Bayesian Image Restoration for Poisson Corrupted Image Using a Latent Variational Method with Gaussian MRF

Hayaru Shouno[1,a)]

Abstract: We treat an image restoration problem with a Poisson noise channel using a Bayesian framework. The Poisson randomness might be appeared in observation of low contrast object in the field of imaging. The noise observation is often hard to treat in a theoretical analysis. In our formulation, we interpret the observation through the Poisson noise channel as a likelihood, and evaluate the bound of it with a Gaussian function using a latent variable method. We then introduce a Gaussian Markov random field (GMRF) as the prior for the Bayesian approach, and derive the posterior as a Gaussian distribution. The latent parameters in the likelihood and the hyperparameter in the GMRF prior could be treated as hidden parameters, so that, we propose an algorithm to infer them in the expectation maximization (EM) framework using loopy belief propagation (LBP). We confirm the ability of our algorithm in the computer simulation, and compare it with the results of other image restoration frameworks.

Keywords: Poisson corrupted image, Bayesian inference, image restoration

1. Introduction

The technique of the noise reducing, which is called image restoration in the field of digital image processing, is an important in the meaning of the pre-processing. In order to reduce the noise, we should focus several clues for image property. The classical methods, such like the Gaussian or median filter methods, are focused on the similarity of the neighbor pixels. The total-variation method also focuses the similarity of the nearest pixel values within the measure of L_1 distance [1]. In these decade, a lot of image restoration procedure were also proposed. Introducing image-block similarity instead of pixel-pair similarity, Dabov et al. proposed a block matching method called BM3D that reduces the noise with less degrading edge-like features [2]. Ahron et al. proposed decomposing the image into the weighted sum of several basis with introducing the clues of the sparse representation [3].

From the theoretical viewpoint of statistical inference, these clues could be considered as knowledge, which is called prior, for the natural images. Thus, it is natural to introduce Bayesian inference into the image restoration. In the framework of Bayesian image restoration, additive white Gaussian noise (AWGN) was mainly discussed as the image corrupting process [4], [5], [6], [7], [8], since the analytical solution could be derived explicitly for the AWGN. However, in the real world, the noise corruption process often could not be described as such Gaussian observation. For example, we could treat the low contrast object observation, such like night vision, as a Poisson noised observation, since the observation of photons might be expressed as a rare event. The

Poisson noised observation also appears in some kinds of medical imaging like positron emission tomography (PET).

The Poisson image restoration methods were also proposed in these decades [9], [10], [11], [12], [13]. Figueiredo & Bioucas-Dias designed the objective function as the the likelihood function with several penalized term, and optimized the objective function with the alternating directional method of multipliers [11]. Ono & Yamada proposed optimization of the similar objective function by use of hybrid steepest decent method [13]. The other methods also designed the similar objective function for their applications.

These objective functions are defined by the likelihood function of Poisson observation with some penalized term. In the Bayesian manner, regarding such penalty term as a prior, we can consider the penalized method as a maximization of a posteriori (MAP) method. MAP method is a effective strategy for the image restoration, however, the strength balance between the prior and the likelihood is hard to determine in its framework. On the contrary, Bayes inference could determine the strength of the penalized term naturally as the hyperparameter inference [4], [6], [7], [8].

In this study, we treat a Poisson corrupted image restoration problem, and solve it in the manner of the Bayesian approach. The Bayesian approach also requires both likelihood and prior probability. We introduce the observation process as a likelihood, and also introduce Gaussian Markov random field (GMRF) as a prior after the fashion of the several works [4], [6].

Assuming the Poisson corruption observation makes difficult to derive the posterior probability in analytic form, since the Poisson variable take discrete and non-negative value. Thus, we introduce a latent variational approximation in the inference derivation [14], [15], [16], [17], [18], [19], [20]. In this study,

1   University of Electro-Communications, Chofu, Tokyo 182–8585, Japan
a)   shouno@uec.ac.jp

we transform the Poisson corruption process as the corresponding Bernoulli process, and introduce local latent variables to approximate the observation process as the Gaussian function for the likelihood in the Bayesian approach [17]. Once, we evaluate the observation likelihood as a Gaussian function, we can derive the posterior probability easily [19], [20]. In this formulation, we should introduce several latent parameters to describe the observation. In order to infer them, we introduce a expectation maximization (EM) algorithm [21], [22], which requires an iterative inference. Our previous work shows the preliminary results [19], [20] of this paper. In this paper, we refine the formulation of Bayesian inference of Ref. [19], and evaluate the accelerated results of Ref. [20] using several images with comaping of other methods.

In the following, we formulate the Bayesian image framework in Section 2 at first. After that, we confirm the abilities of our approach with computer simulation in Section 3. At last, we will conclude and summarize our approach in Section 4.

2. Bayesian Formulation

Our method is based on the Bayesian approach, so that, we explain both image observation process and prior probability in the following. Before the formulation, we define several notations. We consider the 2-dimensional image whose size are L_x and L_y, so that the total number of pixels M is described as $M = L_x L_y$.

2.1 Image Observation Process

The digital image is usually defined by the 2-dimensional pixel array. In the observation, we assume the observation for each pixel is independent, so that we consider single pixel observation at first. We consider each pixel has Poisson parameter λ_i where i means the position of the pixel. Denoting the observed pixel value as z_i, which means the number of of photons for the pixel position i, we regard the observation process as the following Poisson distribution:

$$p(z_i \mid \lambda_i) = \frac{(\lambda_i)^{z_i}}{z_i!} \exp(-\lambda_i). \tag{1}$$

Considering the Poisson process, Watanabe et al. treat the corruption process as a Bernoulli process, which counts the number of on-off event in the proper time bins [17]. Thus, we can translate Eq. (1) as the binomial distribution form:

$$p(z_i \mid \rho_i) = \binom{K}{z_i}(\rho_i)^{z_i}(1 - \rho_i)^{K-z_i}, \tag{2}$$

where $\lambda_i = K\rho_i$, and K means the upper limit of the counting. In this formulation, we can confirm Eq. (2) converges to the Poisson distribution Eq. (1) under the condition $K \to \infty$.
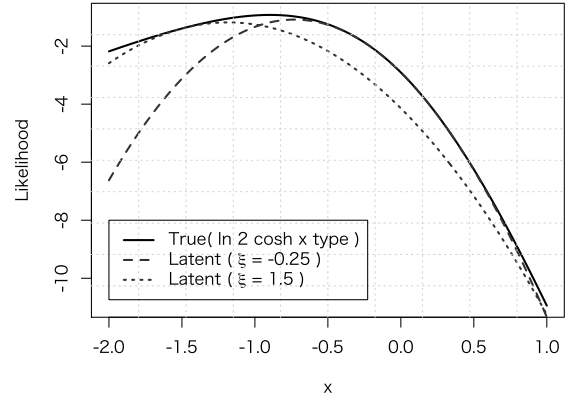
The parameter ρ_i in Eq. (2) is a non-negative parameter, which is just hard to treat for us, so that we introduce the logit transform into the parameter ρ_i, that is:

$$x_i = \frac{1}{2} \ln \frac{\rho_i}{1 - \rho_i}, \tag{3}$$

and obtain the conditional probability for the condition x_i as

$$p(z_i \mid x_i) = \binom{K}{z_i} \exp((2z_i - K)x_i - K \ln 2 \cosh x_i). \tag{4}$$

Comparison of Likelihoods

Fig. 1 Comparison of log-likelihood curves: The solid curve shows the true log-likelihood and both dotted and dashed are the variational curves.

Hence, the image corruption process can be interpreted as observing the z_i under the condition of x_i. Note that Eq. (4) is not a function of a observed value z_i but a parameter x_i.

2.1.1 Evaluation with Local Latent Variables

The term "$\ln 2 \cosh x_i$" in the Eq. (4) looks hard to tract for analysis. Thus, in this study, we introduce a latent variable evaluation [14], [17]. Palmer et al. proposed to evaluate the lower bound of super-Gaussians with multiplied form of the Gaussian distribution and concave parameter function [14], that is, any super-Gaussian, which is denoted by $p(u) = \exp(-g(u^2))$ where $g(\cdot)$ is a concave function, could be described as

$$p(u) = \exp(-g(u^2)) \tag{5}$$

$$= \sup_{\xi > 0} \varphi(\xi) \mathcal{N}\left(u \mid 0, \xi^{-1}\right), \tag{6}$$

$$\varphi(\xi) = \sqrt{\frac{2\pi}{\xi}} \exp\left(g^*\left(\frac{\xi}{2}\right)\right). \tag{7}$$

The function pair $g(u)$ and $g^*(\xi)$ is a convex conjugate relationship which is derived from Legendre's transform

$$g(u) = \inf_{\xi > 0} \xi u - g^*(\xi), \tag{8}$$

$$g^*(\xi) = \inf_{u > 0} \xi u - g(u). \tag{9}$$

Equation (6) consists of the Gaussian part for u and the non-Gaussian part described as the function $\varphi(\xi)$ where ξ is a latent-parameter. We evaluate the bound comes from the term $\ln 2 \cosh x_i$ in the likelihood Eq. (4) with substitution of $\ln 2 \cosh x_i = \ln 2 \cosh \sqrt{u_i^2} = g(u_i^2)$. Introducing the latent parameter form, we obtain the upper bound as:

$$\ln 2 \cosh x_i \leq \frac{\tanh \xi_i}{2\xi_i}(x_i^2 - \xi_i^2) + \ln 2 \cosh \xi_i, \tag{10}$$

where ξ_i is the latent parameter for the i-th pixel. **Figure 1** shows the comparison of log-likelihood curves. The solid curve shows the true log-likelihood, that is, $2(z - K)x - K \ln 2 \cosh x + \ln \binom{K}{z}$ where $K = 7$ and $z = 1$. The dashed and dotted ones show the latent variational curves where $\xi = -0.25$ and $\xi = 1.5$ respectively. We can see the upper bound of each latent variational curve is true log-likelihood function. Thus, we introduce the latent variational method into Eq. (4), we obtain the lower bound of the likelihood

function:

$$p(z_i \mid x_i)$$
$$\geq \binom{K}{z_i} \exp((2z_i - K)x_i - K\frac{\tanh \xi_i}{2\xi_i}(x_i^2 - \xi_i^2) + \ln 2 \cosh \xi_i)$$
$$= p_{\xi_i}(z_i \mid x_i) \tag{11}$$

Assuming the independence for each pixel observation, we can easily evaluate the lower bound of whole image corruption process:

$$p(z \mid x) = \prod_i p(z_i \mid x_i)$$
$$\geq \prod_i \binom{K}{z_i} \exp\left(-\frac{1}{2}x^{\mathrm{T}}\Xi x + z'^{\mathrm{T}}x\right)$$
$$\exp\left(\frac{1}{2}\xi^{\mathrm{T}}\Xi\xi - K \sum_i \ln 2 \cosh \xi_i\right)$$
$$= p_{\xi}(z \mid x), \tag{12}$$

where z' means observation vector

$$z' = (2z_1 - K, \cdots, 2z_i - K, \cdots, 2z_M - K)^{\mathrm{T}}, \tag{13}$$

and ξ means the collection of latent parameter $\{\xi_i\}$, and matrix Ξ means a diagonal matrix whose components are $\{K\frac{\tanh \xi_i}{\xi_i}\}$. Thus, we regard the lower bound of the likelihood $p_{\xi}(z \mid x)$ as the observation process which is denoted as a Gaussian form of x.

2.2 Prior Probability

Introducing the Bayesian inference requires several prior probability for the image in order to compensate for the loss of information through the observation. In this study, we assume a Gaussian Markov random field (GMRF) [4] for the prior. The GMRF prior is one of the popular one in the field of image restoration, and it is not the state-of-art prior in the meaning of the reducing noise performance. However, the GMRF is easy to treat in the analysis, so that we apply it in this study. Usually, we define the GMRF as the sum of neighborhood differential square of parameters $\sum_{(i,j)}(x_i - x_j)^2$ where x_i and x_j are neighborhood parameters. The energy function and the prior probability for the GMRF can be described as following:

$$H_{\mathrm{pri}}(x; \alpha, h) = \frac{\alpha}{2} \sum_{(i,j)}(x_i - x_j)^2 + \frac{h}{2} \sum_i x_i^2 \tag{14}$$

$$= \frac{1}{2}x^{\mathrm{t}}(\alpha\Lambda + hI)x \tag{15}$$

$$p(x \mid \alpha, h) = \frac{1}{Z(\alpha, h)} \exp\left(-H_{\mathrm{pri}}(x; \alpha, h)\right), \tag{16}$$

$$Z(\alpha, h) = \int dx \exp(-H_{\mathrm{pri}}(x; \alpha, h))$$
$$= \sqrt{|2\pi(\alpha\Lambda + hI)^{-1}|} \tag{17}$$

where the sum-up of (i, j) means the neighborhood pixel indices, and the matrix Λ and I mean the adjacent and identical matrices respectively. In Eq. (14), the first term means the GMRF part and the second means the Gaussian prior for the zero-center value for stable calculation.

2.3 Image Restoration Algorithm with Posterior

From the observation Eq. (12) and the prior Eq. (16), we can derive approximated posterior as

$$p_{\xi}(x \mid z, \alpha, h) \propto p_{\xi}(z \mid x) \, p(x \mid \alpha, h), \tag{18}$$

and the observation is evaluated with the latent-valued form, so that we can derive the approximated posterior as Gaussian distribution:

$$p_{\xi}(x \mid z, \alpha, h) \sim \mathcal{N}\left(x \mid m, S^{-1}\right), \tag{19}$$

$$S = \alpha\Lambda + hI + \Xi, \tag{20}$$

$$m = S^{-1}z'. \tag{21}$$

Considering the inference parameter of x as the posterior mean of the x, that is $\hat{x} = \langle x \rangle$, we can obtain the inference parameter explicitly:

$$\langle x \rangle = \sum_x x \, p_{\xi}(x \mid z, \alpha, h) = m. \tag{22}$$

2.4 Inference of Hyperparameters and Latent Variables

In order to obtain appropriate restoration with Eq. (22), the hyper-parameters α, h, and the latent variables $\{\xi\}$ should be adjusted properly. Hereafter we introduce the notation $\theta = \{\alpha, h, \xi\}$ for convenience. In order to solve, we applied a expectation maximization (EM) algorithm for inferring these parameters θ. EM algorithm consists of two-step alternate iterations for the system that has hidden variables [21], [22]. Assuming the notation t as the each time step, the EM algorithm could be described as the following two-steps:

- E-Step: Calculate Q-function that means the average of the likelihood function for the given parameter $\theta^{(t)}$:

$$Q(\theta \mid \theta^{(t)}) = \langle \ln p(x, z \mid \theta) \rangle_{x \mid \theta^{(t)}} \tag{23}$$

- M-Step: Maximize the Q-function for θ, and the arguments are set to the next hyper-parameters $\theta^{(t+1)}$:

$$\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}} \, Q(\theta \mid \theta^{(t)}) \tag{24}$$

Neglecting the constant term for the parameter θ, we can derive the Q-function in the E-step as:

$$Q(\theta \mid \theta^{(t)}) = -\frac{1}{2}(m^{(t)\mathrm{T}}S m^{(t)} + \operatorname{Tr} S S^{(t)^{-1}}) - \frac{1}{2} \ln |\alpha\Lambda + hI|$$
$$+ \frac{1}{2}\xi^{\mathrm{T}}\Xi\xi - K \sum_i \ln 2 \cosh \xi_i \tag{25}$$

$$S^{(t)} = \alpha^{(t)}\Lambda + h^{(t)}I + \Xi^{(t)} \tag{26}$$

$$m^{(t)} = S^{(t)^{-1}}z' \tag{27}$$

In order to maximize the Q-function in the M-step, we solve the saddle point equations $\frac{\partial Q}{\partial \alpha} = 0$, $\frac{\partial Q}{\partial h} = 0$, and $\frac{\partial Q}{\partial \xi_i} = 0$ for any i. Thus, we obtain

$$\sum_i \frac{\eta_i}{\alpha\eta_i + h} = m^{(t)\mathrm{T}}\Lambda m^{(t)} + \operatorname{Tr} \Lambda S^{(t)^{-1}}, \tag{28}$$

$$\sum_i \frac{1}{\alpha\eta_i + h} = \|m^{(t)}\|^2 + \operatorname{Tr} S^{(t)^{-1}}, \tag{29}$$

$$\xi_i = \sqrt{m_i^{(t)2} + S^{(t)^{-1}}_{ii}}, \tag{30}$$

Algorithm 1 Poisson corrupted image restoration using EM algorithm

1: Set the initial hyper-parameters $\alpha^{(0)}$, $\boldsymbol{\xi}^{(0)}$, and h
2: $t \leftarrow 0$
3: **repeat**
4:   Get mean $\boldsymbol{m}^{(t)}$ and accuracy $S^{(t)}$ for the posterior Eq. (21)

$$S^{(t)} = \alpha^{(t)}\Lambda + h + \Xi^{(t)}$$
$$\boldsymbol{m}^{(t)} = S^{(t)-1}\boldsymbol{z}'$$

5:   Update parameters α and $\{\xi_i\}$

$$\xi_i^{(t+1)} = \sqrt{m_i^{(t)2} + S^{(t)-1}_{ii}}$$
$$\frac{1}{\alpha^{(t+1)}} = \frac{1}{M-1}\left(\boldsymbol{m}^{(t)\mathrm{T}}\Lambda\boldsymbol{m}^{(t)} + \mathrm{Tr}\,\Lambda S^{(t)-1}\right)$$

6:   $t \leftarrow t + 1$
7: **until** restoration image $\{m_i\}$ is converged.
8: $\hat{\boldsymbol{x}} \leftarrow \boldsymbol{m}$ as the restored image
9: $\hat{\lambda}_i \leftarrow K\rho(\hat{x}_i)$ where $\rho(\cdot)$ is the inverse logit transform: $\rho(\hat{x}_i) = e^{\hat{x}_i}/(e^{\hat{x}_i} + e^{-\hat{x}_i})$.

where $\{\eta_i\}$ are eigenvalues of the adjacent matrix Λ and $S^{(t)-1}_{ii}$ is the (i,i)th diagonal component of the matrix $S^{(t)-1}$.

In order to obtain the exact hyper parameters α and h, we have to solve Eqs. (28) and (29) simultaneously, however, it makes increasing computational cost. Thus, hereafter, we assume the hyperparameter h is fixed and given as $h \ll \alpha$. Then, we obtain the inference of the hyperparameter α as

$$\frac{1}{\alpha} = \frac{1}{M-1}\left(\boldsymbol{m}^{(t)\mathrm{T}}\Lambda\boldsymbol{m}^{(t)} + \mathrm{Tr}\,\Lambda S^{(t)-1}\right), \qquad (31)$$

since $\{\eta_i\}$, which are the eigenvalues of the adjacent matrix Λ, only has a zero component and other components are positive values. Putting all of them in together, we obtain the Poisson corrupted image restoration algorithm as the Algorithm 1.

2.4.1 Approximating Posterior Mean with Loopy Belief Propagation

In the Algorithm 1, each E-step requires the inverse of accuracy matrix $S^{(t)-1} = (\Xi^{(t)} + \alpha^{(t)}\Lambda + h^{(t)}I)^{-1}$ to calculate the parameters. In general, the computational cost for inverse of a matrix that size is $M \times M$ requires $O(M^3)$ order. In this study, we assume the restoring image size is $M = L_x \times L_y$, so that, in the meaning of calculation scalability, the reduction of the cost is important for the application

In order to reduce the calculation cost, we introduce the loopy belief propagation (LBP) into the E-step in the algorithm. In the manner of the Gaussian graphical model, the efficacy of the LBP were confirmed [6], [20], [23]. Our approximated posterior, that is Eq. (18), is expressed as a kind of Gaussian form, so that we can apply the LBP for the restoration. For applying LBP, we modify the evaluation of restoration value described as Eq. (22) to the marginal posterior $p_\xi(x_i \mid z, \alpha, h)$ mean (MPM):

$$x_i^* = \langle x_i \rangle_{\mathrm{MPM}} = \int dx_i\, x_i\, p_\xi(x_i \mid z, \alpha, h). \qquad (37)$$

Obtaining the marginal posterior mean, we apply a local message passing algorithm defined by LBP. Hereafter, for convenience, we introduce the following notations:

Algorithm 2 Poisson corrupted image restoration using EM algorithm with LBP

1: Set the initial hyper-parameters $\alpha^{(0)}$, $\boldsymbol{\xi}^{(0)}$, and h
2: $t \leftarrow 0$
3: **repeat**
4:   Set $\beta_i^{(t)} = K\frac{\tanh\xi_i^{(t)}}{\xi_i^{(t)}}$, and $y_i^{(t)} = (2z_i - K)/\beta_i^{(t)}$.
5:   Carry out the LBP, where update Eqs. are (43) and (44), under the given hyper-parameters $\alpha^{(t)}$, $\{\beta_i^{(t)}\}$.
6:   After convergence of the LBP, solve several statistics: the restoration pixel values $\{m_i\}$, those of variances $\{(\sigma_i)^2\}$, and the correlations $\{s_{ij}\}$:

$$m_i = \frac{\beta_i^{(t)}y_i^{(t)} + \sum_{j\in N(i)}\gamma_{j\to i}\mu_{j\to i}}{\beta_i^{(t)} + h + \sum_{j\in N(i)}\gamma_{j\to i}} \qquad (32)$$

$$\sigma_i^2 = \left(\beta_i^{(t)} + h + \sum_{j\in N(i)}\gamma_{j\to i}\right)^{-1}, \qquad (33)$$

$$s_{ij} = \frac{(\alpha^{(t)} - \gamma_{i\to j})(\alpha^{(t)} - \gamma_{j\to i})}{\alpha^{(t)3}}. \qquad (34)$$

7:   Update the hyper-parameters:

$$\xi_i^{(t+1)} = \sqrt{m_i^2 + \sigma_i^2} \qquad (35)$$

$$\frac{1}{\alpha^{(t+1)}} = \frac{\left(\sum_{(i,j)}(m_i - m_j)^2 + \sigma_i^2 + \sigma_j^2 - 2s_{ij}\right)}{M-1}. \qquad (36)$$

8:   $t \leftarrow t + 1$
9: **until** restoration image $\{m_i\}$ is converged.
10: $\hat{\boldsymbol{x}} \leftarrow \boldsymbol{m}$ as the restored image
11: $\hat{\lambda}_i \leftarrow K\rho(\hat{x}_i)$ where $\rho(\cdot)$ is the inverse logit transform: $\rho(\hat{x}_i) = e^{\hat{x}_i}/(e^{\hat{x}_i} + e^{-\hat{x}_i})$.
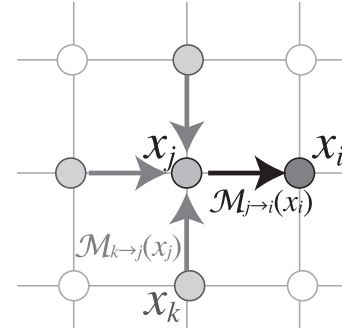
Fig. 2 Schematic diagram of message passing of the LBP: The LBP algorithm can be applied to infer the marginalized posterior. Each circle shows the pixel, which has 4 nearest neighbors. For instance, considering the message from the jth unit to i-th unit named $\mathcal{M}_{j\to i}(x_i)$, the message integrate the messages from the jth nearest neighbor except i-th.

$$\beta_i = K\frac{\tanh\xi_i}{\xi_i}, \qquad (38)$$

$$y_i = \frac{2z_i - K}{\beta_i}. \qquad (39)$$

Then we obtain the observation likelihood Eq. (11) for i-th node as

$$p(y_i \mid x_i) \propto \exp\left(-\frac{\beta_i}{2}(y_i - x_i)^2\right). \qquad (40)$$

The LBP algorithm is a kind of local message passing. Here, we denote the message from the jth node to the i-th node as $\mathcal{M}_{j\to i}(x_i)$. **Figure 2** shows the schematic diagram of the message passing. Here, considering the message $\mathcal{M}_{j\to i}(x_i)$, we should integrate the message of the jth connected units except i-th. In each LBP iteration, this message passing is carried out for each connection. In the GMRF case, the message can be derived as

$$\mathcal{M}_{j\to i}(x_i) \propto \int dx_j\, p(y_j \mid x_j) \exp\left(-\frac{\alpha}{2}(x_i - x_j)^2 - \frac{h}{2}x_j^{\,2}\right)$$
$$\prod_{k\in N(j)\backslash i} \mathcal{M}_{k\to j}(x_j), \qquad (41)$$

where $N(j)$ means the collection of the connected units to the jth unit, and $N(j)\backslash i$ means the collection except i-th unit. From the form of the integral in Eq. (41), we can regard the message from the jth node to the i-th node as the following Gaussian

$$\mathcal{M}_{j\to i}(x_i) \propto \mathcal{N}\left(x_i \mid \mu_{j\to i}, \gamma_{j\to i}^{-1}\right). \qquad (42)$$

Substituting the message form Eq. (42) into Eq. (41), we can derive the message update rule as

$$\mu_{j\to i} = \frac{\beta_j y_j + \sum_{k\in N(j)\backslash i}\gamma_{k\to j}\mu_{k\to j}}{\beta_j + \sum_{k\in N(j)\backslash i}\gamma_{k\to j} + h} \qquad (43)$$

$$\frac{1}{\gamma_{j\to i}} = \frac{1}{\alpha} + \frac{1}{\beta_j + \sum_{k\in N(j)\backslash i}\gamma_{k\to j} + h}. \qquad (44)$$

The LBP requires iterations for convergence of the message values. After the convergence, the marginal posterior required for the EM algorithm can be evaluated as

$$p(x_i \mid \mathbf{y}, \alpha, h) \propto p(y_i \mid x_i)\prod_{j\in N(i)}\mathcal{M}_{j\to i}(x_i), \qquad (45)$$

$$p(x_i, x_j \mid \mathbf{y}, \alpha, h) \propto p(y_i \mid x_i)p(y_j \mid x_j)$$
$$\exp\left(-\frac{\alpha}{2}(x_i - x_j)^2 - \frac{h}{2}(x_i^2 + x_j^2)\right)$$
$$\prod_{k\in N(i)\backslash j}\mathcal{M}_{k\to i}(x_i)\prod_{l\in N(j)\backslash i}\mathcal{M}_{l\to j}(x_j). \qquad (46)$$

Thus, the Q-function for the proposing EM algorithm is

$$Q(\theta \mid \theta^{(t)}) = \langle \ln p(\mathbf{x}, \mathbf{y} \mid \theta)\rangle_{\mathrm{MPM}}$$
$$= \frac{1}{2}\sum_i \ln\beta_i - \sum_i \frac{\beta_i}{2}\left\langle (y_i - x_i)^2\right\rangle_{\mathrm{MPM}}$$
$$+ \frac{M-1}{2}\ln\alpha - \frac{\alpha}{2}\sum_{(i,j)}\left\langle (x_i - x_j)^2\right\rangle_{\mathrm{MPM}}, \qquad (47)$$

where $\langle\cdot\rangle_{\mathrm{MPM}}$ means the average over the marginal posterior (MPM) denoted as Eqs. (45) and (46). Deriving Eq. (47), we also assume the hyper-parameter h is enough small $h/\alpha \ll 1$.

Let put them all together, the proposing LBP approximated solution is shown as the algorithm 2.

3. Computer Simulation Results

We evaluate the restoration performance with computer simulation. In the following, at first, we compare our latent variational restoration abilities between the EM method solution (algorithm 1) and its LBP approximated one (algorithm 2) from the viewpoint of the time consumption and restored image quality. After that, we compare our latent variational restoration with LBP solution (algorithm 2) with the conventional median filter restoration and the standard Gaussian LBP (GLBP) solution [6].

For the evaluation, we extract several image patches from the standard images, called "Cameraman," "Lena," "Barbara," "Fingerprint" and "Peppers."

We resample each image into the half-size with weak Gaussian

blurring in order to increase smoothness since we assume observing object would be smooth. We regard the resampled images as the observing images, and extract several image patches with size of $L_x \times L_y$. The Poisson corruption process is influenced with the contrast of the observing images, so that, we control the maximum and minimum of the image in the simulation. Here, we regard the patch image as $\mathbf{I} = \{I_i\}$, where i means the position of the pixel. In the simulation, in order to control the contrast of the image, we introduce the pixel value range $(\lambda_{\mathrm{Min}}, \lambda_{\mathrm{Max}})$ which mean the minimum and the maximum values of the Poisson parameters image. Assuming the minimum and the maximum values of the patch image as I_{Min}, and I_{Max} respectively, we define the source image λ^* of the i-th pixel λ_i^* as a linear transform (see **Fig. 3** (e)):

$$\lambda_i^* = \frac{\lambda_{\mathrm{Max}} - \lambda_{\mathrm{Min}}}{I_{\mathrm{Max}} - I_{\mathrm{Min}}}(I_i - I_{\mathrm{Min}}) + \lambda_{\mathrm{Min}}. \qquad (48)$$

Thus, the difference between λ_{Max} and λ_{Min} becomes large, the source image becomes high contrast which means low noise case. Hereafter, we fix the $\lambda_{\mathrm{Min}} = 2$, and only control the parameter λ_{Max} as the strength of the accuracy of the observation.

3.1 Evaluation of LBP Restoration Abilities

We compare our latent variational solutions between the exact EM algorithm (algorithm 1) and its LBP approximated one (algorithm 2). First, we compare them from the viewpoint of the hyperparameter inference and the restoration quality. The results of the hyperparameter α, which can be regarded as the strength of the prior Eq. (16), is shown as **Table 1** We use a cropped image of "cameraman" with patch size $L_x = L_y = 64$ as the source image, and apply several contrast parameters λ_{Max}, which are chosen from $\{20, 40, 80, 160\}$, to control the noise strength. The smaller the parameter λ_{Max} becomes, the larger the noise strength becomes. We make 10 trials for each condition, and show the results of the averages with standard deviation in the table. From the table, We can see the results in each λ_{Max} are almost same.

We also compare these algorithms with restored image quality by use of the peak signal to noise ratio (PSNR). The PSNR is defined as a kind of similarity between the reference image λ^* and the test image λ as:
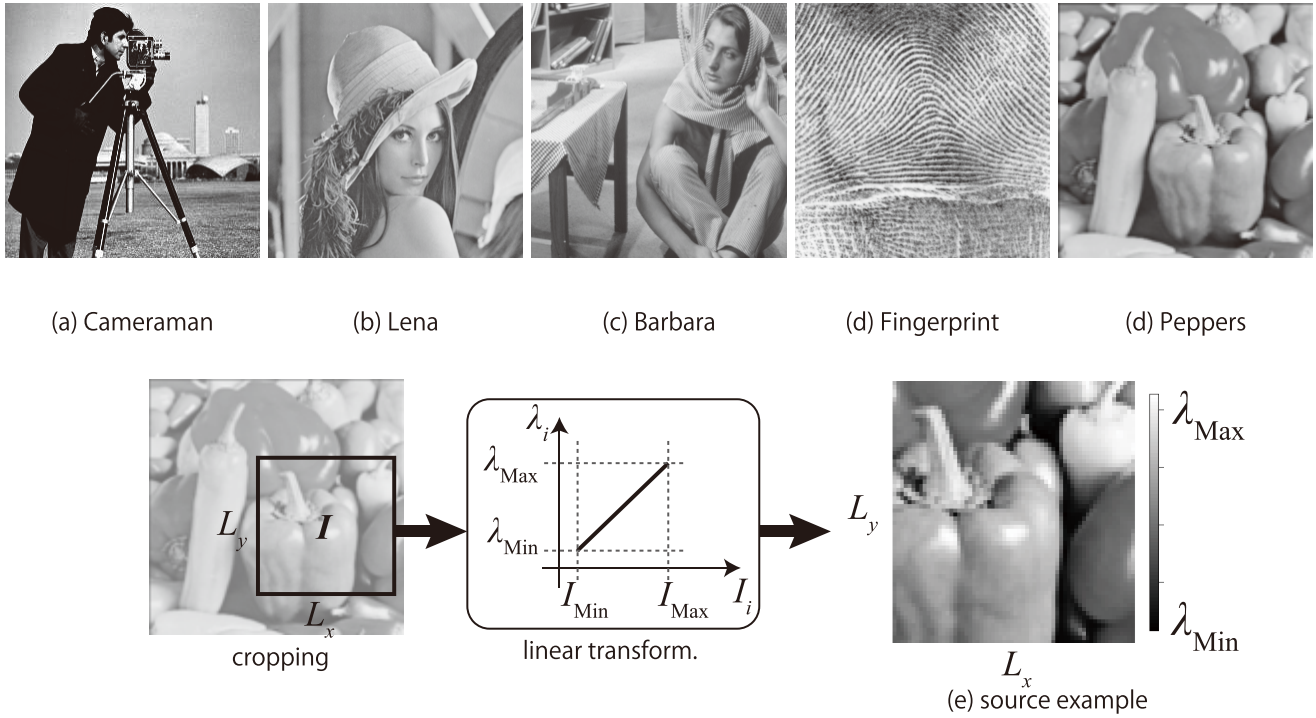
$$\mathrm{PSNR}(\lambda, \lambda^*) = 10\,\log_{10}\frac{(\max\lambda^* - \min\lambda^*)^2}{\mathrm{MSE}(\lambda, \lambda^*)}, \qquad (49)$$

$$\mathrm{MSE}(\lambda, \lambda^*) = \frac{1}{M}\sum_i(\lambda_i - \lambda_i^*)^2, \qquad (50)$$

where M means the image size $M = L_x L_y$. **Table 2** shows the results. In each condition, we show PSNRs of the noise corrupted values $\mathrm{PSNR}(z, \lambda^*)$, and restoration results $\mathrm{PSNR}(\hat{\lambda}, \lambda^*)$ for algorithms 1 and 2. We can see the PSNRs of restored image of algorithms 1 and 2 take very similar values, and the PSNRs are improved from the observed one.

At last, we compare the computational time costs of algorithms 1 and 2. The computer simulation is carried out on the Apple MacBook Pro, which has 2.7 GHz Intel Core i7, 8 GBytes 1.6 GHz DDR3 memory, with OS version is OS X 10.9.2. We implement these algorithms with R language which version is 3.0.0.

Figure 4 shows the result. The horizontal axis shows the image

(a) Cameraman   (b) Lena   (c) Barbara   (d) Fingerprint   (d) Peppers

λ_i

λ_{Max}

λ_{Min}

I_{Min} I_{Max} I_i

L_y I

L_x

cropping

linear transform.

λ_{Max}

L_y

L_x

λ_{Min}

(e) source example

Fig. 3   Images for evaluation: From (a) to (d) show the well known evaluation image obtained from standard image database [5]. (e) shows an example of source image cropped from "Peppers." The cropping size are L_y for vertical and L_x for horizontal. The image intensity is controlled by the parameters λ_{Min} and λ_{Max} shown as Eq. (48).
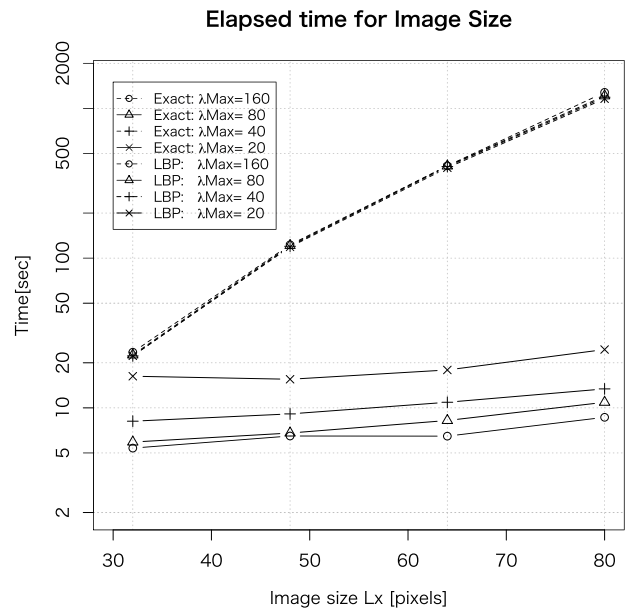
Table 1   Comparison of inferred hyperparameters α between the algorithm 1 and 2. We use a cropped image of "cameraman" with patch size $L_x = L_y = 64$, and apply several contrast conditions, that is, $\lambda_{\text{Max}} = \{20, 40, 80, 160\}$. Each result is evaluated with 10 trials.

	Algorithm 1	Algorithm 2
$\lambda_{\text{Max}} = 20$	54.06 ± 1.39	53.94 ± 1.38
$\lambda_{\text{Max}} = 40$	16.34 ± 0.50	16.34 ± 0.50
$\lambda_{\text{Max}} = 80$	7.61 ± 0.09	7.61 ± 0.09
$\lambda_{\text{Max}} = 160$	4.61 ± 0.07	4.61 ± 0.07

Table 2   Comparison of restored image quality with PSNR Eq. (49): Each column shows the PSNR for corrupted image, for restored with the algorithm 1 and 2 respectively. The trial conditions are same with the one for Table 1.

	Corrupted	Algorithm 1	Algorithm 2
$\lambda_{\text{Max}} = 20$	16.82 ± 0.12	18.62 ± 0.13	18.62 ± 0.13
$\lambda_{\text{Max}} = 40$	20.44 ± 0.15	21.86 ± 0.12	21.86 ± 0.12
$\lambda_{\text{Max}} = 80$	23.86 ± 0.18	24.52 ± 0.18	24.52 ± 0.18
$\lambda_{\text{Max}} = 160$	27.30 ± 0.15	27.58 ± 0.16	27.58 ± 0.16

Elapsed time for Image Size

Exact: λMax=160
Exact: λMax= 80
Exact: λMax= 40
Exact: λMax= 20
LBP: λMax=160
LBP: λMax= 80
LBP: λMax= 40
LBP: λMax= 20

Time[sec]

Image size Lx [pixels]

Fig. 4   Comparison of restoring times: The vertical axis shows the log-scaled elapsed time [sec] for the restoration, which means convergence of the EM algorithms. The horizontal one shows the image size $L_x(= L_y)$ [pixels], which means image size of one side. Each solid line shows the result for the LBP restoration, and dashed line shows the result for the exact solution for the algorithm 1. For both LBP and exact method, we investigate the convergence time for several noise levels, which are $\lambda_{\text{Max}} = \{20, 40, 80, 160\}$.

scale of one side L_x, which equals to the other side L_y. Thus, the horizontal axis shows the square root of the total number of image pixels. The vertical one shows the elapsed time of seconds for the algorithms within the log scale. In the figure, the solid lines show the results for the LBPs approximation (algorithm 2), and the dashed one show the exact solution (algorithm 1) which solve the inverse of accuracy matrix in each E-step. In the exact solution expressed as the dashed lines, the larger the image size is, the larger elapsed time becomes. We can also see the corruption level does not affect to the calculation cost for the algorithm 1. On the contrary, in the LBP solutions, the calculation cost looks insensitive to the image size. Instead, the LBP solutions are affected to the corruption level, when it becomes large, which means small

λ_{Max}, the more calculation cost is required. However, in the large scale image, the LBP solutions has advantage to the exact solutions.

From these comparisons, we show the LBP approximated solution (algorithm 2) has similar performance to the algorithm 1, in
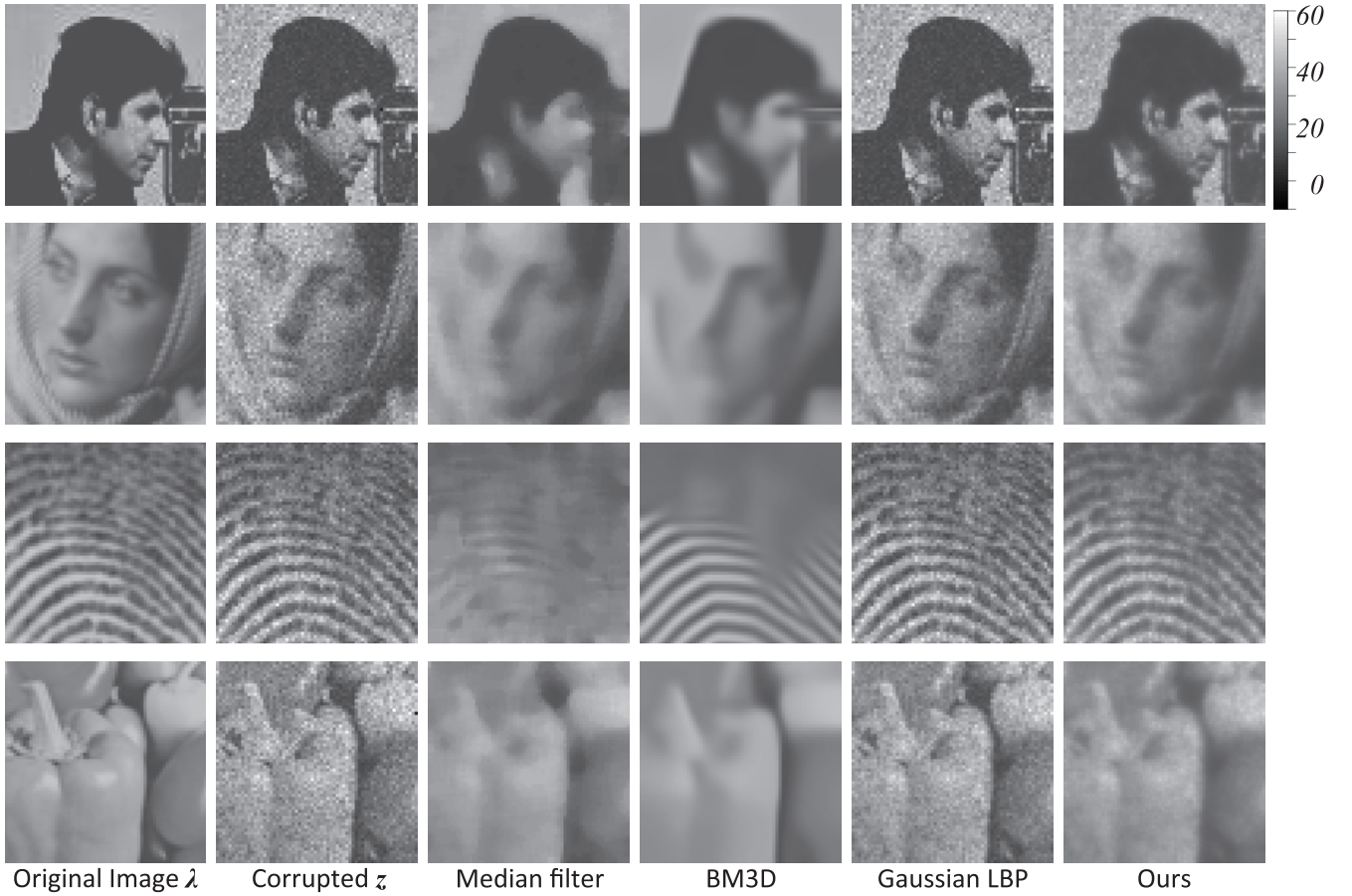
Fig. 5 Comparison of restored image examples: The first column shows the original image λ with $L_x = L_y = 64$. The second shows the corrupted images through the Poisson observation z where the contrast parameter $\lambda_{\mathrm{Max}} = 40$. The third shows the image restoration results with Median filter with 3×3 size. The fourth shows the results of BM3D [2]. The fourth shows the results of Gaussian LBP [6]. The fifth shows our latent variational method results.

the meaning of the image restoration quality in spite of the lower computational time cost. Thus, we only discuss about the LBP approximated solution following section.

3.2   Comparison with Other Image Restoration Methods

In this section, we compare the restored image quality with the algorithm 2 and other image restoration methods, which are conventional median filter, BM3D method [2], and the Gaussian LBP (GLBP) solution [6]. In these restoration methods, we assume restoration of the parameter $\{\lambda_i\}$ from the observed values $\{z_i\}$. Thus, we apply the median filter, BM3D and GLBP methods to the observation $\{z_i\}$ in order to obtain the restoration result $\{\lambda_i\}$. Especially, in the Gaussian LBP solution, the observation of the image assumes the pixel values $\{z_i\}$ are the result of the observation of the corresponding parameters $\{\lambda_i\}$ through the Gaussian channel, that is,

$$p(z \mid \lambda) = \prod_i p(z_i \mid \lambda_i, \beta_G), \tag{51}$$

$$p(z_i \mid \lambda_i, \beta_G) = \sqrt{\frac{\beta_G}{2\pi}} \exp\left(-\frac{\beta_G}{2}(z_i - \lambda_i)^2\right) \tag{52}$$

instead of Eq. (1). In the GLBP solution, we also adopt the GMRF as the prior

$$p(\lambda \mid \alpha_G, h_G) = \frac{1}{Z_{\mathrm{Gpri}}(\alpha_G, h_G)} \exp\left(-H_{\mathrm{Gpri}}(\lambda \mid \alpha_G, h_G)\right), \tag{53}$$

instead of Eq. (16), where

$$H_{\mathrm{Gpri}}(\lambda \mid \alpha_G, h_G) = \frac{\alpha_G}{2} \sum_{(i,j)} (\lambda_i - \lambda_j)^2 + \frac{h_G}{2} \sum_i \lambda_i^2, \tag{54}$$

$$Z_{\mathrm{Gpri}}(\alpha_G, h_G) = \sum_\lambda \exp\left(-H_{\mathrm{Gpri}}(\lambda \mid \alpha_G, h_G)\right). \tag{55}$$

The hyperparameters α_G, β_G, and h_G are inference parameters which are solved by the EM algorithm using LBP [6].

Figure 5 shows the comparison of result examples. In the figure, we show several cropped images with $L_x = L_y = 64$ from "cameraman," "barbara," "fingerprint," and "peppers." In the evaluation, we fix the contrast control parameter $\lambda_{\mathrm{Max}} = 40$ that controls the noise strength in the Poisson corruption. The first column shows the original images, and the second one shows the Poisson corrupted images. The third shows the conventional median filter restoration results with the size of 3×3 filter. The fourth shows the restoration results with BM3D method [2]. The BM3D method is one of the state-of-art method for the AWGN, however, in this case, the variance of noise is not uniform in the image so that the estimation of thresholding parameter looks insufficient. The fifth shows the restoration results with GLBP method, and

Table 3 Restoration quality evaluation with PSNR [dB] in Fig. 5 and median filters.

	Corrupted	Median	BM3D	GLBP	Ours
Cameraman	20.24	16.76	18.75	20.72	**21.92**
Barbara	19.18	22.10	21.91	22.48	**23.89**
Fingerprint	19.47	13.83	18.75	19.86	**21.85**
Peppers	18.44	21.38	21.63	21.03	**22.81**

the sixth shows the one with our LBP method. We can see our result images are more smooth than those of the GLBP results. **Table 3** shows the PSNR evaluations for each original image. Our latent variational method shows better restoration results rather than those of the GLBP solutions. We also restoration results with conventional median filter with 3×3 in Table 3. The median filter restoration make the image too much smooth, so that the PSNR evaluation tends to be the small value.

In order to compare the quantitative restoration evaluations, we introduce the following improvement of PSNR (ISNR) index for two type of restoration results λ_1 and λ_2,

$$\mathrm{ISNR}(\lambda_1, \lambda_2; \lambda^*) = \mathrm{PSNR}(\lambda_2, \lambda^*) - \mathrm{PSNR}(\lambda_1, \lambda^*)$$
$$= 10 \log \frac{\mathrm{MSE}(\lambda_1, \lambda^*)}{\mathrm{MSE}(\lambda_2, \lambda^*)}, \qquad (56)$$

where λ^* means the ground-truth source image. This index shows the improvement of the λ_2 against the λ_1 in the meaning of PSNR. The positive index shows the improvement of the method λ_2 from λ_1.

We evaluate ISNR between the noised image z and our results λ_{ours} that is $\mathrm{ISNR}(z, \lambda_{\mathrm{ours}})$, and also evaluate the one with other restoration results, that is median filter result λ_{Med}, BM3D restoration result λ_{BM3D}, and GLBP result λ_{GLBP}, with our result, that is $\mathrm{ISNR}(\lambda_{\mathrm{Med}}, \lambda_{\mathrm{ours}})$ $\mathrm{ISNR}(\lambda_{\mathrm{BM3D}}, \lambda_{\mathrm{ours}})$ $\mathrm{ISNR}(\lambda_{\mathrm{GLBP}}, \lambda_{\mathrm{ours}})$ respectively. In the image preparation, we crop 10 patch images with the size of $L_x = L_y = 64$ from random locations of each original image shown in Fig. 3. Thus, the total number for the evaluation images is 50 image patches. In the evaluation, we apply several contrast parameter cases, that is $\lambda_{\mathrm{Max}} = \{20, 40, 60, 80, 100, 120, 140, 160\}$.

Figure 6 shows the improvement of our results for the corrupted images z. The contrast parameter λ_{Max}, which denotes the horizontal value of the figure, controls the noise strength of the Poisson noise. In the figure, the boxplot shows the median, which are described as the thick line in the box, with quantiles for each contrast levels. we obtain $2 \sim 3$ [dB] improvement from the corrupted image in the meaning of the median. From the range of the control parameter $\lambda_{\mathrm{Max}} \geq 40$, our method shows good performance, however, the result shows large quantile variance in the $\lambda_{\mathrm{Max}} = 20$. The reason of the large variance comes from the input image property. The low improvement results only come from the "fingerprint" input image, so that, high spatial frequency with low contrast image might prevent restoration.

We also show the comparison results with applying conventional median filter. We apply 3×3 median filter for the noised image z, and evaluate the PSNR with the original Poisson parameters λ. **Figure 7** shows the results. The horizontal axis shows the same range of Fig. 6, and the vertical one shows the ISNR with the range of $[-4, 15]$ [dB]. In many cases, applying the median
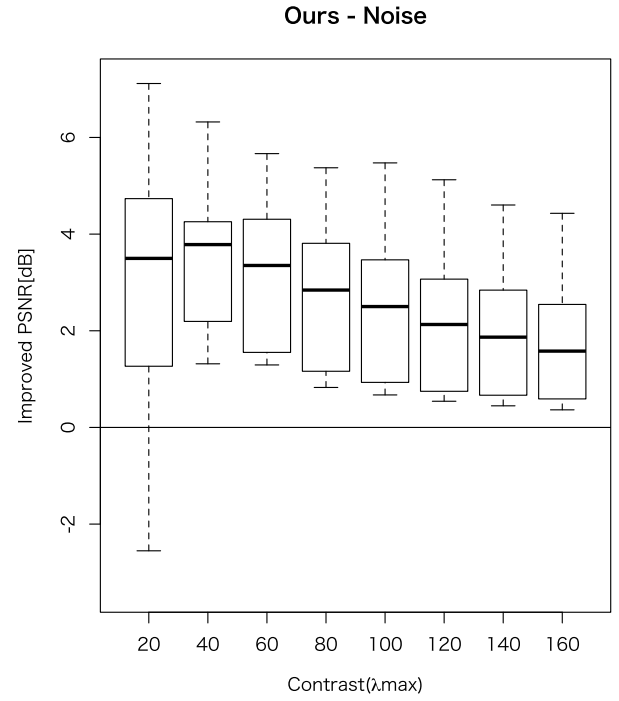
Fig. 6 PSNR improvement of our results λ_{ours} from the observed image z. The horizontal axis shows the contrast parameter λ_{Max}. The vertical one shows the ISNR value whose range is $[-4, 7]$ [dB]. The box-plot shows the medians with quantiles.
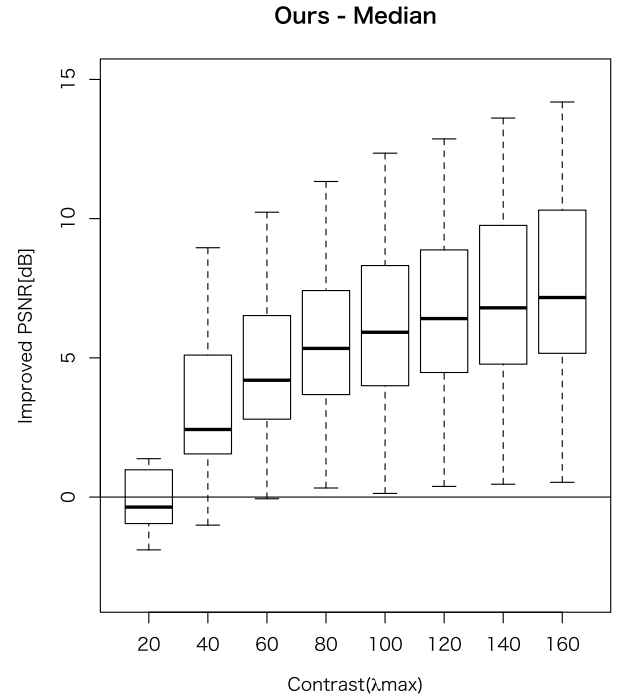
Fig. 7 PSNR improvement of our results λ_{ours} from the median filter results λ_{Med}. The range of vertical axis is $[-4, 15]$ [dB].

filter, the restored image becomes too much smooth, so that, the PSNR evaluation becomes worse.

Figure 8 shows the improvement of our results for the BM3D restoration [2]. The horizontal axis is identical to Figs. 6 and 7, and the vertical shows the ISNR with range of $[-4, 7]$ [dB]. In the figure, the white circles show the outliers which differ twice the standard deviation or more from each average. In the $\lambda_{\mathrm{Max}} = 20$ case, the BM3D restoration looks better than that of ours. How-
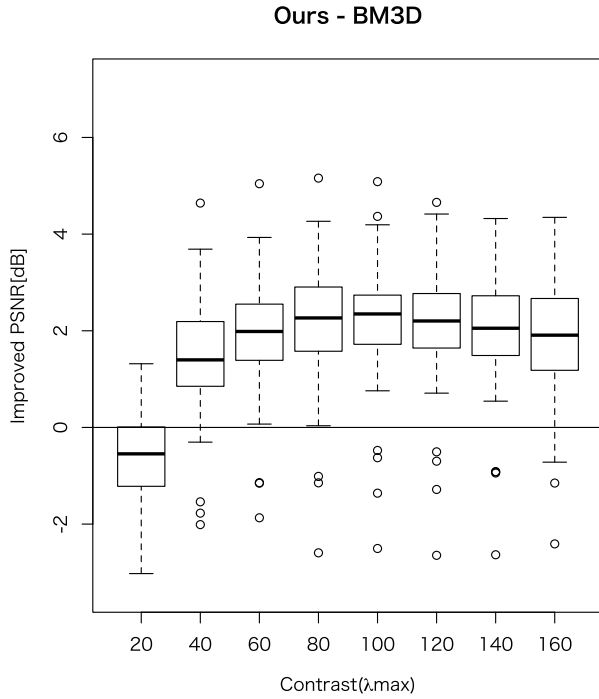
Ours - BM3D

Fig. 8   PSNR improvement of our results λ_{ours} from the BM3D solutions λ_{BM3D}. The range of vertical axis is $[-4, 7]$ [dB]. The white circles show the outliers.
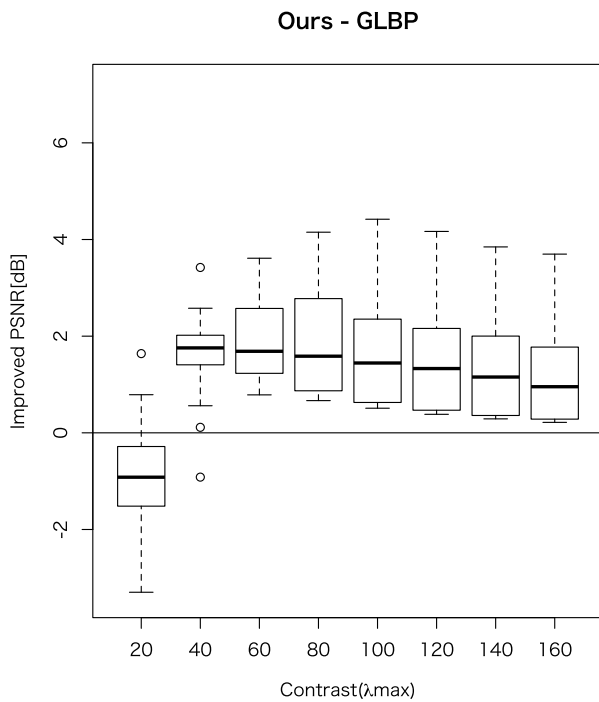
Ours - GLBP

Fig. 9   PSNR improvement of our results λ_{ours} from the GLBP solutions λ_{GLBP}. The range of vertical axis is $[-4, 7]$ [dB]. The white circles show the outliers.

ever, in other cases, our method shows better performance.

Figure 9 shows the improvement of our results for the GLBP restoration. This result also shows our method shows better results than that of those of the GLBP in the $\lambda_{\mathrm{Max}} \geq 40$. In the range of $\lambda_{\mathrm{Max}} \geq 40$, we obtain $1 \sim 2$ [dB] improvement from the GLBP solution in the meaning of the median. On the contrary, in the $\lambda_{\mathrm{Max}=20}$ case, almost all the restoration results of the GLBP are better than those of ours. We consider the results comes from the

accuracy of the hyperparameter inference. In the GLBP method, the noise variance, which is the inverse of the accuracy parameter β_G in Eq. (51), is a single parameter. In the low λ_{Max} case, the variances of each observation z_i might be described as a single value. However, the differences of variances of the observation values z_i would be large when the contrast parameter λ_{Max} becomes large. As the result, the Gaussian model that is Eq. (51) could not describe the observation value z_i within a single hyperparameter β_G. As the result, in many cases of the large λ_{Max}, the inference value of the β_G becomes large, which make low efficacy of the prior.

4. Conclusion

In this study, we propose an image restoration method for Poisson corrupted image. Introducing the latent variable method, we derive the corruption process, which denote the likelihood function, as a Gaussian form. Using Bayesian image restoration framework, we derive the posterior probability for the restoration with introducing GMRF as a prior. The posterior includes several hyperparameters α, h, and latent variables $\{\xi_i\}$. In order to solve the restoration problem with determining these parameters, we construct an algorithm as the algorithm 1 in the manner of the EM method. Thus, our algorithm could determine all the parameters from the observed data z. The algorithm 1 requires $O(M^3)$ computational cost. Hence, in order to accelerate the algorithm, we approximate the posterior mean as the marginal posterior mean, and derive LBP method for hyperparameter inference that is described as the algorithm 2. We introduce the two-body marginalized posterior described as Eq. (46) in order to infer the correlation between connected two units denoted as s_{ij} in Eq. (34). Without the two-body interactions, the inference ignores the correlations, that is, $s_{ij} = 0$ for any indices. It is identical to the naive mean field approximation. The naive mean field approximation, which only applies the single-body marginal posterior, occurs the underestimation of the hyperparameter of α. The correlation s_{ij} update rule is derived as Eq. (34), which only requires the local message, so that the cost for the inference does not increase so much. Solving exact correlation between two units requires considering not only the connected bodies effect but also all the other bodies effect. This is the reason for the requiring the inverse of the accuracy matrix in the EM algorithm. We only consider the two-bodies effect, however, the hyper-parameter inference looks work well, and the restoration performance becomes same or more than the that of the exact solution in the previous work. Hence, we propose the LBP method is a good approximation for our Poisson corrupted image restoration framework.

In the computer simulation, at first, we compare the restoration abilities of the algorithms 1 and 2. The algorithm 2 is a LBP approximated version of the latent variable method described as the algorithm 1. From the viewpoint of the hyperparameter inference and the quality of the restored image, we confirm these two algorithms have similar abilities, however, the computational cost of the algorithm 2 is lower than that of the algorithm 1.

Then, we compare our algorithm with other methods, that is, median filter, BM3D and, GLBP restorations [2], [6]. The BM3D

and GLBP method regards the obtained variables z are observed from the Gaussian noise channel. In these solutions, the very low contrast image ($\lambda_{\text{Max}} = 20$) shows slightly better restoration result than that of ours, however, the larger contrast becomes, the lower the performance becomes. From Eq. (40), our method could express the observation accuracy β_i for each pixel value, however, the GLBP solution has only single hyperparameter β_G in Eq. (51). The BM3D also assumes the variance of the pixels in a image might be denoted as a single parameter in its algorithm. The variance and the mean of the Poisson observation, however, depends on the single parameter λ_i, so that, the assumption of GLBP and BM3D observation might not satisfy. As the result, in the GLBP case, the β_G tends to be overestimation. In the numerical evaluation, our method shows better performance rather than that of the GLBP except $\lambda_{\text{Max}} = 20$. Thus, our results suggests that considering the correct Poisson observation model is important as well as the choosing of the prior.

Our latent variable method evaluate the Poisson likelihood function as the Gaussian form. Thus, in future works, we can easily extend our method into other image restoration framework. For example, when we could express the the parameter values of x as the some linear transformation:

$$x = As \qquad (57)$$

where A and s are the transformation matrix and the expression vector respectively. Then, we could substitute it into Eq. (11) and consider the prior for the transformed vector s such like sparse prior, such like K-SVD [3].

References

[1]   Rudin, L., Osher, S. and Fatermi, E.: Nonlinear total variation based noise removal algorithms, *Physica D.*, Vol.60, pp.259–268 (1992).

[2]   Dabov, K., Foi, A., Katkovnik, V. and Egiazarian, K.: Image Denoising by Sparse 3D Transform-Domain Collaborative Filtering, *IEEE Trans. Image Processing*, Vol.16, No.8, pp.2080–2095 (2007).

[3]   Aharon, M., Elad, M. and Bruckstein, A.: K-SVD: An Algorithm for Desiging Overcomplete Dictionaries for Sparse Representation, *IEEE Transacions on Signal Processing*, Vol.54, No.11, pp.4311–4322 (2006).

[4]   Tanaka, K.: Statistical-mechanical approach to image processing, *Journal of Physics A: Mathematical and General*, Vol.35, No.37, pp.R81–R150 (2002).

[5]   Portilla, J., Strela, V., Wainwright, M.J. and Simoncelli, E.P.: Image denoising using scale mixtures of gaussians in the wavelet domain, *IEEE Trans. Image Processing*, Vol.12, pp.1338–1351 (2003).

[6]   Tanaka, K. and Titterington, D.M.: Statistical trajectory of an approximate EM algorithm for probabilistic image processing, *Journal of Physics A: Mathematical and Theoretical*, Vol.40, No.37, p.11285 (2007).

[7]   Shouno, H. and Okada, M.: Bayesian Image Restoration for Medical Images Using Radon Transform, *Journal of the Physical Society of Japan*, Vol.79, p.074004 (2010).

[8]   Shouno, H., Yamasaki, M. and Okada, M.: A Bayesian Hyperparameter Inference for Radon-Transformed Image Reconstruction, *International Journal of Biomedical Imaging*, ArticleID 870252, 10 pages (2011).

[9]   Rodrigues, I., Sanches, J. and Bioucas-Dias, J.: Denoising of medical images corrupted by Poisson noise, *IEEE International Conference on Image Processing*, pp.1756–1759 (2008).

[10]   de Decker, A., Lee, J.A. and Velysen, M.: Vairance stabilizing transformations in patch-based bilateral filters for Poisson noise image denoising, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp.3673–3676 (2009).

[11]   Figueiredo, M.A.T. and Bioucas-Dias, J.: Restoration of Poissonian Images Using Alternating Direction Optimization, *IEEE Trans. Image Processing*, Vol.19, No.12, pp.3133–3145 (2010).

[12]   Hadj, B.S., Blanc-Feraud, L., Aubert, G. and Engler, G.: Blind restoration of confocal microscopy images in presence of a depth-variant blur and Poisson noise, *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp.915–919 (2013).

[13]   Ono, S. and Yamada, I.: Poisson image restoration with likelihood constraint via hybrid steepest decent method, *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp.5929–5933 (2013).

[14]   Palmer, J.A., Kerutz-Delgado, K., Wipf, D.P. and Rao, B.D.: Variational em algorithm for non-gaussian latent variable models, *Advances in Neural Information Processing Systems 18*, Vol.18, pp.1059–1066, MIT Press (2005).

[15]   Bishop, C.M.: *Pattern Recogition and Machine Learning*, Springer (2006).

[16]   Seeger, M.W.: Bayesian inference and optimal design for the sparse linear model, *Journal of Machine Learning Research*, Vol.9, pp.759–813 (2008).

[17]   Watanabe, K. and Okada, M.: Approximate bayesian estimation of varying binomial process, *IEICE Trans. Inf. Syst.*, Vol.E94-A, No.12, pp.2879–2885 (2011).

[18]   Sawada, H., Kameoka, H., Araki, S. and Ueda, N.: Efficient algorithms for multichannel extensions of Itakura-Saito nonnegative matrix factorization, *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp.261–264 (2012).

[19]   Shouno, H. and Okada, M.: Poisson observed image restoration using a latent variational approximation with gaussian mrf, *PDPTA*, pp.201–208, PDPTA (2013).

[20]   Shouno, H. and Okada, M.: Acceleration of poisson corrupted image restoration with loopy belief propagation, *PDPTA*, Vol.1, pp.165–170, PDPTA (2014).

[21]   Dempster, A.P., Larid, N.M. and Rubin, D.B.: Maximum likelihood estimation from incomplete data via the em algorithm, *Journal of the Royal Statistical Society: Series B* (*Statistical Methodology*), Vol.39, pp.1–38 (1977).

[22]   Mackay, D.J.C. and Cavendish Laboratory: Hyperparameters: optimize, or integrate out, *Maximum Entropy and Bayesian Methods, Santa Barbara*, pp.43–60, Kluwer (1996).

[23]   Murphy, K.P., Weiss, Y. and Jordan, M.I.: Loopy belief propagation for approximate inference: An empirical study, *Proc. 15th Conference on Uncertainty in Artificial Intelligence*, UAI'99, pp.467–475, San Francisco, CA, USA, Morgan Kaufmann Publishers Inc. (1999).

Hayaru Shouno was born in 1968.   He received his M.E. and Ph.D. from Osaka University in 1994 and 1999. His current research interest is image processing and artificial neural network. He is a member of IEEE, IEICE, and IPSJ.