1	20	18.68	9.01	18.33	8.97	18.59	2.98	18.74	**0.56**
		40	19.15	6.26	18.91	4.88	18.54	2.46	19.76	**0.74**
		60	19.53	6.36	19.35	3.96	19.19	2.80	20.01	**0.84**
		80	19.92	6.15	19.76	3.82	19.40	2.72	20.14	**0.89**
	10	20	18.47	9.07	18.07	8.12	18.35	3.14	18.27	**0.75**
		40	19.54	7.66	19.27	5.17	19.53	3.53	19.33	**0.82**
		60	19.95	6.27	19.79	4.48	19.65	3.29	19.79	**0.87**
		80	20.21	6.50	20.01	3.69	19.89	2.69	20.22	**0.91**

Difference images

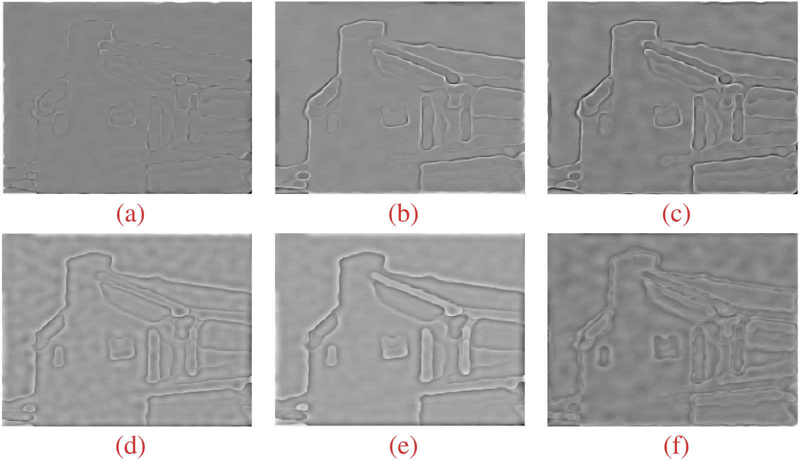

Figure: Difference images of House image among different algorithms in Figure 1. (a) Difference image between PIDAL and PIDSplit+. (b) Difference image between PD and PIDAL. (c) Difference image between PD and PIDSplit+. (d) Difference image between Algorithm 1 and PIDAL. (e) Difference image between Algorithm 1 and PIDSplit+. (f) Difference image between Algorithm 1 and PD.

Gaussian blur

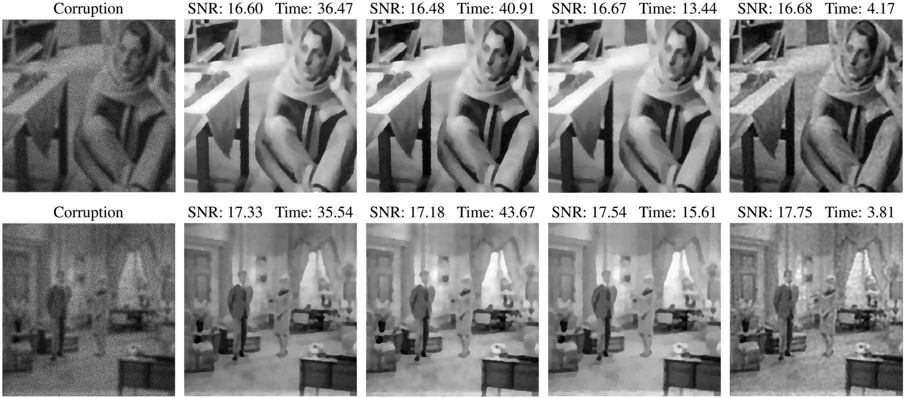

Figure: Recovered images (with SNR(dB) and CPU time(s)) of different algorithms on the Barbara and Livingroom images corrupted by Gaussian blur and Poisson noise with peak intensity $M = 40$ and $c = 1$. First column: Noisy and blurry images. Second column: Images recovered by PIDAL. Third column: Images recovered by PIDSplit+. Fourth column: Images recovered by PD. Fifth column: Images recovered by Algorithm 1.

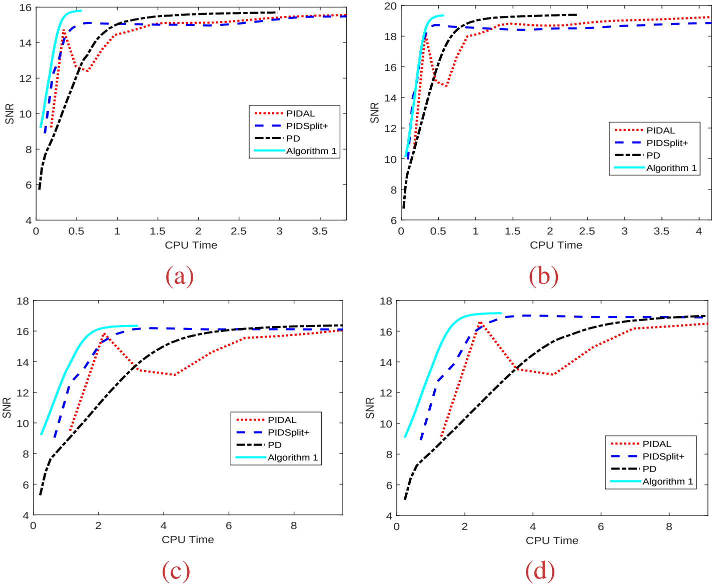

Figure: SNR(dB) values versus CPU time(s) for the images corrupted by Gaussian blur and Poisson noise with $M = 20, c = 1$. (a) Cameraman. (b) House. (c) Barbara. (d) Livingroom.

Outline

Conclusions

- Proposed an alternating minimization algorithm for the restoration of blurred images corrupted by Poisson noise
- Established the convergence of the proposed algorithm under very mild assumptions
- Preliminary numerical experiments showed the efficiency the proposed algorithm

Thank you for your attention!