Fast Bayesian Restoration of Poisson Corrupted Images with INLA

Takahiro Kawashima[1,a)] Hayaru Shouno[1,b)]

Abstract: Photon-limited images are often seen in fields such as medical imaging. Although the number of collected photons on an image sensor statistically follows Poisson distribution, this type of noise is intractable, unlike Gaussian noise. In this study, we propose a Bayesian restoration method of Poisson corrupted image using Integrated Nested Laplace Approximation (INLA), which is a computational method to evaluate marginalized posterior distributions of latent Gaussian models (LGMs). When the original image can be regarded as ICAR (intrinsic conditional auto-regressive) model reasonably, our method performs much faster than well-known ones such as loopy belief propagation-based method and Markov chain Monte Carlo (MCMC) without decreasing the accuracy.

1. Introduction

Estimating the original image from a noisy observation is one of the representative issues in the field of image processing. So far, a lot of image denoising methods have been proposed, but many of them such as Wiener filter require the assumption of additive white Gaussian noise. However, for example, in medical X-ray imaging systems, the number of detected photons stochastically fluctuates following Poisson distribution[4].

From a Bayesian viewpoint, it is natural to adopt Poisson observation modeling explicitly, and then solve the inverse problem to deal with such images. In fact, some Bayesian methods to restore Poisson corrupted images have been already proposed. Le et al. introduced a variational model with total-variation regularization[7]. Lefkimmiatis et al. applied quadtree decomposition and then estimated parameters using expectation-maximization (EM) algorithm to handle Poisson noise[8]. Shouno derived the fixed point equations of loopy belief propagation (LBP) by approximating Poisson distribution with binomial distribution[11]. Furthermore, Tachella et al. compared the performances of multiple MCMC methods to restore Poisson corrupted images [12].

EM, LBP, and MCMC methods are frequently used to evaluate the posterior (or its point estimate) of each pixel. In more general applications, Markov chain Monte Carlo (MCMC) methods succeeded due to their usability and accuracy. However, these iterative restoring methods take too long computational time because X-ray images are usually high resolution. On the other hand, recently, integrated nested laplace approximation (INLA) is proposed by Rue et al. [9], then the validity has been reported mainly in the fields of spatial statistics and epidemiology [10].

In this study, we try applying INLA to image restoration of Poisson corrupted images. When the original image seems to be reasonable to assume intrinsic CAR (ICAR) model, which means almost of adjacent pixels do not vary sharply, the proposed method rapidly obtained equal to the result by an MCMC simulation.

2. Methodology

2.1 Latent Gaussian Models

A latent Gaussian Models (LGM) is a class of statistical models, which include many commonly used statistical models such as auto-regressive (AR) model, conditional auto-regressive (CAR) model, and generalized linear mixture (GLM) model.

Fig. 1 shows the outline of an LGM. In LGMs, latent variables $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)^\top$ follow Gaussian Markov random field (GMRF)

$$p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}(\boldsymbol{\theta}_{\rm lat}), \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_{\rm lat})), \tag{1}$$

where $\boldsymbol{\theta}_{\rm lat}$ is a set of hyperparameters of latent GMRF. In addition, mean vector $\boldsymbol{\mu}$ and precision matrix $\boldsymbol{\Sigma}^{-1}$ are parameterized by $\boldsymbol{\theta}_{\rm lat}$.

Observations $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n)^\top$ should be conditionally independent and identically distributed on a parameter set $\boldsymbol{\theta}_{\rm obs}$

$$p(\boldsymbol{y}) = \prod_{i=1}^{n} p(\boldsymbol{y}_i|\boldsymbol{x}, \boldsymbol{\theta}_{\rm obs}), \tag{2}$$

and conditional distribution $p(\boldsymbol{y}_i|\boldsymbol{x}, \boldsymbol{\theta}_{\rm obs})$ must be belong to exponential family. Hereafter, for simplicity, we will put hyperparameter sets into $\boldsymbol{\theta}$,

[1] The University of Electro-Communications
[a)] kawashima@uec.ac.jp
[b)] shouno@uec.ac.jp

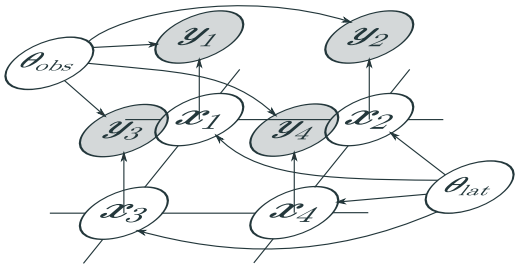


Fig. 1 An illustrated outline of a LGM. Highlighted nodes denote observated variables. Latent variables compose GMRF which is parameterized by $\boldsymbol{\theta}_{\text{lat}}$ and observations should be independent when latent variables $\boldsymbol{x}$ and hyperparameters $\boldsymbol{\theta}_{\text{obs}}$ are given.

$$\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\text{lat}}, \boldsymbol{\theta}_{\text{obs}}\}. \tag{3}$$

2.2 Integrated Nested Laplace Approximation

Integrated Nested Laplace Approximation (INLA) is a fast and accurate method to approximate posterior distributions of LGMs[9]. Especially when the number of hyperparamters is small (empirically $|\boldsymbol{\theta}| \lesssim 5$[10]), INLA performs very faster than Markov chain Monte Carlo (MCMC).

In LGMs, our aim is to get marginalized posteriors of hyperparameters

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \frac{p(\boldsymbol{\theta})p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})}{p(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})} \tag{4}$$

and latent variables

$$p(\boldsymbol{x}_i|\boldsymbol{y}) = \int p(\boldsymbol{x}_i|\boldsymbol{y},\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{y})\mathrm{d}\boldsymbol{\theta}. \tag{5}$$

In equation (4), each term in the numerator can be calculated by forward computation, but the denominator can not. Hence, first we apply Laplace (or simplified Laplace, which can fit with a skew normal distribution[1], [9]) approximation to the denominator

$$\tilde{p}(\boldsymbol{\theta}|\boldsymbol{y}) \propto \frac{p(\boldsymbol{\theta})p(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})}{\tilde{p}(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{y})}, \tag{6}$$

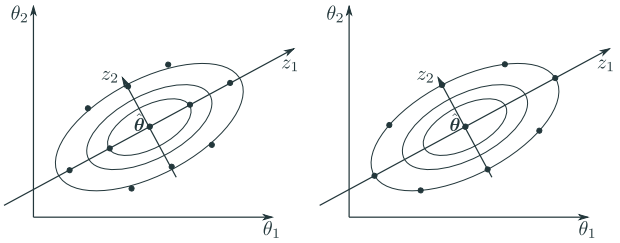
where $\tilde{p}(\cdot)$ denotes an approximation of $p(\cdot)$.

For latent variables (5), applying numerical integration as following:

(1) Exploring the mode of $\log \tilde{p}(\boldsymbol{\theta}|\boldsymbol{y})$ using an optimization algorithm (e.g., quasi-Newton method) with respect to $\boldsymbol{\theta}$.
(2) Arranging integration points in accordance with grid or central composite design (CCD)[3] strategy (see **Fig. 2**).
(3) Calculating $\log \tilde{p}(\boldsymbol{\theta}_h|\boldsymbol{y})$ at arranged points, where $h \in \{1, \ldots, H\}$ is the index of each point.
(4) Approximating equation (5) with weighted summation on θ_h:

$$p(\boldsymbol{x}_i|\boldsymbol{y}) \approx \sum_{h=1}^{H} p(\boldsymbol{x}_i|\boldsymbol{y},\boldsymbol{\theta}_h)p(\boldsymbol{\theta}_h|\boldsymbol{y})\Delta_h, \tag{7}$$

where Δ_h is the weight of $\boldsymbol{\theta}_h$.



(a) grid (b) CCD

Fig. 2 Strategies for numerical integration of $\log \tilde{p}(\boldsymbol{\theta}|\boldsymbol{y})$ when $|\boldsymbol{\theta}| = 2$. $\boldsymbol{z} = (z_1, z_2)^\top$ is the linear transformation of $\boldsymbol{\theta}$. If $\tilde{p}(\boldsymbol{\theta}|\boldsymbol{y})$ is Gaussian, $\mathbb{E}[\boldsymbol{z}] = \boldsymbol{0}$ and $\mathrm{Var}[\boldsymbol{z}] = I$ are satisfied.

3. Our Model

3.1 ICAR Models

To use INLA, we should choose a latent structure carefully. This is because the computaional performance of INLA strongly depends on the number of hyperparameters of the LGM.

ICAR model is one of the commonly used representations of spatial interaction[2]. In ICAR model, x_i has interaction with x_j $(j \in C_i)$, where C_i indicates the set of indices of nearest neighbor nodes of x_i (for simplicity, each variable is assumed to be a scalar value). Here it assumed that x_i follows

$$p(x_i) = \mathcal{N}\left(x_i \,\middle|\, \frac{1}{|C_i|}\sum_{j\in C_i} x_j, \frac{\sigma^2}{|C_i|}\right), \tag{8}$$

where σ^2 is the hyperparameter indicates variance between nodes. This model means some assumptions on $\boldsymbol{x}$:

- Each node hardly differs from adjacent ones.
- The similarity between adjacent nodes is controlled by the fixed effect σ^2.
- As $|C_i|$ is increasing, the variance of x_i decreases.

At a glance, this model seems not to be inappropriate. However, the precision matrix of $\boldsymbol{x}$ does not satisfy the condition of positive definiteness, then ICAR models become improper. Therefore, we regularize them to guarantee its computational stability by adding extra hyperparameter $d > 0$:

$$p(x_i) = \mathcal{N}\left(x_i \,\middle|\, \frac{1}{|C_i|+d}\sum_{j\in C_i} x_j, \frac{\sigma^2}{|C_i|+d}\right). \tag{9}$$

3.2 Observation Model

In photon-limited images, the number of observated photons y_i fluctuates as following Poisson distribution with the *true* value x_i. Hence, we use

$$p(y_i) = \mathrm{Poisson}(y_i|x_i) \tag{10}$$

as the observation model straightforwardly, and y_i is assumed to be independent from other observations with given $\boldsymbol{x}$. Therefore, the joint distribution of $\boldsymbol{y}$ becomes

$$p(\boldsymbol{y}) = \prod_{i=1}^{n} \mathrm{Poisson}(y_i|x_i). \tag{11}$$

(a) Lena (b) Boat (c) Cameraman

Fig. 3 The original images to evaluate parameter estimating methods for denoising Poisson corrupted images. Each image has been cropped into 128x128.

Table 1 Comparison of parameter estimating methods with PSNR, SSIM, and CPU time.

Image	Method	PSNR	SSIM	Time(s)
Lena	INLA	22.24	0.966	226.8
	MCMC	24.97	0.968	2083.5
	LBP	21.77	0.928	18794.7
Boat	INLA	21.76	0.947	209.5
	MCMC	24.03	0.960	1585.3
	LBP	21.50	0.915	18891.5
Cameraman	INLA	10.88	0.916	113.5
	MCMC	21.82	0.956	1360.7
	LBP	19.61	0.923	18809.8

3.3 Hyperpriors

In our model described above, we have two hyperparameters σ^2 and d. As their hyperpriors, we use uninformative prior on $[0, \infty)$.

4. Experiments

In this study, we compared denoising simulation results with INLA, MCMC (Hamiltonian Monte Carlo with No U-Turn Sampler[5], [6]), and LBP using the described model in Section 3. We simulated MCMC for 2000 steps and set first 1000 steps as burn-in. As the original images, we use "Lena", "Boat", and "Cameraman" for the evaluation (in **Fig. 3**). The each image is cropped into $n = 128 \times 128$(px). To compute point estimates of $\boldsymbol{x}$, we used expected a posteriori (EAP) estimation in INLA and MCMC. **Table 1** shows the comparison of results of INLA, MCMC, and LBP. In Table 1, we can find that INLA is very faster than other algorithms with little decreasing PSNR and SSIM from MCMC, expect for "Cameraman". LBP is too slow because it requires spectral decomposition of an $n \times n$ matrix in its algorithm. **Fig. 4** is the denoised images. Qualitatively, the images restored by INLA look clearer and unhazier than by other methods on a whole. On the other hand, we could not estimate the original "Cameraman" appropriately using INLA (Fig.4(c)). Specifically, the result of INLA is apparently different from others in its brightness. Then, to give additional consideration, we present histograms of intensities in **Fig. 5**. Seeing Fig.5(b), we can find that brighter pixels have been hardly restored in "Cameraman". In "Lena", the shape of *true* histogram is slightly flat. In contrast, the *true* histogram of "Cameraman" has two sharp peaks. Therefore, it is conceivable that the error of INLA mainly caused by this illness. In this case, INLA might fail to estimate the hyperparameters due to the illness of the original image, then also failed to get appropriate posteriors with respect to the latent variables.

5. Conclusion

In this study, we proposed and evaluated a denoising methods for Poisson corrupted images based on INLA. INLA requires an assumption of GMRF to latent variables. Hence, its accuracy decreases when the assumption is inappropriate. Whereas when the assumption is suitable for the original image, we can reduce much time compared to other Bayesian computational algorithms, such as LBP or MCMC with enough accuracy.

As a future work, we can also consider to use other models for latent models. For example, to apply segmentation to noisy image firstly, and then evaluate posteriors of each segmentation independently. In addition, we should adopt our method to real X-ray imaging data, and then evaluate its effectivity.

References

[1] Azzalini, A. and Capitanio, A.: Statistical applications of the multivariate skew normal distribution, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 61, No. 3, pp. 579–602 (online), DOI: 10.1111/1467-9868.00194 (1999).

[2] Besag, J. and Kooperberg, C.: On Conditional and Intrinsic Autoregression, *Biometrika*, Vol. 82, No. 4, pp. 733–746 (online), DOI: 10.2307/2337341 (1995).

[3] Box, G. E. P. and Wilson, K. B.: On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 13, No. 1, pp. 1–45 (online), available from ⟨http://www.jstor.org/stable/2983966⟩ (1951).

[4] Boyat, A. K. and Joshi, B. K.: A Review Paper: Noise Models in Digital Image Processing, *arXiv:1505.03489 [cs]*, (online), available from ⟨http://arxiv.org/abs/1505.03489⟩ (2015). arXiv: 1505.03489.

[5] Duane, S., Kennedy, A. D., Pendleton, B. J. and Roweth, D.: Hybrid Monte Carlo, *Physics Letters B*, Vol. 195, No. 2, pp. 216 – 222 (online), DOI: https://doi.org/10.1016/0370-2693(87)91197-X (1987).

[6] Hoffman, M. D. and Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo., *Journal of Machine Learning Research*, Vol. 15, No. 1, pp. 1593–1623 (2014).

[7] Le, T., Chartrand, R. and Asaki, T. J.: A variational approach to reconstructing images corrupted by Poisson noise, *Journal of mathematical imaging and vision*, Vol. 27, No. 3, pp. 257–263 (2007).

[8] Lefkimmiatis, S., Petros, M. and George, P.: Bayesian Inference on Multiscale Models for Poisson Intensity Estimation: Applications to Photon-Limited Image Denoising, *IEEE Transactions on Image Processing*, Vol. 18, No. 8, pp. 1724–1741 (online), DOI: 10.1109/TIP.2009.2022008 (2009).

[9] Rue, H., Martino, S. and Chopin, N.: Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 71, No. 2, pp. 319–392 (online), DOI: 10.1111/j.1467-9868.2008.00700.x (2009).

[10] Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P. and Lindgren, F. K.: Bayesian Computing with INLA: A Review, *Annual Review of Statistics and Its Application*, Vol. 4, No. 1, pp. 395–421 (online), DOI: 10.1146/annurev-statistics-060116-054045 (2017).

[11] Shouno, H.: Bayesian Image Restoration for Poisson Corrupted Image Using a Latent Variational Method with Gaussian MRF, *IPSJ Online Transactions*, Vol. 8, No. 0, pp. 15–24 (online), available from ⟨https://www.jstage.jst.go.jp/article/ipsjtrans/8/0/8_15/_article⟩ (2015).

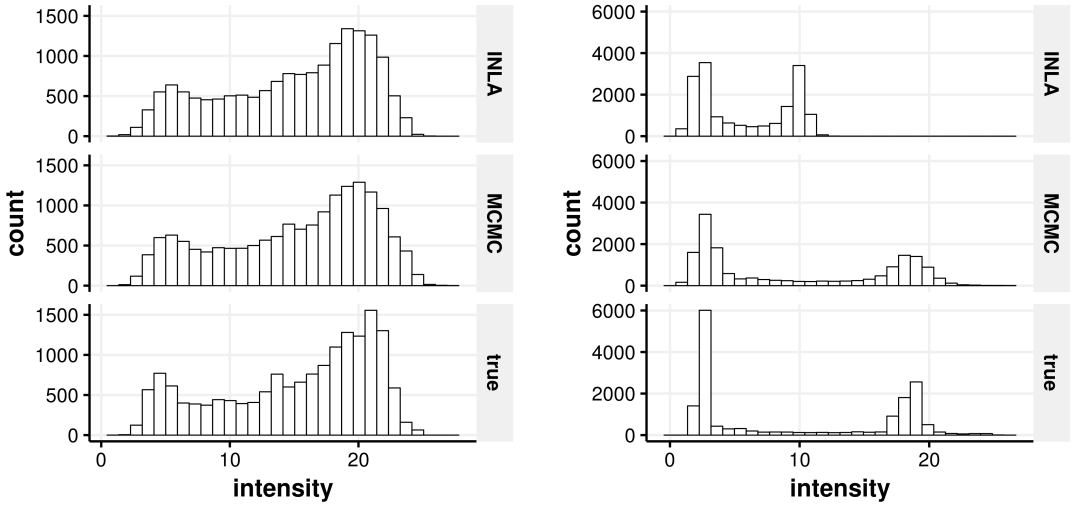
[12] Tachella, J., Altmann, Y., Pereyra, M., McLaughlin, S. and Tourneret, J. Y.: Bayesian Restoration of High-Dimensional

(a) Lena

(b) Boat

(c) Cameraman

Fig. 4 Poisson corrupted images and denoised ones using INLA, MCMC, and LBP. (a), (b), and (c) shows the results of "Lena", "Boat", and "Cameraman" respectively.



(a) Lena

(b) Cameraman

Fig. 5 The histograms of intensities of (a) "Lena" and (b) "Cameraman". Upper plot shows the estimated intensities by MCMC, and then middle and lower ones show INLA's and original's respectively.

Photon-Starved Images, *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome, IEEE, pp. 747–751 (online), DOI: 10.23919/EUSIPCO.2018.8553175 (2018).