Variational Bayesian Image Restoration with Transformation Parameter Estimation

MOTOHARU SONOGASHIRA[1,a]   TAKUYA FUNATOMI[2,b]   MASAAKI IIYAMA[1,c]   MICHIHIKO MINOH[3,d]

Abstract: Image restoration is essential for high-quality photography and used in many applications. Its objective is to estimate a clean image from one or more degraded images affected by problem-specific transformation with an unknown parameter. Traditionally, the estimation of this transformation parameter has limited the effectiveness of restoration, since it can be unstable in the presence of severe or complex degradation. In this work, we overcome this difficulty by estimating multiple parameters simultaneously via stable joint inference enabled by variational Bayes. Specifically, we developed novel methods for three problems of restoration: multiframe denoising, deblurring, and devignetting. The established methodology can address various restoration problems in a unified Bayesian framework, and also achieves high quality restoration for a wide range of images owing to stable and fully automatic parameter tuning, thereby extending the real-world applicability of restoration.

Keywords: Image restoration, variational Bayes, multiframe denoising, deblurring, devignetting

| (a) Noise | (b) Blur | (c) Vignetting |

Fig. 1: Image degradation types.

Fig. 2: Image degradation process.

1. Introduction

Image restoration is a long-studied topic in the field of image processing [1]. Essentially, its goal is to remove various types of degradation from images. Noise (Fig. 1(a)) is one of the most typical types of image degradation, which randomly alters the light intensity values of image pixels. Blur (Fig. 1(b)) is another type of degradation that obscures image details, such as contours and textures. Vignetting (Fig. 1(c)) is also common in real images, which makes the periphery of an image darker than its center. Digital image restoration for reduction of such undesirable degradation effects is important for high-quality photography, and thus has a wide range of applications [2].

In general, a problem of image restoration can be considered as estimation of a clean image without degradation from one or more degraded images [1], assuming the degradation process depicted in Fig. 2. Here, each degraded image is affected by image

transformation, which describes optical effect in image capturing such as blur. To constrain the possible range of transformation, it is common to assume that the transformation can be uniquely determined by an image-dependent parameter. While this transformation parameter has traditionally been assumed to be known, its true value is often unknown in reality; thus, to perform restoration, we need to guess its value. The most näive approach is manual parameter tuning, which is obviously time-consuming and often impractical. Alternatively, a reasonable value of a transformation parameter may be found automatically by estimation from degraded images. However, automatic parameter tuning is not trivial in the real world, where images are affected by a variety of transformation, and also by noise. Since such degradation can damage informative image features that help distinguish transformation, pre-estimation of a transformation parameter before restoration can be quite unstable and often inaccurate, leading to poor restoration quality. Although restrictive parameterization schemes have been proposed to alleviate this difficulty [3], they have been found too inflexible to represent the variety of transformation in the real world. Therefore, the effectiveness of traditional image restoration has been strictly limited due to unknown transformation parameters in image degradation.

In this work, we aim to overcome the difficulty with unknown transformation parameters in image restoration. To this end, we follow the approach of automatic parameter tuning to achieve

1   Kyoto University, Kyoto, Kyoto 606–8501, Japan
2   Nara Institute of Science and Technology, Ikoma, Nara 630–0192, Japan
3   RIKEN, Wako, Saitama 351–0198, Japan
a)   sonogashira@mm.media.kyoto-u.ac.jp
b)   funatomi@is.naist.jp
c)   iiyama@mm.media.kyoto-u.ac.jp
d)   michihiko.minoh@riken.jp

higher applicability than manual parameter tuning; however, we consider multiple unknown parameters not separately but jointly, thereby dealing with transformation, noise, and other unknown factors simultaneously. Specifically, we establish a methodology to jointly estimate a transformation parameter and a clean image from degraded images by making full use of the family of statistical techniques called *variational Bayes* (VB) [4], which enables stable statistical inference on multiple latent variables in consideration of their dependency and uncertainty. To realize the VB methodology of image restoration and demonstrate its effectiveness for practical imaging scenarios, we developed novel methods for three important problems of image restoration: multiframe denoising, deblurring, and devignetting. Through extensive experiments, we confirmed that the proposed VB methods are highly effective and often outperform previous methods, especially when degradation is severe, dealing with a variety of transformation without manual parameter tuning.

The proposed VB methodology of image restoration, which is the main contribution of this work, has various advantages over traditional image restoration. First, the proposed methodology can address various image restoration problems in a unified manner, i.e., by specifying a transformation type for each problem in a Bayesian model and then performing principled VB inference. Thus, it facilitates the development of novel methods for various image restoration problems. Second, the VB approach enables more effective image restoration, i.e., achieves higher image quality, than traditional non-VB approaches. This is because VB restoration can deal with a variety of transformation without resorting to restrictive parameterization, even if images to be restored are affected by severe degradation, owing to stable parameter estimation via joint VB inference. Third, VB enables automatic tuning of other image-dependent parameters than transformation parameters, which have been manually tuned in traditional image restoration. Hence, the proposed VB methods can restore a wide range of images without troublesome manual parameter tuning, thereby extending the applicability of image restoration in the real world.

The rest of this paper is organized as follows. In Section 2, a general overview on VB image restoration is provided, motivating the above-mentioned three problems of image restoration with different transformation parameters. Then, these problems are addressed in the following three chapters, i.e., multiframe denoising in Section 3, deblurring in Section 4, and devignetting in Section 5. Finally, this work is concluded with a discussion on future work in Section 6.

2. Variational Bayesian Image Restoration

2.1 Image Restoration

The imaging system of a digital camera captures an image by recording the intensity of light from a targeted scene using an optical sensor, which consists of photon-detecting cells arranged in a two-dimensional grid on a plane [2]; here, all light rays from a single point in the scene should converge exactly on a single cell, which is represented by a single pixel of the captured image, after being narrowed by an aperture and bent by a lens. In reality, however, this expectation does not hold when some objects in

the scene have gone out of focus, or when the camera or objects have moved during exposure. Then, the light rays spread over multiple cells, resulting in blur. Moreover, if multiple images of the same scene are taken from different viewpoints or at different times, a single scene point may be looked at by more than one cell, producing displacement between images. Besides, since physical components of the imaging system do not accept incoming light uniformly, the sensitivity to light can also vary between cells, leading to vignetting. Finally, the light intensity recorded at each cell is affected by noise from optical, electric, and electronic components of the imaging system, before being converted into a digital pixel value.

Image restoration is the inverse problem of estimating a clean image without degradation from one or more degraded images. Considering the degradation process of the typical imaging system explained above, each degraded image is commonly modeled as a version of the original clean image affected by linear transformation and additive noise [1]; here, the transformation describes optical effect that light has undergone before arriving at the sensor, e.g., displacement, blur, and vignetting. Since each digital image is a two-dimensional array of pixels with light intensity values, it can mathematically be represented as a vector by stacking the pixel values; in the multiframe case, the vectors of multiple degraded images are further stacked to form a single vector. Similarly, an arbitrary linear transformation on such a vector can be represented as a matrix. Using these notations, the standard degradation model of image restoration can be formulated as follows [1]:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the vectors of the clean and degraded images, respectively, $\mathbf{H}$ is the matrix of an arbitrary linear transformation, and $\mathbf{n}$ is a noise vector (commonly assumed to follow a zero-mean Gaussian distribution). Here, each row of $\mathbf{y}$ represents a single pixel of the degraded images, which is the sum of the pixels of $\mathbf{x}$ weighted by the corresponding row of $\mathbf{H}$ plus the noise component of $\mathbf{n}$. Image restoration basically aims to invert this degradation process to recover the clean image $\mathbf{x}$ from the degraded images $\mathbf{y}$.

The fundamental observation in this work is that image restoration problems for various degradation types can be formulated in a unified manner, i.e., by letting the image transformation in the standard degradation model represent different operations. For example, in multiframe denoising, where more than one image are used to improve noise removal, $\mathbf{H}$ is considered to be a warping operation, which aligns the desired clean image $\mathbf{x}$ with respect to each of the degraded images $\mathbf{y}$ (Fig. 3(a)). Here, to produce the pixel represented by each row of $\mathbf{y}$, the corresponding row of $\mathbf{H}$ interpolates the value of $\mathbf{x}$ at a single spatial point on the image plane, from which the pixel value of $\mathbf{y}$ (before noise $\mathbf{n}$ is added) is originated. This effectively describes the displacement of a single scene point projected onto the image plane, which must be compensated for to enable the use of multiple images. Another example is deblurring, where $\mathbf{H}$ is a convolution operation (Fig. 3(b)). Here, at each pixel of $\mathbf{y}$, the corresponding row of $\mathbf{H}$ averages the pixel values of $\mathbf{x}$ in a small spatial neighborhood, i.e., multiplies

(a) Multiframe denoising.

(b) Deblurring.

(c) Devignetting.
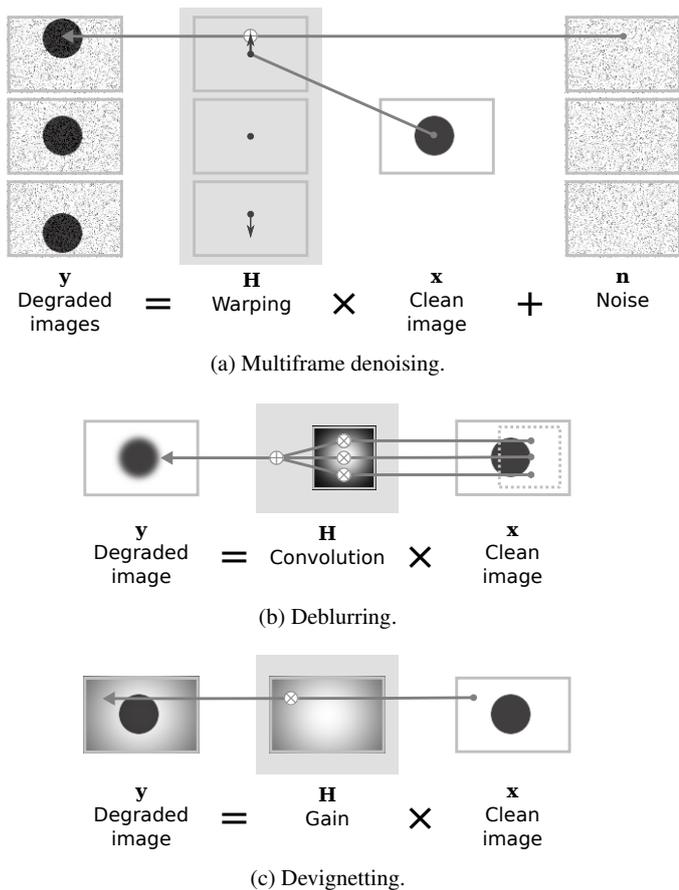
Fig. 3: Degradation models in image restoration problems.

the original values with weights and sums them up. Physically, this operation mixes light rays from different scene points into a single cell, resulting in a blurry pixel. Furthermore, devignetting is the problem where **H** represents a pixelwise gain operation, which decreases the light intensity value at each pixel (Fig. 3(c)). In this case, **H** becomes a diagonal matrix, each of whose rows has only one element that holds a multiplicative factor applied to the pixel at the same row of **x**, yielding the corresponding pixel of **y**. This factor measures the amount of brightness reduction at a sensor cell due to vignetting, which makes the degraded pixel darker than the original clean pixel. To constrain **H** as the intended operation for each problem, the large matrix **H** is commonly parameterized with a smaller number of elements, which is denoted by vector **w** in this work. In addition, regularization on **w** is often performed to alleviate the ill-posedness of parameter estimation.

2.2 Variational Bayes

Bayesian inference provides a principled way to estimate unknown or uncertain parameters from known parameters given as observation. To perform inference, a problem-specific Bayesian model that describes the relationship between known and unknown parameters is first constructed, treating them as observed and latent variables, respectively, and defining their probability distributions. Then, an algorithm of Bayesian inference is derived, whose typical objective is to find out the most probable value of a latent variable, i.e., to maximize its posterior

probability given the observed variables. To obtain the necessary posterior distribution when more than one latent variable is present, marginalization, i.e., integration of a joint probability distribution with respect to other latent variables, needs to be performed. However, this integration can be quite difficult and often intractable (at least analytically) in practical problems, due to complex dependency between parameters.

The main objective of variational Bayes is to resolve such mutual dependence of multiple latent variables in Bayesian inference. Technically, VB bypasses intractable marginalization by approximating posterior distributions of latent variables [4]. To obtain approximate posteriors that are as close to the exact ones as possible, the Kullback-Leibler divergence between the exact and approximate distributions are typically minimized. The resulting algorithm reduces to iterative update of parameterized approximate posteriors, which is guaranteed to converge [5], starting at some initial estimates. In this joint inference, dependency between all variables are naturally considered, since each posterior is updated using estimates of other parameters, including those that are initially unknown. Note that VB can be interpreted as a generalization of traditional techniques such as *maximum a posteriori* (MAP) and *expectation maximization* (EM), which assume that all or some posteriors are degenerate, effectively treating them as deterministic variables and thus ignoring their uncertainty. Unlike these techniques, however, VB can yield estimates of full posterior distributions, i.e., not only point estimates but also information on their uncertainty, which contributes to the robustness of inference [4].

Bayesian inference has been proved to be effective for image restoration, which deals with random degradation factors such as noise [6], [7], [8]. In Bayesian image restoration, observed variables consist of one or more degraded images, and latent variables include a clean images, a transformation parameter, and other image-dependent parameters, e.g., strength of noise and smoothness of the clean image; here, the distribution of the degraded image should incorporate assumptions on observation, including the specific degradation model for the problem, while the prior distributions of the clean image and transformation parameter should reflect additional assumptions on these unknowns for regularization. In traditional image restoration, the parameters other than the desired clean image have been assumed to be known, or separately estimated in advance of restoration; by contrast, VB image restoration performs joint inference on these variables to obtain their approximate posteriors, effectively estimating all unknown parameters simultaneously. After this inference, a restored image can be obtained by maximizing the resulting approximate posterior of the clean image.

3. Variational Bayesian Multiframe Denoising

3.1 Introduction

Classical image restoration assumes that each degraded image is affected by noise and blur, and aims to recover a clean image from it [1], mainly focusing on denoising rather than deblurring. To effectively remove noise, the majority of traditional methods have parameters that control restoration, e.g., a parameter for image smoothing, which distinguishes a noise-free clean

image from a degraded image and thereby regularizes the ill-posed restoration problem. Since the optimal value of such a parameter varies between images, it should be carefully tuned for each image to achieve high-quality restoration. Although there have been studies on automatic tuning of smoothing parameters [9], [10], [11] along with several attempts to simultaneously estimate multiple parameters via MAP [12], [13], [14], [15], [16], such point-estimation methods can yield poor parameter estimates in the presence of degradation.

To overcome the difficulty of parameter tuning, variational Bayes has recently been introduced to denoising [6], [17], [18]; here, VB can jointly estimate a clean image and other parameters from degraded images, thus enabling fully automatic parameter tuning. It is also known that VB methods often outperform traditional non-VB methods in the presence of severe degradation [6]. However, state-of-the-art VB denoising has focused on the single-frame setting, where one clean image is recovered using only one degraded image. In practice, information from a single image is often insufficient for high-quality restoration, especially when noise is strong [19]. Therefore, despite the practical advantage of automatic parameter tuning, the effectiveness of VB denoising has been strictly limited.

In this section, multiframe denoising is addressed via a VB approach. Unlike previous single-frame methods, the proposed multiframe method exploits rich information in multiple images to restore a single image, thus achieving higher image quality. This VB approach also maintains the key advantage over traditional non-VB approaches, i.e., the ability to automatically tune parameters. More specifically, we make full use of VB techniques to enable Bayesian inference for multiframe denoising, where multiple parameters are jointly estimated, including one for image warping that compensates for displacement between images, effectively performing image registration simultaneously with image restoration; here, the warping is flexibly parameterized with an optical flow, which can handle a wide range of displacement in the real world [20], [21]. Through experiments, the effectiveness of the proposed multiframe method was verified in comparison with its single-frame counterpart, as well as the advantage of the VB approach over non-VB approaches.

3.2 Model

As in the previous work on single-frame VB denoising [6], [17], [18], we assume that degraded images are affected by additive zero-mean Gaussian noise, and additionally blur by a known convolution kernel. Furthermore, in multiframe denoising, displacement between the images also needs to be considered; hence, we introduce image warping, which geometrically aligns one image with another, effectively performing registration. More specifically, each degraded image is assumed to be generated by warping, blurring, and then adding noise to a single clean image. To determine an image to be restored, the clean image is required to be a noise-free and blur-free version of one of the degraded images, which is called the *reference image*.

Under these assumptions, we estimate the clean image from the degraded images, considering all other unknown parameters including an optical flow as a warping parameter. To describe
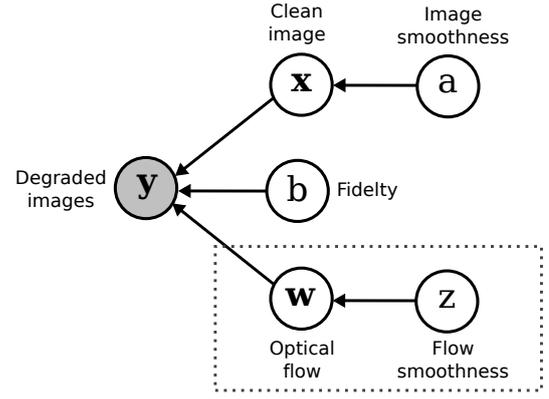
Fig. 4: Bayesian model for multiframe denoising. Gray and white circles represent observed and latent variables, respectively. The part enclosed by dots is specific to the multiframe case.

their relationship, we define a Bayesian model by treating these parameters as random variables and assigning a probability distribution to the each variable, as described in the following. The resulting Bayesian model is shown graphically in Fig. 4.

3.2.1 Degraded Images

Let $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^m$ be vectors that represent the degraded images and the clean image, respectively, where n and m are the total numbers of their pixels. Then, the following equality describes the relationship between $\mathbf{x}$ and $\mathbf{y}$:

$$\mathbf{y} = \mathbf{H}_0\mathbf{H_w}\mathbf{x} + \mathbf{n}, \qquad (2)$$

where $\mathbf{H_w} \in \mathbb{R}^{n \times m}$ is the warping matrix with parameter $\mathbf{w}$, $\mathbf{H}_0 \in \mathbb{R}^{n \times n}$ is the convolution matrix for blur (assumed to be known), and $\mathbf{n} \in \mathbb{R}^n$ is the zero-mean Gaussian noise vector. Note that $\mathbf{H}_0\mathbf{H_w}$ can be considered as a single transformation matrix to recover the standard degradation model. To make $\mathbf{H_w}$ a flexible warping operator, we parameterize it with optical flow $\mathbf{w} \in \mathbb{R}^{2n}$, which has two velocity components for each degraded pixel. Next, to enable optical flow estimation, $\mathbf{w}$ is related to $\mathbf{H_w}$ and $\mathbf{x}$ by the following linearized optical flow constraint [21], [22]:

$$\mathbf{H_w}\mathbf{x} \simeq \mathbf{I}'\left(\mathbf{w}' \circ \mathbf{x}'\right), \qquad (3)$$

where $\circ$ denotes elementwise multiplication, $\mathbf{I}' \in \mathbb{R}^{n \times 3n}$, and

$$\mathbf{w}' = \mathbf{J}(\mathbf{w} - \mathbf{w}^*) + \mathbf{j}, \qquad (4)$$

$$\mathbf{x}' = \mathbf{K}\mathbf{x}; \qquad (5)$$

here, $\mathbf{J} \in \mathbb{R}^{3n \times 2n}$, $\mathbf{j} \in \mathbb{R}^{3n}$, $\mathbf{K} \in \mathbb{R}^{3n \times m}$, and $\mathbf{w}^* \in \mathbb{R}^{2n}$ is the point at which $\mathbf{w}$ is linearized. Combining Eqs. (2) and (3), the probability distribution of $\mathbf{y}$ is defined as follows:

$$p(\mathbf{y}|\mathbf{x}, \mathbf{w}, b) \propto \prod_{i=1}^{n} \sqrt{b} \exp\left(-\frac{b}{2}\left(\mathbf{y} - \mathbf{H}'(\mathbf{w}' \circ \mathbf{x}')\right)_i^2\right), \qquad (6)$$

where $b \in \mathbb{R}$ is a fidelity parameter (i.e., an inverse noise variance), and $\mathbf{H}' = \mathbf{H}_0\mathbf{I}'$.

3.2.2 Clean Image

Exploiting the redundancy of natural images to regularize the ill-posedness of image restoration, it is commonly assumed that the clean image is smooth except for edges, which can be realized by minimizing the *total variation* (TV) of the image [23]. To

enable a Bayesian treatment of the TV prior, we assume that the gradient magnitude at each pixel of $\mathbf{x}$ follows a zero-mean Laplacian distribution, and define the probability distribution of $\mathbf{x}$ as follows:

$$p(\mathbf{x}|a) \propto \prod_{i=1}^{m} a \exp\left(-a\sqrt{\sum_{k=1}^{2}(\mathbf{G}_k\mathbf{x})_i^2 + \epsilon}\right), \qquad (7)$$

where $a \in \mathbb{R}$ is an image smoothness parameter (i.e., a global smoothing weight), $\mathbf{G}_1, \mathbf{G}_2 \in \mathbb{R}^{m \times m}$ are horizontal and vertical differentiation matrices, respectively, and $\epsilon = 0.001$ is a small constant for numerical stability.

3.2.3 Optical Flow

The multiframe problem is more ill-posed and thus difficult than the single-frame case, since not only the clean image $\mathbf{x}$ but also the optical flow $\mathbf{w}$ have to be estimated. As TV-based smoothing is also known to work well for optical flows between natural images [22], we define the prior distribution of $\mathbf{w}$ to be similar to that of $\mathbf{x}$ as follows:

$$p(\mathbf{w}|z) \propto \prod_{i=1}^{2n} z \exp\left(-z\sqrt{\sum_{k=1}^{2}(\mathbf{F}_k\mathbf{w})_i^2 + \epsilon}\right), \qquad (8)$$

where $z \in \mathbb{R}$ is a flow smoothness parameter, and $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{R}^{2n \times 2n}$ are horizontal and vertical differentiation matrices, respectively.

3.2.4 Weights

For the rest of the variables, i.e., fidelity parameter b and smoothness parameters a, z, no prior information is available, since they can vary a lot between different images to be restored. Thus, we use noninformative conjugate priors [4] for these weight parameters, which are Gamma distributions for Gaussian precisions. Here, we make these priors noninformative by using the shape parameter $\alpha_0 = 0$ and the rate parameter $\beta_0 = \epsilon$.

3.3 Inference

In Bayesian multiframe denoising, the goal is to find out the most probable clean image given a set of degraded images. That is, we maximize the posterior probability of $\mathbf{x}$ given $\mathbf{y}$ as follows:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}). \qquad (9)$$

To obtain the posterior probability in the right-hand side, the other latent variables need to be marginalized out from the joint posterior probability of all latent variables as follows:

$$p(\mathbf{x}|\mathbf{y}) = \int p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})d\mathbf{w}dbdadz. \qquad (10)$$

By Bayes' theorem [4], the joint posterior in the right-hand side can be obtained as

$$p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{w}, b)p(\mathbf{x}|a)p(\mathbf{w}|z)p(b)p(a)p(z). \quad (11)$$

The exact marginalization, however, is intractable due to the complex dependency between the variables; for instance, the degraded images $\mathbf{y}$ is dependent on $\mathbf{x}$, $\mathbf{w}$ and b, as shown in Fig. 4, and the posterior of $\mathbf{x}$ given $\mathbf{y}$ cannot be obtained by marginalizing out $\mathbf{w}$ and b at once.

To handle the intractable marginalization, we invoke a VB

technique called *mean-field approximation* [4]. The basic idea is to resolve the complex dependency by regarding the latent variables as independent after observation, i.e., assuming the conditional independence of their posterior distributions. More specifically, the exact joint posterior is approximated by the product of the independent posteriors of individual latent variables as follows:

$$p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y}) \simeq q(\mathbf{x}, \mathbf{w}, b, a, z) = q(\mathbf{x})q(\mathbf{w})q(b)q(a)q(z). \qquad (12)$$

This factorized approximation yields the approximate posterior of each variable without explicit marginalization, and $q(\mathbf{x})$ provides an approximation of the desired posterior $p(\mathbf{x}|\mathbf{y})$. Then, the exact joint posterior becomes close to the approximate one by minimizing the Kullback-Leibler divergence of $p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})$ from $q(\mathbf{x}, \mathbf{w}, b, a, z)$. This is equivalent to maximizing the variational lower bound of the log-evidence under the model, i.e.,

$$\begin{aligned} &\text{LB}[q(\mathbf{x}, \mathbf{w}, b, a, z)] \\ &= \int q(\mathbf{x}, \mathbf{w}, b, a, z) \ln \frac{p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})}{q(\mathbf{x}, \mathbf{w}, b, a, z)} d\mathbf{x}d\mathbf{w}dbdadz + \text{const}. \end{aligned} \qquad (13)$$

In performing inference under the model in Section 3.2, the nonlinearity in the TV priors of $\mathbf{w}$ and $\mathbf{x}$ prevents direct maximization of LB in Eq. (13). Hence, we invoke another VB technique called *local approximation* [4]. The idea is to approximate the original lower bound LB with further lower bound LB' such that $\text{LB}' \le \text{LB}$, replacing the troublesome TV priors with Gaussian-like priors with additional parameters. To this end, the inequality of arithmetic and geometric means [6] is applied to the logarithms of the TV priors on $\mathbf{w}$ and $\mathbf{x}$, introducing auxiliary parameters $\mathbf{u} \in \mathbb{R}^{2n}$ and $\mathbf{v} \in \mathbb{R}^{m}$, respectively; then, the following new lower bound is obtained:

$$\begin{aligned} &\text{LB}'[q(\mathbf{x}, \mathbf{w}, b, a, z)] \\ &= \int q(\mathbf{x}, \mathbf{w}, b, a, z) \ln \frac{p'(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})}{q(\mathbf{x}, \mathbf{w}, b, a, z)} d\mathbf{x}d\mathbf{w}dbdadz + \text{const}., \end{aligned} \qquad (14)$$

where

$$p'(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{w}, b)p'(\mathbf{x}|a)p'(\mathbf{w}|z)p(b)p(a)p(z) \qquad (15)$$

with modified priors

$$p'(\mathbf{w}|z) \propto \prod_{i=1}^{2n} z \exp\left(-\frac{z}{2}u_i\left(\sum_{k=1}^{2}(\mathbf{F}_k\mathbf{w})_i^2 + \epsilon + u_i^{-2}\right)\right), \qquad (16)$$

$$p'(\mathbf{x}|a) \propto \prod_{i=1}^{m} a \exp\left(-\frac{a}{2}v_i\left(\sum_{k=1}^{2}(\mathbf{G}_k\mathbf{x})_i^2 + \epsilon + v_i^{-2}\right)\right). \qquad (17)$$

Since these priors are Gaussian of $\mathbf{w}$ and $\mathbf{x}$, it is easy to perform VB inference with these priors. In addition, the precision parts of Eqs. (16) and (17) consist of global smoothness parameters (z and a) and auxiliary parameters ($\mathbf{u}$ and $\mathbf{v}$), which effectively work as local (pixelwise) smoothness parameters for edge preserving [6] when adapted to $\mathbf{w}$ and $\mathbf{x}$ through inference.

To perform VB inference, we maximize the modified lower

Algorithm 1 Algorithm of VB inference for multiframe denoising.

1: Input $\mathbf{y}$.
2: Initialize $\boldsymbol{\mu}_\mathbf{x}, \boldsymbol{\Sigma}_\mathbf{x}, \boldsymbol{\mu}_\mathbf{w}$, and $\boldsymbol{\Sigma}_\mathbf{w}$.
3: **repeat**
4: Update $\mathbf{u}$ and $\mathbf{v}$ with Eqs. (18) and (19).
5: Update $\alpha_b, \alpha_a, \alpha_z, \beta_b, \beta_a$, and β_z with Eqs. (23) to (28).
6: Update $\boldsymbol{\mu}_\mathbf{w}$ and $\boldsymbol{\Sigma}_\mathbf{w}$ with Eqs. (30) and (31).
7: Update $\boldsymbol{\mu}_\mathbf{x}$ and $\boldsymbol{\Sigma}_\mathbf{x}$ with Eqs. (33) and (34).
8: **until** convergence.
9: Output $\hat{\mathbf{x}} = \boldsymbol{\mu}_\mathbf{x}$.

---

bound LB' in Eq. (14) with respect to approximate posteriors $q(\mathbf{x}), q(\mathbf{w}), q(b), q(a), q(z)$ (each of which is determined with a few parameters) and auxiliary parameters $\mathbf{u}, \mathbf{v}$. Since the mutual dependency of the latent variables makes it impossible to obtain all the parameters at once in closed form, we iteratively update them one by one until convergence. In the following, update formulas for the auxiliary parameters and approximate posteriors that maximize the lower bound are derived. The resulting algorithm is summarized in Algorithm 1.

3.3.1 Auxiliary Parameters

Taking the derivative of the right-hand side of Eq. (14) with respect to $\mathbf{u}$ and $\mathbf{v}$ and setting it to zero, the auxiliary parameters are obtained as follows:

$$u_i = \frac{1}{\sqrt{\sum_{k=1}^{2} \text{E}\left[(\mathbf{F}_k\mathbf{w})_i^2\right] + \epsilon}}, \quad (18)$$

$$v_i = \frac{1}{\sqrt{\sum_{k=1}^{2} \text{E}\left[(\mathbf{G}_k\mathbf{x})_i^2\right] + \epsilon}}, \quad (19)$$

where we have taken the expectations with respect to the approximate posteriors. Here, v_i is basically the inverse of the local image variation at the ith pixel captured by high-pass filtering via $\mathbf{G}_1, \mathbf{G}_2$. Thus, the coefficients of the auxiliary parameters become small around large variations. This adaptation of the auxiliary parameters automatically reduces local smoothing power of the modified TV priors in Eqs. (16) and (17) around edges, effectively achieving edge-preserved smoothing. In addition, the small positive term ϵ works as a supplementary regularizer that prevents division by zero in Eqs. (16) and (17), thereby maintaining numerical stability.

3.3.2 Weights

According to the variational principle [4], the optimal approximate posterior of each variable can be obtained by taking the logarithmic expectation of the modified joint posterior $p'(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})$ with respect to the other approximate posteriors. Taking the logarithmic expectation of the right-hand side of Eq. (15) with respect to the other variables, the approximate posteriors of the smoothness parameters z, a, and the fidelity parameter b are obtained as gamma distributions as follows:

$$q(z) = \mathcal{G}(z|\alpha_z, \beta_z), \quad (20)$$

$$q(a) = \mathcal{G}(a|\alpha_a, \beta_a), \quad (21)$$

$$q(b) = \mathcal{G}(b|\alpha_b, \beta_b), \quad (22)$$

where

$$\alpha_z = \alpha_0 + 2n, \quad (23)$$

$$\alpha_a = \alpha_0 + m, \quad (24)$$

$$\alpha_b = \alpha_0 + \frac{1}{2}n, \quad (25)$$

$$\beta_z = \beta_0 + \frac{1}{2}\sum_{i=1}^{2n} u_i \left(\sum_{k=1}^{2} \text{E}\left[(\mathbf{F}_k\mathbf{w})_i^2\right] + \epsilon + u_i^{-2} \right), \quad (26)$$

$$\beta_a = \beta_0 + \frac{1}{2}\sum_{i=1}^{m} v_i \left(\sum_{k=1}^{2} \text{E}\left[(\mathbf{G}_k\mathbf{x})_i^2\right] + \epsilon + v_i^{-2} \right), \quad (27)$$

$$\beta_b = \beta_0 + \frac{1}{2}\sum_{i=1}^{n} \text{E}\left[(\mathbf{y} - \mathbf{H}'(\mathbf{w}' \circ \mathbf{x}'))_i^2\right]. \quad (28)$$

3.3.3 Optical Flow

Taking the logarithmic expectation of the right-hand side of Eq. (15) with respect to the variables other than $\mathbf{w}$, the approximate posterior of the optical flow $\mathbf{w}$ is obtained as a Gaussian distribution as follows:

$$q(\mathbf{w}) = \mathcal{N}(\mathbf{w}|\boldsymbol{\mu}_\mathbf{w}, \boldsymbol{\Sigma}_\mathbf{w}), \quad (29)$$

where

$$\boldsymbol{\mu}_\mathbf{w} = \text{E}[b]\boldsymbol{\Sigma}_\mathbf{w}\mathbf{J}^\mathsf{T}\left(\text{diag}(\text{E}[\mathbf{x}'])\mathbf{H}'^\mathsf{T}\mathbf{y} + (\mathbf{H}'^\mathsf{T}\mathbf{H}' \circ \text{E}[\mathbf{x}'\mathbf{x}'^\mathsf{T}])(\mathbf{J}\mathbf{w}^* - \mathbf{j})\right), \quad (30)$$

$$\boldsymbol{\Sigma}_\mathbf{w}^{-1} = \text{E}[b]\mathbf{J}^\mathsf{T}(\mathbf{H}'^\mathsf{T}\mathbf{H}' \circ \text{E}[\mathbf{x}'\mathbf{x}'^\mathsf{T}])\mathbf{J} + \text{E}[z]\sum_{k=1}^{2}\mathbf{F}_k^\mathsf{T}\text{diag}(\mathbf{u})\mathbf{F}_k. \quad (31)$$

Thus, the posterior of $\mathbf{w}$ depends not only on degraded images $\mathbf{y}$ but also on the posterior of clean image $\mathbf{x}$ and other parameters through the expectations in Eqs. (30) and (31). Hence, the dependency between the variables is naturally incorporated in VB inference. Moreover, the full posterior approximation of $\mathbf{w}$, i.e., not only the mean $\boldsymbol{\mu}_\mathbf{w}$ but also the covariance $\boldsymbol{\Sigma}_\mathbf{w}$ is available, which represents the uncertainty in estimation and contributes to the robustness of inference [4]. For example, $\boldsymbol{\Sigma}_\mathbf{w}$ is considered in evaluating the squared expectations of variations in Eq. (18), producing an additional positive term in the denominator as a regularizer for stable inference [18].

3.3.4 Clean Image

Taking the logarithmic expectation of the right-hand side of Eq. (15) with respect to the variables other than $\mathbf{x}$, the approximate posterior of the clean image $\mathbf{x}$ is obtained as a Gaussian distribution as follows:

$$q(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_\mathbf{x}, \boldsymbol{\Sigma}_\mathbf{x}), \quad (32)$$

where

$$\boldsymbol{\mu}_\mathbf{x} = \text{E}[b]\boldsymbol{\Sigma}_\mathbf{x}\mathbf{K}^\mathsf{T}\text{diag}(\text{E}[\mathbf{w}'])\mathbf{H}'^\mathsf{T}\mathbf{y}, \quad (33)$$

$$\boldsymbol{\Sigma}_\mathbf{x}^{-1} = \text{E}[b]\mathbf{K}^\mathsf{T}(\mathbf{H}'^\mathsf{T}\mathbf{H}' \circ \text{E}[\mathbf{w}'\mathbf{w}'^\mathsf{T}])\mathbf{K} + \text{E}[a]\sum_{k=1}^{2}\mathbf{G}_k^\mathsf{T}\text{diag}(\mathbf{v})\mathbf{G}_k, \quad (34)$$

Hence, this posterior also depends on the posterior of the optical flow $\mathbf{w}$, and has the variance $\boldsymbol{\Sigma}_\mathbf{x}$. It is trivial to maximize this approximate posterior, since the mode $\hat{\mathbf{x}}$ of Gaussian $q(\mathbf{x})$ coincides with the mean $\boldsymbol{\mu}_\mathbf{x}$ [4].

Table 1: Image quality of Lena (PSNR [dB]).

Sequence			Method		
Dis-place-ment	Blur	Noise	Degraded	Single	Multi (proposed)
Trans-lation	Gaus-sian	Weak	32.8	37.2	**39.8**
		Strong	25.0	29.8	**34.2**
	Uni-form	Weak	32.0	36.8	**39.7**
		Strong	24.9	29.5	**33.6**
Rota-tion	Gaus-sian	Weak	32.8	37.3	**39.7**
		Strong	25.0	29.8	**34.1**
	Uni-form	Weak	32.0	36.8	**39.7**
		Strong	24.9	30.0	**33.5**

(a) (b) (c)

Fig. 5: Images of Lena with translation, uniform blur, and strong noise: (a) degraded; (b) restored by single-frame denoising; and (c) restored by multiframe denoising (proposed) using 15 images.

3.4 Experiments

3.4.1 Comparison of Single-Frame and Multiframe Denoising

The effectiveness of the proposed VB multiframe denoising method was evaluated through experiments. To validate the benefit of using multiple images, the proposed VB method was compared with a single-frame method, which is a special case of the proposed method when the number of input images is one and no warping is involved. This single-frame method is in principle equivalent to the previous method proposed in [6], which also takes a VB approach and uses a TV-based image prior. In this experiment, the standard test image *Lena*, which consisted of 256×256 pixels, was used as a ground-truth clean image. Assuming displacement due to camera motion, this image was transformed by translation and rotation to obtain two sequences of warped images, each of which had 15 images; here, the amounts of displacement through the sequences were 1.5 pixels horizontally for translation and 1.5 degrees around the image center for rotation. Next, blur and noise were added to each sequence; here, each image was convolved with an uniform kernel and a Gaussian blur kernel of standard derivation 1, both of whose sizes were 3×3 pixels, while Gaussian noise was generated at 40 and 20 dB in *blurred-signal-to-noise ratio* (BSNR) [6], which represent weak and strong noise, respectively. Then, the single-frame and multiframe methods were applied to each of these degraded image sequences, selecting the image in the middle of the sequence as a reference image.

Image quality of each restored image was assessed using *peak-signal-to-noise ratio* (PSNR), which is a standard image quality metric for image restoration [24]. The resulting PSNR values are summarized in Table 1, while example images are shown in Fig. 5. As seen from the tables, for any original image and degradation type, the proposed multiframe method achieved higher quality of the restored image than the corresponding degraded image; thus, the denoising by the proposed method was successful regardless of input sequences. Furthermore, the proposed multiframe method achieved higher PSNR values than those of the single-frame version, thereby demonstrating the effectiveness of the proposed multiframe method compared with previous single-frame one.

3.4.2 Comparison of VB and Non-VB Approaches

The advantage of the proposed VB approach to multiframe denoising over non-VB approaches was examined. For a fair comparison between approaches, the proposed VB method was compared with a non-VB method that uses the same model, i.e., the multiframe extension of classical TV-based denoising with manual parameter tuning [25], [26], which directly maximizes the original lower bound LB with respect to $\hat{\mathbf{x}}$ by nonlinear optimization. As with the majority of non-VB methods, this method requires manual parameter tuning, since it is unable to estimate parameters other than clean image $\mathbf{x}$; thus, the ground-truth flow was used for warping parameter $\mathbf{w}$, and image smoothness parameter a was varied between 1 and 128, while ignoring fidelity b and flow smoothness z (which have no effect in this method). Moreover, a comparison with a state-of-the-art non-VB multiframe method with partially automatic parameter tuning [16] was made, which performs optical-flow-based registration and iterative MAP estimation. Since the main objective of this method was to reconstruct a high-resolution image from multiple low-resolution images, we adapted this method to our denoising problem by omitting downsampling and blur kernel estimation, and employed the same iteration scheme as the proposed method. While this non-VB method can automatically set parameters for registration and noise, it still requires several free parameters to be given, for which the values suggested in its original paper were used.

Table 2 shows the PSNR values resulting from the non-VB manual method with different image smoothness parameter values (where only the results for $a = 1, 16$ are shown to summarize the overall trend), the non-VB automatic method, and the proposed VB automatic method. Closeup examples of the restored images are also shown in Fig. 6, which are obtained from the sequence with translation, uniform blur, and strong noise. First, the optimal values of the smoothness parameter for the manual method that yielded the best results varied a lot between noise levels. This result confirms the fact that traditional methods of image restoration require careful manual parameter tuning for each input. Note that, since this method used the true value of the warping parameter, no better results could have been expected from this method, even if state-of-the-art registration methods had been used. Compared with this manual method, the proposed automatic method achieved higher PSNR values regardless of degradation without knowing the ground-truth warping parameter. These results shows that the automatic parameter tuning of the proposed method was quite successful, which is a huge advantage of our VB approach over traditional approaches. Moreover, the proposed VB automatic method always outperformed the adaptation of the state-of-the-art non-VB automatic method

Table 2: Image quality by multiframe methods (PSNR [dB]).

Sequence			Method			
Dis-place-ment	Blur	Noise	Non-VB			VB
			Manual with a		Auto	
			1	16		
Trans-lation	Uni-form	Weak	38.5	34.2	34.6	**39.7**
		Strong	26.0	33.4	30.1	**33.6**
	Gaus-sian	Weak	38.9	34.7	34.8	**39.8**
		Strong	26.5	33.8	30.8	**34.2**
Rota-tion	Uni-form	Weak	37.5	34.2	34.3	**39.7**
		Strong	25.7	33.4	30.1	**33.5**
	Gaus-sian	Weak	38.9	34.7	34.3	**39.7**
		Strong	26.6	33.8	30.8	**34.1**

(a)   (b)   (c)   (d)
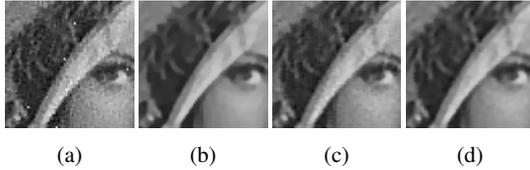
Fig. 6: Closeups of images restored by multiframe methods: the non-VB manual method with (a) $\alpha = 1$ (worst quality) and (b) $\alpha = 16$ (best quality); (c) the non-VB automatic method; and the proposed VB automatic method.

for all the sequences, further demonstrating the advantage of the VB approach.

3.5 Conclusion

In this section, a method of multiframe denoising via a VB approach has been presented. Extending previous work on single-frame denoising via VB approaches [6], [17], [18], image warping has been introduced to a Bayesian model for denoising, effectively integrating registration into restoration. Then, VB techniques such as mean-field and local approximation have been used to overcome difficulties due to the complexity of the model in deriving an algorithm of Bayesian inference, where full posteriors of multiple latent variables are iteratively updated. The proposed VB multiframe denoising method can achieve higher image quality than single-frame one, while tuning multiple parameters automatically and stably. Through experiments, the effectiveness of the proposed VB multiframe method was demonstrated in comparison with its single-frame version and non-VB approaches.

4. Variational Bayesian Deblurring

4.1 Introduction

The basic assumption in deblurring is that blurring on an image can be modeled as convolution with a blur kernel. Deblurring, which is often called deconvolution, is essentially the inverse of this process, and effectively restores the sharp image from the blurry one. In practice, the kernel is often unknown [3], and needs to be estimated to perform deconvolution. This motivates the problem called blind deconvolution, i.e., deblurring in the presence of unknown blur kernels.

Traditional shift-invariant deblurring assumes uniform blur that does not vary spatially across an image. In reality, however, blur is often nonuniform [3]; for example, independently moving objects are differently affected by motion blur, and close and dis-

tant objects are subject to different amounts of defocus blur. Blur is often more complex since both motion and defocus blur can simultaneously occur. In such cases, shift-invariant methods fail in blur kernel estimation and also in blur removal [3]. Although specialized methods for nonuniform blur have been recently developed, they can only handle certain types of blur, e.g., motion, defocus, or locally uniform blur in each image segment [8]. Consequently, the ability of deblurring to restore general images with nonuniform blur has been limited.

In this section, shift-variant deblurring that can handle nonuniform blur regardless of type is discussed. The idea is to model spatially varying blur with a field of kernels that assigns a local blur kernel to each image pixel. By allowing different kernels between pixels, a wide range of nonuniform blur can be flexibly represented without making assumptions about its type. To alleviate the ill-posedness of the shift-variant problem, smoothing of the field is also introduced, which sufficiently regularizes kernel estimation without losing flexibility. Under this model, both a sharp clean image and the field of kernels are estimated from a blurry degraded image using techniques of variational Bayes. As demonstrated through experiments, the proposed field-based model is flexible enough to deal with complex blur beyond the ability of previous shift-invariant and shift-variant approaches.

4.2 Model

Given a blurry degraded image of n pixels, denoted by a vector $\mathbf{y} = [y_1 \ldots y_n]^\mathsf{T} \in \mathbb{R}^n$, the goal of deblurring is to recover a sharp clean image, i.e., a blur-free version of $\mathbf{y}$, denoted by $\mathbf{x} = [x_1 \ldots x_n]^\mathsf{T} \in \mathbb{R}^n$. To achieve this, we assume that each degraded pixel y_i is the convolution of the clean image $\mathbf{x}$ and a local blur kernel with m coefficients $\mathbf{w}_i = [w_{i1} \ldots w_{im}]^\mathsf{T} \in \mathbb{R}^m$:

$$y_i \simeq \sum_{j=1}^m w_{ij} x_{i \oplus j}, \qquad (35)$$

where $i \oplus j$ denotes the jth coefficient in the spatial support of the kernel at the ith pixel, and the approximate equality implies the presence of noise. In practice, it is assumed that the support is square and centered at the ith pixel. Then, the n local kernels from all the pixels are gathered into a single vector $\mathbf{w} = [\mathbf{w}_1^T \ldots \mathbf{w}_m^T]^\mathsf{T} \in \mathbb{R}^{mn}$, which is referred to as a field of kernels because it has one kernel at each pixel in space.

In the deblurring problem, both the field of kernels and the clean image are unknown, and thus need to be estimated in order to perform deconvolution. To describe the relationship of these parameters, we construct a Bayesian model in the following. The resulting Bayesian model is shown graphically in Fig. 7.

4.2.1 Degraded Image

Let $\mathbf{H_w} \in \mathbb{R}^{n \times n}$ be the shift-variant convolution matrix with respect to $\mathbf{w}$ such that $(\mathbf{H_w}\mathbf{x})_i$ equals the right-hand side of Eq. (35). Then, assuming additive noise, we can rewrite Eq. (35) as follows:

$$\mathbf{y} = \mathbf{H_w}\mathbf{x} + \mathbf{n}, \qquad (36)$$

where $\mathbf{n} \in \mathbb{R}^n$ is a zero-mean Gaussian noise vector, which is a standard assumption in deblurring [8]. We rephrase this degradation model in Bayesian terms by defining an elementwise Gaussian distribution on $\mathbf{y}$ as follows:
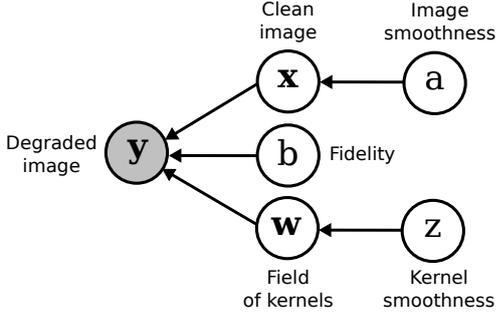
Fig. 7: Bayesian model for deblurring. Gray and white circles represent observed and latent random variables, respectively.

$$p(\mathbf{y}|\mathbf{x}, \mathbf{w}, b) \propto \prod_{i=1}^{n} \sqrt{b} \exp\left(-\frac{b}{2}(\mathbf{y} - \mathbf{H_w}\mathbf{x})_i^2\right), \quad (37)$$

where $b \in \mathbb{R}$ is a fidelity parameter.

4.2.2 Clean Image

Exploiting natural image statistics, we assume that the clean image is smooth except for edges, and minimize its TV. This is achieved by defining an elementwise zero-mean Laplacian distribution on the gradient magnitude of $\mathbf{x}$ as follows:

$$p(\mathbf{x}|a) \propto \prod_{i=1}^{n} a \exp\left(-a\sqrt{\sum_{k=1}^{2}(\mathbf{G}_k\mathbf{x})_i^2}\right), \quad (38)$$

where $a \in \mathbb{R}$ is an image smoothness parameter, and $\mathbf{G}_1, \mathbf{G}_2 \in \mathbb{R}^{m \times m}$ are horizontal and vertical differentiation matrices, respectively.

4.2.3 Field of Kernels

In shift-invariant deblurring, it is known that kernels can be estimated even without priors [8]. In our shift-variant case, however, the field of kernels has multiple coefficients at each pixels, which makes the problem highly ill-posed. To overcome this difficulty, we also impose a prior on the field of kernels $\mathbf{w}$. While we could use shift-invariant kernel priors [8], [27] for each local kernel, such priors limit the range of tractable blur types, as in restrictive parameterization. Instead, we assume that the field of kernels is smooth, i.e., local kernels at adjacent pixels are similar. This assumption effectively regularizes the ill-posed shift-variant deblurring problem, but still allows for kernel variation between pixels. This is realized by defining an elementwise zero-mean Gaussian distribution on the gradient magnitudes of the field as follows:

$$p(\mathbf{w}|z) \propto \prod_{i=1}^{mn} \sqrt{z} \exp\left(-\frac{z}{2}\sum_{k=1}^{2}(\mathbf{F}_k\mathbf{w})_i^2\right), \quad (39)$$

where $z \in \mathbb{R}$ is a kernel smoothness parameter for $\mathbf{w}$, and $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{R}^{mn \times mn}$ are horizontal and vertical differentiation matrices, respectively. While this smoothing prior is not edge-preserving, it is sufficient to enable shift-variant deblurring with the field of kernels.

To reflect the physical properties of real kernels [8], each local kernel $\mathbf{w}_i$ is also assumed to be normalized, i.e.,

$$0 \leq w_{ij} \leq 1, \quad (40)$$

$$\sum_{j=1}^{m} w_{ij} = 1. \quad (41)$$

4.2.4 Weights

For the weight-like parameters in the model, i.e., b, a, and z, we simply assume uniform distributions as noninformative priors, i.e., $p(b)$, $p(a)$, and $p(z)$ are constant.

4.3 Inference

The objective of Bayesian deblurring is to find the most probable clean image given a degraded image. Thus, under the model defined in Section 4.2, we maximize the posterior probability of $\mathbf{x}$ given $\mathbf{y}$ to obtain a restored image $\hat{\mathbf{x}}$ as follows:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}). \quad (42)$$

The posterior distribution of $\mathbf{x}$ is obtained by marginalizing out latent variables other than $\mathbf{x}$ from the joint posterior of all variables:

$$p(\mathbf{x}|\mathbf{y}) = \int p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})d\mathbf{w}dbdadz, \quad (43)$$

where

$$p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{w}, b)p(\mathbf{x}|a)p(\mathbf{w}|z)p(b)p(a)p(z). \quad (44)$$

Since exact marginalization is difficult due to the mutual dependency between variables, we invoke the VB technique of mean-field approximation, where the exact joint posterior is approximated by the product of independent posteriors of individual variables as follows:

$$p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y}) \simeq q(\mathbf{x}, \mathbf{w}, b, a, z) = q(\mathbf{x})q(\mathbf{w})q(b)q(a)q(z). \quad (45)$$

Here, $q(\mathbf{x})$ approximates $p(\mathbf{x}|\mathbf{y})$, which is required to evaluate Eq. (42). Following previous work on VB deblurring [8], we also assume that $q(\mathbf{x})$ is Gaussian and $q(\mathbf{w}), q(b), q(a)$, and $q(b)$ are degenerate. This assumption makes all approximate posteriors well-parameterized, thereby simplifying optimization [28]. Let $\boldsymbol{\mu}_\mathbf{x} \in \mathbb{R}^n$ and $\boldsymbol{\Sigma}_\mathbf{x} \in \mathbb{R}^{n \times n}$ be the mean and covariance of $q(\mathbf{x})$, respectively, and let $\hat{b}, \hat{a}, \hat{z}$, and $\hat{\mathbf{w}}$ be the modes of $q(b), q(a), q(z)$, and $q(\mathbf{w})$, respectively. Then, optimal approximate posteriors in terms of Kullback-Leibler divergence can be obtained by maximizing the following log-evidence lower bound of the model:

$$\text{LB}[q(\mathbf{x}, \mathbf{w}, b, a, z)]$$
$$= \int q(\mathbf{x}, \mathbf{w}, b, a, z) \ln \frac{p(\mathbf{x}, \mathbf{w}, b, a, z|\mathbf{y})}{q(\mathbf{x}, \mathbf{w}, b, a, z)} d\mathbf{x}d\mathbf{w}dbdadz + \text{const}. \quad (46)$$

As the non-Gaussian image prior defined in Eq. (38) prevents direct optimization of LB, we invoke another VB technique, i.e., local approximation, which is also common in VB deblurring [8]. Then, the non-Gaussian $p(\mathbf{x}|a)$ in Eq. (44) is approximated with Gaussian-like distribution $p'(\mathbf{x}|a, \mathbf{v})$ with auxiliary parameter $\mathbf{v} \in \mathbb{R}^n$, which is derived from lower-bounding of $p(\mathbf{x}|a)$ [29] and defined as follows:

$$p'(\mathbf{x}|a) \propto \prod_{i=1}^{n} \sqrt{a} \exp\left(-\frac{av_i}{2}\left(\sum_{k=1}^{2}(\mathbf{G}_k\mathbf{x})_i^2 + v_i^{-2}\right)\right). \quad (47)$$

Intuitively, v_i is a local smoothing weight at the ith pixel, which is

---

Algorithm 2 Algorithm of VB inference for deblurring.

1: Input $\mathbf{y}$.
2: Initialize $\boldsymbol{\mu}_\mathbf{x}$, $\boldsymbol{\Sigma}_\mathbf{x}$, and $\hat{\mathbf{w}}$.
3: **repeat**
4:    Update $\mathbf{v}$ with Eq. (49).
5:    Update $\hat{z}$, $\hat{a}$, and $\hat{b}$ with Eqs. (50) to (52).
6:    Update $\boldsymbol{\mu}_\mathbf{x}$ and $\boldsymbol{\Sigma}_\mathbf{x}$ with Eqs. (53) and (54).
7:    Update $\hat{\mathbf{w}}$ with Eq. (55).
8: **until** convergence.
9: Output $\hat{\mathbf{x}} = \boldsymbol{\mu}_\mathbf{x}$.

---

adapted to the clean image automatically through VB inference, thereby enabling edge-preserving smoothing.

Combining Eqs. (44) and (46) and replacing $p(\mathbf{x}|a)$ with $p'(\mathbf{x}|a)$, a modified lower bound is obtained, which depends on both the approximate posteriors and the auxiliary parameter $\mathbf{v}$. This can be evaluated using the parameters of approximate posteriors and the definition of the distributions, i.e., Eqs. (37), (39) and (47) as follows:

$$
\begin{aligned}
&\mathrm{LB}'(\boldsymbol{\mu}_\mathbf{x}, \boldsymbol{\Sigma}_\mathbf{x}, \hat{\mathbf{w}}, \hat{b}, \hat{a}, \hat{z}, \mathbf{v}) \\
=\ & \frac{n}{2} \ln \hat{b} - \frac{\hat{b}}{2} \left(\|\mathbf{y} - \mathbf{H}_{\hat{\mathbf{w}}} \boldsymbol{\mu}_\mathbf{x}\|_2^2 + \mathrm{tr}(\mathbf{H}_{\hat{\mathbf{w}}} \boldsymbol{\Sigma}_\mathbf{x} \mathbf{H}_{\hat{\mathbf{w}}}^\mathsf{T}) \right) \\
& + \frac{n}{2} \ln \hat{a} - \frac{\hat{a}}{2} \left(\sum_{k=1}^{2} \left(\|\mathbf{V}^{\frac{1}{2}} \mathbf{G}_k \boldsymbol{\mu}_\mathbf{x}\|_2^2 + \mathrm{tr}(\mathbf{V} \mathbf{G}_k \boldsymbol{\Sigma}_\mathbf{x} \mathbf{G}_k^\mathsf{T}) \right) + \mathrm{tr}(\mathbf{V}^{-1}) \right) \\
& + \frac{mn}{2} \ln \hat{z} - \frac{\hat{z}}{2} \sum_{k=1}^{2} \|\mathbf{F}_k \hat{\mathbf{w}}\|_2^2 \\
& + \frac{1}{2} \ln |\boldsymbol{\Sigma}_\mathbf{x}| + \text{const.},
\end{aligned}
$$
(48)

where $\mathbf{H}_{\hat{\mathbf{w}}}$ is the version of $\mathbf{H}_\mathbf{w}$ constructed with $\hat{\mathbf{w}}$, and $\mathbf{V} = \mathrm{diag}(\mathbf{v})$. Then, this modified lower bound can be maximized with respect to each parameter. Since each parameter depends on others, they are iteratively updated until convergence, as described in the following. The resulting algorithm is summarized in Algorithm 2.

4.3.1 Auxiliary Parameter

Taking the derivative of Eq. (48) with respect to v_i and setting it to zero, each element v_i of the auxiliary parameter $\mathbf{v}$ is obtained as follows:

$$
v_i = \frac{1}{\sqrt{\sum_{k=1}^{2} \left((\mathbf{G}_k \boldsymbol{\mu}_\mathbf{x})_i^2 + (\mathbf{G}_k \boldsymbol{\Sigma}_\mathbf{x} \mathbf{G}_k^\mathsf{T})_{ii} \right)}}.
$$
(49)

Thus, v_i becomes small around an edge, thereby weakening local smoothing to preserve the edge.

4.3.2 Weights

Taking the derivative of Eq. (48) with respect to each of $\hat{z}$, $\hat{a}$, and $\hat{b}$ and setting it to zero, the parameters of $q(z)$, $q(a)$, and $q(b)$ are obtained as follows:

$$
\hat{z} = \frac{mn}{\sum_{i=1}^{mn} \sum_{k=1}^{2} (\mathbf{F}_k \hat{\mathbf{w}})_i^2},
$$
(50)

$$
\hat{a} = \frac{n}{\sum_{i=1}^{n} \left(v_i \sum_{k=1}^{2} \left((\mathbf{G}_k \boldsymbol{\mu}_\mathbf{x})_i^2 + (\mathbf{G}_k \boldsymbol{\Sigma}_\mathbf{x} \mathbf{G}_k^\mathsf{T})_{ii} \right) + v_i^{-1} \right)},
$$
(51)

$$
\hat{b} = \frac{n}{\sum_{i=1}^{n} \left((\mathbf{y} - \mathbf{H}_{\hat{\mathbf{w}}} \boldsymbol{\mu}_\mathbf{x})_i^2 + (\mathbf{H}_{\hat{\mathbf{w}}} \boldsymbol{\Sigma}_\mathbf{x} \mathbf{H}_{\hat{\mathbf{w}}}^\mathsf{T})_{ii} \right)}.
$$
(52)

4.3.3 Clean Image

Taking the derivative of Eq. (48) with respect to $\boldsymbol{\mu}_\mathbf{x}$, $\boldsymbol{\Sigma}_\mathbf{x}$ and setting it to zero, the parameters of $q(\mathbf{x})$ are obtained as follows:

$$
\boldsymbol{\mu}_\mathbf{x} = \left(\mathbf{H}_{\hat{\mathbf{w}}}^\mathsf{T} \mathbf{H}_{\hat{\mathbf{w}}} + \frac{\hat{a}}{\hat{b}} \sum_{k=1}^{2} \mathbf{G}_k^\mathsf{T} \mathbf{V} \mathbf{G}_k \right)^{-1} \mathbf{H}_{\hat{\mathbf{w}}}^\mathsf{T} \mathbf{y},
$$
(53)

$$
\boldsymbol{\Sigma}_\mathbf{x} = \left(\hat{b} \mathbf{H}_{\hat{\mathbf{w}}}^\mathsf{T} \mathbf{H}_{\hat{\mathbf{w}}} + \hat{a} \sum_{k=1}^{2} \mathbf{G}_k^\mathsf{T} \mathbf{V} \mathbf{G}_k \right)^{-1}.
$$
(54)

Here, the covariance $\boldsymbol{\Sigma}_\mathbf{x}$ represents uncertainty in estimates of $\mathbf{x}$, working as regularizers in the denominators of Eqs. (49), (51) and (52). In deblurring, this property of VB inference helps avoid trivial solutions and improve performance [3], [29], [30]. By replacing the exact posterior $p(\mathbf{x}|\mathbf{y})$ with the obtained approximate posterior $q(\mathbf{x})$ in Eq. (42), a restored image is obtained as the mode $\hat{\mathbf{x}}$ of $q(\mathbf{x})$, which coincides with $\boldsymbol{\mu}_\mathbf{x}$ since $q(\mathbf{x})$ is Gaussian.

4.3.4 Field of Kernels

Taking the derivative of Eq. (48) with respect to $\hat{\mathbf{w}}$ and setting it to zero, the parameter of $q(\mathbf{w})$ is obtained as follows:

$$
\hat{\mathbf{w}} = \left(\mathbf{H}_{\boldsymbol{\mu}_\mathbf{x}}^\mathsf{T} \mathbf{H}_{\boldsymbol{\mu}_\mathbf{x}} + \mathbf{H}_{\boldsymbol{\Sigma}_\mathbf{x}}^2 + \frac{\hat{z}}{\hat{b}} \sum_{k=1}^{2} \mathbf{F}_k^\mathsf{T} \mathbf{F}_k \right)^{-1} \mathbf{H}_{\boldsymbol{\mu}_\mathbf{x}}^\mathsf{T} \mathbf{y},
$$
(55)

and $\mathbf{H}_{\boldsymbol{\mu}_\mathbf{x}}$ and $\mathbf{H}_{\boldsymbol{\Sigma}_\mathbf{x}}^2$ are the matrices such that

$$
\mathbf{H}_{\boldsymbol{\mu}_\mathbf{x}} \hat{\mathbf{w}} = \mathbf{H}_{\hat{\mathbf{w}}} \boldsymbol{\mu}_\mathbf{x},
$$
(56)

$$
\hat{\mathbf{w}}^\mathsf{T} \mathbf{H}_{\boldsymbol{\Sigma}_\mathbf{x}}^2 \hat{\mathbf{w}} = \mathrm{tr}(\mathbf{H}_{\hat{\mathbf{w}}} \boldsymbol{\Sigma}_\mathbf{x} \mathbf{H}_{\hat{\mathbf{w}}}^\mathsf{T}), .
$$
(57)

4.4 Experiments

4.4.1 Comparison of Shift-Invariant and Shift-Variant Deblurring

We conducted experiments to evaluate the effectiveness of the proposed method. For a fair comparison, the proposed shift-variant method was modified into a shift-invariant version with nearly the same model, assuming a single local kernel common to all pixels in the convolution model Eq. (35) and disabling kernel smoothing. Ignoring minor differences in optimization, this shift-invariant version is equivalent to state-of-the-art VB deblurring methods with TV image priors [7], [29], [30]. In addition, the proposed field-based method was compared with a segmentation-based variant that assumes the same kernel in each segment in Eq. (35). To enable ground-truth-based evaluation, a degraded image was prepared by blurring the standard test image *Lena* as a clean image via shift-variant convolution with a field of kernels consisting of horizontally-varying line kernels in the foreground and a box kernel in the background, which simulates rotational motion blur and uniform defocus blur, respectively. The sizes of the image and local kernel were $n = 256 \times 256$ pixels and $m = 5 \times 5$ coefficients per pixel, respectively. Several samples of the local kernels are shown in Fig. 8(a). For the segmentation-based method, the ground-truth segments (the foreground and background) were given as input.

The resulting images are shown in Fig. 9 with their PSNR values in Table 3. Here, the shift-invariant method could not fully remove blur and even degraded image quality in terms of PSNR, attempting deconvolution based on shift-invariant kernel estimation
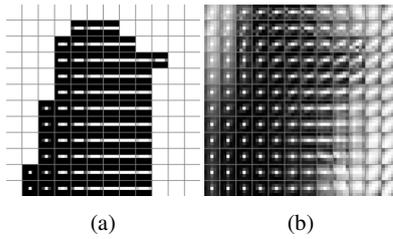
Fig. 8: Kernels for Lena: (a) ground truth and (b) estimated by the proposed method. From the 256×256 pixels of each field, 12×12 pixels were sampled at an equal interval. The value range of each local kernel was maximized for visualization. Note that the white blocks in (a) indicate uniform kernels.
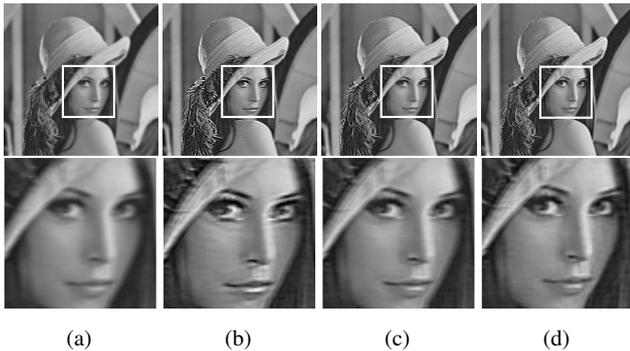
Fig. 9: Images of Lena (top) and closeups (bottom): (a) degraded; (b) restored by shift-invariant deblurring; (c) restored by segmentation-based shift-variant deblurring; and (d) restored by field-based shift-variant deblurring (proposed).

Table 3: Image quality of Lena (PSNR [dB]).

Degraded	Shift-invariant	Segmentation-based shift-variant	Field-based shift-variant (proposed)
28.3	23.2	28.6	**30.2**

against nonuniform blur. Meanwhile, the segmentation-based shift-variant deblurring could make only a slight PSNR improvement despite the use of the perfect segmentation result, since the smoothly varying blur in the foreground was difficult to approximate by segmentation. Note that such composite blur is out of the domain of most parametric deblurring methods that globally assume a single type of blur, e.g., motion or defocus. By contrast, the proposed field-based shift-variant method successfully recovered a visually sharper image (e.g., around the edge of the face) with a higher PSNR value, dealing with both the smooth variation and the complexity of the blur. For reference, the field of kernels estimated by the proposed method is shown in Fig. 8(b). Here, the outline of the field, i.e., the difference between the foreground and background, and also the smooth variation in the foreground along the horizontal axis are successfully captured. Because of the field smoothing via the non-edge-preserving Gaussian prior, discontinuities around object boundaries were smoothed out. Still, such a kernel prior was sufficient for achieving PSNR improvement in the image domain, which is the ultimate objective of deblurring.

4.4.2 Comparison of Shift-Variant Methods

Using real images with complex blur, the proposed method
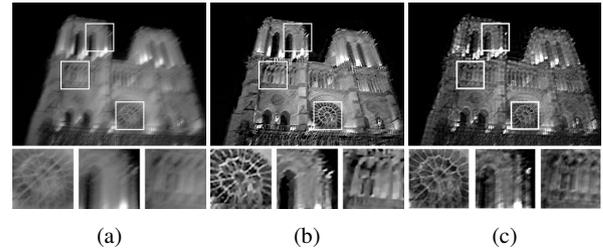
Fig. 10: Images for comparison of shift-variant methods: (a) degraded; (b) restored by the previous specialized method; and (c) restored by the proposed general method. The degraded image input to the proposed method is from [32], while the restored image output from the previous method is from [31]. These images were gamma-corrected for visual comparison.

for general nonuniform blur was compared with a previous shift-variant deblurring method specialized to camera shake, i.e., the method of Hirsch et al. [31], which is based on segmentwise blur parameterization with motion paths. As no implementations of this method were available, we used the same image as in their experiments for comparison, i.e., an image with camera shake used in Whyte et al. [32], assuming local kernels of size 11×11 pixels. Since this real image do not have the corresponding ground truth image and kernel, only visual comparison is possible here.

The results are shown in Fig. 10. Overall, the deblurring performances of both methods were visually comparable. However, the previous specialized method exhibited oversmoothing in regions with textures (in the middle and right closeups of Fig. 10(b)), which could happen when the assumptions under their parameterization did not hold exactly. Meanwhile, the proposed general method produced some artifacts in highly blurry regions (in the middle closeup of Fig. 10(c)) possibly due to inaccuracy in kernel estimation, which could be seen as the price for the flexibility offered by our nonparametric approach; still, such artifacts might be suppressed by using more effective image priors [7]. Note that the proposed method has a clear advantage over this previous method, i.e., while the previous method is useful only for images with camera motion blur, the proposed method has the potential to handle defocus blur, object motion blur, and even their mixture, without the need for blur type identification for each image.

4.5 Conclusion

In this section, a method of shift-variant deblurring with a field-based model has been presented. By modeling a spatially varying blur kernel as a smooth field of local kernels, the proposed VB-based method can flexibly handle different kernels between pixels. Experimental results confirmed that the proposed method can successfully deblur images affected by nonuniform blur that is intractable for previous approaches such as shift-invariant and segmentation-based shift-variant deblurring.

5. Variational Bayesian Devignetting

5.1 Introduction

Vignetting arises from various optical components of real imaging systems; thus, it is difficult to eliminate all possible

causes of vignetting before actually capturing images. Hence, it is useful to reduce undesirable vignetting effect in a postprocess. While calibration is a basic approach to vignetting correction [33], it requires additional work during image capturing, such as taking reference images [34], and cannot be applied to previously captured images. In this work, the problem of single-frame devignetting using no other images is addressed, thereby affording broader applicability than calibration techniques.

The vignetting effect in an image is commonly modeled with a vignetting function, which describes the radial decrease of brightness from the optical center of an imaging system. In devignetting, it is essential to estimate this function accurately. Previous work on single-frame devignetting [34], [35], [36], [37], [38] has basically assumed a rather ideal imaging condition, i.e., degradation other than vignetting is absent or negligible. In this case, a correct vignetting function can often be determined from a vignetted image, and then a desired vignetting-free image can be obtained by inverting the vignetting process with that function. In reality, however, images captured by cameras are often affected by noise, which is inherent in imaging systems. When noise is dominant in a vignetted image, previous devignetting methods that perform näive inversion of vignetting can remove only vignetting but not noise, thereby being limited in the quality of the restored image. Furthermore, estimation of a vignetting function from a noisy image can become unstable due to noise, which leads to poor devignetting performance.

In this section, stable single-frame devignetting in the presence of noise is discussed. Specifically, we introduce a general image prior that exploits natural image statistics, whose effectiveness for noise removal has been well-confirmed in image restoration. Then, in Bayesian inference enabled by VB, we jointly estimate a vignetting function and a clean image without vignetting nor noise from a degraded image, considering their relationship. This prior-driven VB devignetting can achieve high image quality even in the presence of noise because it seeks a clean image free from both vignetting and noise, which further benefits stable estimation of the vignetting function through the joint inference. In experiments, we confirmed the effectiveness of the proposed VB approach to devignetting, especially when noise is strong, comparing the proposed method with a state-of-the-art method both qualitatively and quantitatively.

5.2 Model

In devignetting, the goal is to estimate a clean image with no vignetting from a degraded image affected by vignetting. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be vectors of the clean and degraded images, respectively, where n denotes the total number of pixels per image. Meanwhile, the vignetting effect is commonly described using a one-dimensional vignetting function [34], [36], [37], since natural vignetting is mostly radial, i.e., constant at each radius around an optical center, which is assumed to be known here, as in the majority of previous studies [33], [34], [35]. After discretization, this function is represented as another vector $\mathbf{w} \in \mathbb{R}^m$, where m is the number of discrete radii and the first element at the radial origin corresponds to the optical center.

In the following, a Bayesian model is constructed to describe
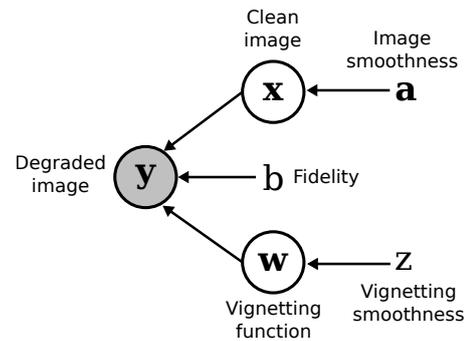
Fig. 11: Bayesian model for devignetting. Gray, white, and no circles represent observed random, latent random, and deterministic variables, respectively.
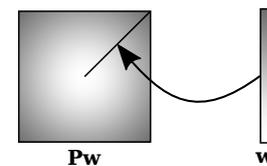
Fig. 12: Prolongation of one-dimensional vignetting function $\mathbf{w}$ into two-dimensional version $\mathbf{Pw}$ via operator $\mathbf{P}$. Note that algorithmically $\mathbf{Pw}$ is treated as a flattened one-dimensional vector.

the relationship between parameters, i.e., the degraded and clean images, the vignetting function, and other weight parameters. The resulting Bayesian model is shown graphically in Fig. 11.

5.2.1 Degraded Image

Unlike previous vignetting studies, we assume that each pixel of the degraded image is affected by additive zero-mean Gaussian noise. Note that this is a standard assumption in image restoration [29]. By denoting the vector of pixelwise noise intensity values by $\mathbf{n} \in \mathbb{R}^n$, the following relationship between parameters is obtained:

$$\mathbf{y} = \mathbf{H_w}\mathbf{x} + \mathbf{n} = \mathbf{Pw} \circ \mathbf{x} + \mathbf{n}, \qquad (58)$$

where $\mathbf{H_w} = \mathrm{diag}(\mathbf{Pw})$, and $\mathbf{P} \in \mathbb{R}^{n \times m}$ is a matrix that prolongs the one-dimensional vignetting function into a two-dimensional function, as shown in Fig. 12. Using this degradation model, we define the distribution of the degraded image $\mathbf{y}$ as the following elementwise Gaussian distribution:

$$p(\mathbf{y}|\mathbf{x}, \mathbf{w}) \propto \prod_{i=1}^{n} \sqrt{b} \exp\left(-\frac{b}{2}(\mathbf{y} - \mathbf{Pw} \circ \mathbf{x})_i^2\right), \qquad (59)$$

where $b \in \mathbb{R}$ is a fidelity parameter.

5.2.2 Clean Image

In the presence of noise, estimation of a clean image from a degraded image under the degradation model in Eq. (58) is an ill-posed problem, i.e., $\mathbf{x}$ cannot be uniquely determined from $\mathbf{y}$ alone (even if $\mathbf{w}$ is known), since the complete information of $\mathbf{x}$ is already lost in $\mathbf{y}$ due to noise. Thus, unlike previous devignetting methods that ignore noise, additional information is necessary to recover a noise-free image. To this end, we exploit the statistical regularity of natural images, i.e., an image in the real world is generally smooth except for edges [23]. This means that local variations in the image, i.e., the magnitudes of image gradients,

should be small at most pixels, but can be large at a few discontinuities. To minimize local image variations, we define the following elementwise zero-mean Gaussian distribution on the gradient magnitudes of the clean image $\mathbf{x}$:

$$p(\mathbf{x}) \propto \prod_{i=1}^{n} \sqrt{a_i} \exp\left(-\frac{a_i}{2} \sum_{k=1}^{2} (\mathbf{G}_k \mathbf{x})_i^2\right), \qquad (60)$$

where $\mathbf{G}_1, \mathbf{G}_2 \in \mathbb{R}^{n\times n}$ are horizontal and vertical differentiation matrices, respectively, and $\mathbf{a} \in \mathbb{R}^n$ is a local image smoothness parameter; here, each element a_i measures the local smoothness of $\mathbf{x}$ at the ith pixel. Thus, edge-preserving smoothing can be achieved by adapting $\mathbf{a}$ to the structure of $\mathbf{x}$ through inference [30].

5.2.3 Vignetting Function

To regularize the estimation of the vignetting function without losing flexibility, we perform smoothing along radii by imposing a Gaussian distribution on the magnitudes of the radial gradients of $\mathbf{w}$ as follows:

$$p(\mathbf{w}) \propto \prod_{i=1}^{m} \sqrt{z} \exp\left(-\frac{z}{2} (\mathbf{F}\mathbf{w})_i^2\right), \qquad (61)$$

where $\mathbf{F} \in \mathbb{R}^{m\times m}$ is a radial differentiation matrix, and $z \in \mathbb{R}$ is a global vignetting smoothness parameter.

Moreover, $\mathbf{w}$ is constrained to obtain a valid vignetting function that reflect the natural vignetting characteristics, i.e., it takes values between zero and one, decreases with respect to radii, and is one at the origin corresponding to the optical center:

$$0 \leq \mathbf{w} \leq 1, \qquad (62)$$

$$\mathbf{F}\mathbf{w} < 0, \qquad (63)$$

$$w_1 = 0, \qquad (64)$$

where vector inequalities are elementwise.

5.3 Inference

Under the model defined in Section 5.2, we estimate the most probable clean image $\hat{\mathbf{x}}$ given the degraded image $\mathbf{y}$. This can be achieved by maximizing the marginal posterior probability of $\mathbf{x}$ given $\mathbf{y}$ while marginalizing out $\mathbf{w}$ as follows:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) = \arg\max_{\mathbf{x}} \int p(\mathbf{x}, \mathbf{w}|\mathbf{y}) d\mathbf{w}. \qquad (65)$$

Here, the joint posterior distribution of the latent variables can be obtained as the product of the distributions in the model:

$$p(\mathbf{x}, \mathbf{w}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{w}) p(\mathbf{x}) p(\mathbf{w}), \qquad (66)$$

where Bayes' theorem has been used [39].

Since exact marginalization of $\mathbf{w}$ is difficult under the constraints in Eqs. (62) to (64), we use mean-field variational Bayes. Specifically, we factorize the exact joint posterior distribution into approximate posterior distributions of the individual variables as follows:

$$p(\mathbf{x}, \mathbf{w}|\mathbf{y}) \simeq q(\mathbf{x}, \mathbf{w}) = q(\mathbf{x})q(\mathbf{w}). \qquad (67)$$

To obtain a tractable distribution, we also restrict each approximate posterior as an elementwise Gaussian, whose covariance

---

Algorithm 3 Algorithm of VB inference for devingetting.

1: Input $\mathbf{y}$.
2: Initialize $\boldsymbol{\mu}_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \boldsymbol{\mu}_{\mathbf{w}}, \sigma_{\mathbf{w}}^2, b, \mathbf{a}, z$.
3: **repeat**
4:     Update $\boldsymbol{\mu}_{\mathbf{x}}$ and $\sigma_{\mathbf{x}}^2$ with Eqs. (70) and (71).
5:     Update $\boldsymbol{\mu}_{\mathbf{w}}$ and $\sigma_{\mathbf{w}}^2$ with Eqs. (72) and (73).
6:     Update b, $\mathbf{a}$, and z with Eqs. (74) to (76).
7: **until** convergence.
8: Output $\hat{\mathbf{x}} = \boldsymbol{\mu}_{\mathbf{x}}$.

---

matrix is diagonal. The mean and elementwise variance of $q(\mathbf{x})$ are denoted by vectors $\boldsymbol{\mu}_{\mathbf{x}}$ and $\sigma_{\mathbf{x}}^2$, respectively (i.e., the covariance is $\mathrm{diag}(\sigma_{\mathbf{x}}^2)$), and those of $q(\mathbf{w})$ by $\boldsymbol{\mu}_{\mathbf{w}}$ and $\sigma_{\mathbf{w}}^2$. Once the approximate posterior $q(\mathbf{x})$ is obtained, its mode can be taken as the MAP solution $\hat{\mathbf{x}}$, which coincides with the mean $\boldsymbol{\mu}_{\mathbf{x}}$ in the Gaussian case.

To make the approximation as accurate as possible, we maximize the log-evidence lower bound under our model, i.e.,

$$\mathrm{LB}[q(\mathbf{x}, \mathbf{w})] = \int q(\mathbf{x}, \mathbf{w}) \ln \frac{p(\mathbf{x}, \mathbf{w}|\mathbf{y})}{q(\mathbf{x}, \mathbf{w})} d\mathbf{x} d\mathbf{w} + \mathrm{const}. \qquad (68)$$

Substituting Eqs. (66) and (67) along with Eqs. (59) to (61) into Eq. (68), we obtain

$$\begin{aligned}
&\mathrm{LB}(\boldsymbol{\mu}_{\mathbf{x}}, \sigma_{\mathbf{x}}^2, \boldsymbol{\mu}_{\mathbf{w}}, \sigma_{\mathbf{w}}^2, b, \mathbf{a}, z) \\
&= \frac{n}{2}\ln b - \frac{b}{2}\mathbf{j} \cdot \left(\begin{array}{c} (\mathbf{y} - \mathbf{P}\boldsymbol{\mu}_{\mathbf{w}} \circ \boldsymbol{\mu}_{\mathbf{x}})^{\circ 2} + \mathbf{P}^{\circ 2}\sigma_{\mathbf{w}}^2 \circ \sigma_{\mathbf{x}}^2 \\ + (\mathbf{P}\boldsymbol{\mu}_{\mathbf{w}})^{\circ 2} \circ \sigma_{\mathbf{x}}^2 + \mathbf{P}^{\circ 2}\sigma_{\mathbf{w}}^2 \circ \boldsymbol{\mu}_{\mathbf{x}}^{\circ 2} \end{array}\right) \quad (69) \\
&\quad + \frac{1}{2}\mathbf{j} \cdot \ln \mathbf{a} - \frac{1}{2}\mathbf{a} \cdot \sum_{k=1}^{2}\left((\mathbf{G}_k\boldsymbol{\mu}_{\mathbf{x}})^{\circ 2} + \mathbf{G}_k^{\circ 2}\sigma_{\mathbf{x}}^2\right) \\
&\quad + \frac{m}{2}\ln z - \frac{z}{2}\mathbf{i} \cdot \left((\mathbf{F}\boldsymbol{\mu}_{\mathbf{w}})^{\circ 2} + \mathbf{F}^{\circ 2}\sigma_{\mathbf{w}}^2\right) \\
&\quad + \frac{1}{2}\mathbf{j} \cdot \ln \sigma_{\mathbf{x}}^2 + \frac{1}{2}\mathbf{i} \cdot \ln \sigma_{\mathbf{w}}^2 + \mathrm{const}.,
\end{aligned}$$

where vector logarithms are elementwise, $\circ$ in superscripts denotes elementwise power, and $\mathbf{j} \in \mathbb{R}^n$ and $\mathbf{i} \in \mathbb{R}^m$ are vectors of ones used to express summation in terms of dot product. Since the parameters of the approximate posteriors and weight parameters depend on each other, they are updated iteratively as follows. The resulting algorithm is summarized in Algorithm 3.

5.3.1 Clean Image

Setting the derivative of Eq. (69) with respect to $\boldsymbol{\mu}_{\mathbf{x}}$ and $\sigma_{\mathbf{x}}^2$ to zero, the parameters of $q(\mathbf{x})$ are obtained as follows:

$$\boldsymbol{\mu}_{\mathbf{x}} = \left(\begin{array}{c} \mathrm{diag}\left((\mathbf{P}\boldsymbol{\mu}_{\mathbf{w}})^{\circ 2} + \mathbf{P}^{\circ 2}\sigma_{\mathbf{w}}^2\right) \\ + \frac{1}{b}\sum_{k=1}^{2}\mathbf{G}_k^{\mathsf{T}}\mathrm{diag}(\mathbf{a})\mathbf{G}_k \end{array}\right)^{-1} (\mathbf{P}\boldsymbol{\mu}_{\mathbf{w}} \circ \mathbf{y}), \qquad (70)$$

$$\sigma_{\mathbf{x}}^2 = \left(b\left((\mathbf{P}\boldsymbol{\mu}_{\mathbf{w}})^{\circ 2} + \mathbf{P}^{\circ 2}\sigma_{\mathbf{w}}^2\right) + \sum_{k=1}^{2}\mathbf{G}_k^{\circ 2\mathsf{T}}\mathbf{a}\right)^{\circ -1}. \qquad (71)$$

5.3.2 Vignetting Function

Setting the derivative of Eq. (69) with respect to $\boldsymbol{\mu}_{\mathbf{w}}$ and $\sigma_{\mathbf{w}}^2$ to zero, the parameters of $q(\mathbf{w})$ are obtained as follows:

$$\boldsymbol{\mu}_{\mathbf{w}} = \left(\mathbf{P}^{\mathsf{T}}\mathrm{diag}\left(\boldsymbol{\mu}_{\mathbf{x}}^{\circ 2} + \sigma_{\mathbf{x}}^2\right)\mathbf{P} + \frac{z}{b}\mathbf{F}^{\mathsf{T}}\mathbf{F}\right)^{-1}\mathbf{P}^{\mathsf{T}}(\boldsymbol{\mu}_{\mathbf{x}} \circ \mathbf{y}), \qquad (72)$$

$$\sigma_{\mathbf{w}}^2 = \left(b\mathbf{P}^{\circ 2\mathsf{T}}\left(\boldsymbol{\mu}_{\mathbf{x}}^{\circ 2} + \sigma_{\mathbf{x}}^2\right) + z\mathbf{F}^{\circ 2\mathsf{T}}\mathbf{i}\right)^{\circ -1}. \qquad (73)$$

Here, the mean of the vignetting function $\mu_{\mathbf{w}}$ depends on clean image estimate $\mu_{\mathbf{x}}$ in Eq. (60), Thus, the vignetting function is stably estimated in VB inference by considering the noise-free version of the degraded image $\mathbf{y}$, unlike in previous MAP-based methods that consider the noisy degraded image $\mathbf{y}$ only.

5.3.3 Weights

While previous non-VB methods have required manual tuning of weight parameters [29], our VB method can tune such parameters automatically by estimation through inference. Since the optimal values of such parameters greatly vary between images, this grants wider applicability to the proposed VB method than non-VB methods. Specifically, setting the derivative of Eq. (69) with respect to b, a_i, and z to zero, these weight parameters can be obtain as follows:

$$b = \frac{n}{\mathbf{j} \cdot \left(\begin{array}{c} (\mathbf{y} - \mathbf{P}\mu_{\mathbf{w}} \circ \mu_{\mathbf{x}})^{\circ 2} + \mathbf{P}^{\circ 2}\sigma_{\mathbf{w}}^2 \circ \sigma_{\mathbf{x}}^2 \\ + (\mathbf{P}\mu_{\mathbf{w}})^{\circ 2} \circ \sigma_{\mathbf{x}}^2 + \mathbf{P}^{\circ 2}\sigma_{\mathbf{w}}^2 \circ \mu_{\mathbf{x}}^{\circ 2} \end{array} \right)}, \tag{74}$$

$$a_i = \frac{1}{\sum_{k=1}^{2} \left((\mathbf{G}_k\mu_{\mathbf{x}})^{\circ 2} + \mathbf{G}_k^{\circ 2}\sigma_{\mathbf{x}}^2 \right)_i}, \tag{75}$$

$$z = \frac{m}{\mathbf{i} \cdot \left((\mathbf{F}\mu_{\mathbf{w}})^{\circ 2} + \mathbf{F}^{\circ 2}\sigma_{\mathbf{w}}^2 \right)}. \tag{76}$$

Here, a_i depends on the inverse of the local image variation at the ith pixel; thus, it becomes small around an edge (a large variation) and weakens smoothing, thereby preserving the edge. Moreover, VB-specific variance terms appear in the denominators, e.g., terms dependent on the variance of the clean image $\sigma_{\mathbf{x}}^2$ in the update of the local smoothness parameter $\mathbf{a}$ in Eq. (75). This positive term effectively work as additional regularizers [7] that prevent the weight parameter from reaching infinity, thereby contributing to stability.

5.4 Experiments

To evaluate the effectiveness of the proposed devignetting method experimentally, we used synthetic images to assess image quality quantitatively with respect to ground truth, and also real images to examine real-world performance qualitatively.

The proposed method was compared with the previous method used for initialization of the proposed method, i.e., the state-of-the-art method proposed by Cho et al. [34], which reportedly outperforms more classical methods such as the one proposed by Zheng et al. [36]. For the parameters of this previous method, the default values suggested in its original paper were used.

5.4.1 Evaluation for Synthetic Images

A degraded images was prepared synthetically by processing the *Lena* standard test image, whose size was $n = 512 \times 512$ pixels. First, following previous devignetting studies [34], [36], [37], vignetting was produced by the off-axis illumination component of the Kang-Weiss model [40], where the optical center of the vignetting function was assumed to be the center of each image. Then, as in typical restoration studies [17], [29], zero-mean Gaussian noise was added, whose level was measured via the signal-to-noise ratio (SNR) [18]. Here, the focal length parameter f of the Kang-Weiss vignetting was set to 250 to obtain noticeable vignetting, and the standard deviation $\sigma = \frac{1}{\sqrt{b}}$ of the noise was chosen to obtain SNR values of 40 and 20, which correspond to

Table 4: Image quality of Lena (PSNR [dB]).

Noise	Degraded	Previous	Proposed
Weak	9.9	**26.9**	26.1
Strong	9.9	20.8	**23.5**

(a)   (b)   (c)

Fig. 13: Lena images with strong noise: (a) degraded; (b) restored by the previous method; and (c) restored by the proposed method.

(a)   (b)
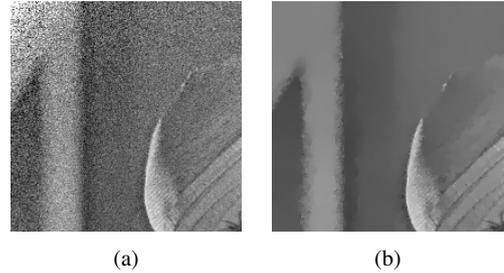
Fig. 14: Closeups of the top-left corners of Lena images with strong noise: (a) restored by the previous method; and (b) restored by the proposed method.

weak and *strong* noise, respectively. For each of these degraded images, restoration by each method was performed to obtain an estimate of the original clean image.

In Table 4, the PSNR values of the degraded and restored images are shown. For visual comparison, the images with strong noise are also shown in Fig. 13. As seen from these results, both methods successfully improved the quality of each degraded image in terms of PSNR. However, the proposed method consistently achieved higher image quality than the previous method when noise was strong. Looking at the restored images, it can be visually seen that the proposed method could remove not only vignetting but also noise. These results demonstrate the stability of the proposed VB approach, which exhibits good restoration performance regardless of noise. Although the previous method achieved a slightly higher PSNR value than that of the proposed method for weak noise, this can be attributed to its naïve devignetting scheme, i.e., pixelwise division with an estimated vignetting function without using any image priors, which does not remove and even preserves noise in ground-truth images. When noise was increased, the previous method overestimated pixel values, as observed in the corners of the closeups in Fig. 14(a). While the whited-out pixels in these regions could be attributed to underestimated values of vignetting functions, they were successfully corrected by the proposed method after refining these estimates through joint inference, as shown in Fig. 14(b). Overall, the performance of the proposed method was stable and comparable with the previous method for weak noise, and better than it for strong noise, as intended in designing the prior-based VB method.
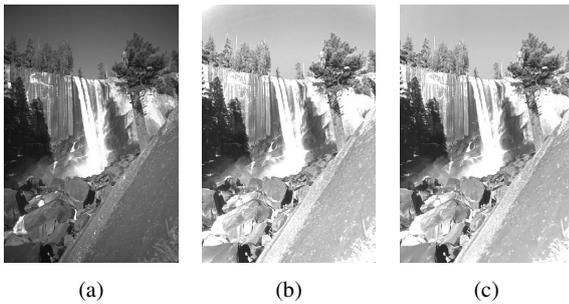
Fig. 15: Real vignetted images: (a) degraded; (b) restored by the previous method; and (c) restored by the proposed method.
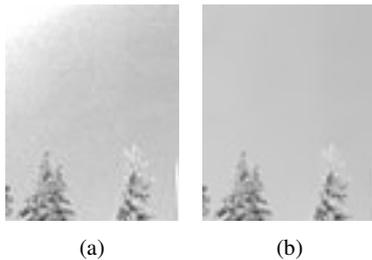
Fig. 16: Closeups of the top-left corners of real vignetted images: (a) restored by the previous method; and (b) restored by the proposed method.

5.4.2 Evaluation for Real Images

Following previous studies [34], [36], an image with real vignetting effect from the Berkeley Segmentation Dataset [41] was used. The optical center of each image was estimated by a previously proposed technique [37] and fed to each method as a known parameter. Since the degradation-free image as ground truth was not available, only qualitative evaluation by visual comparison was possible here.

The results are shown in Fig. 15. Here, both of the two methods yielded images whose peripheral parts were brighter than the corresponding degraded images, which indicate successful devignetting, As seen from the closeup in Fig. 16, however, the proposed method also corrected whiteout artifacts produced by the previous method, and also removed noise in the background sky and compression artifacts around trees, thereby demonstrating the effectiveness of our VB approach in the real world.

5.5 Conclusion

In this section, a method of single-frame devignetting in consideration of noise has been presented. Through prior-based VB inference, the proposed method jointly estimates both a vignetting function and a clean image without vignetting nor noise. Thus, it maintains good performance even when noise is not negligible, as confirmed experimentally.

6. Conclusion

The standard degradation model of image restoration can be adapted to various problems by considering different types of transformation. In this work, we established a VB image restoration methodology, which stably estimates an unknown transformation parameter jointly with a clean image and other parameters from degraded images under flexible parameterization. To this end, we made full use of variational Bayesian techniques to derive an algorithm of principled statistical inference under a problem-specific Bayesian model. Specifically, we developed VB methods for three problems, i.e., multiframe denoising, deblurring, and devignetting. As confirmed in experiments, the proposed VB methods can consistently perform high-quality restoration for a wide range of images, dealing with a variety of transformation stably even in the presence of degradation. Therefore, our VB approach realizes effective image restoration with automatic parameter tuning, which facilitates various image-based applications in the real world.

There are many future perspectives on this work. For example, we may improve the proposed VB methods so that they can handle a wider range of image degradation by introducing more flexible nonlinear transformations and non-Gaussian noise distributions to degradation models. While we used traditional edge-preserving smoothing priors for image models, modern priors such as ones based on deep learning [42] may help the proposed methods to achieve higher image quality. Stochastic approximation techniques such as random sampling [43] may also enable more accurate inference when combined with deterministic VB techniques. Moreover, the proposed VB methodology is general enough to cover other image restoration problems beyond our current scope, e.g., correction of image distortion [2] and exposure [33], both of which can be described by linear transformations with unknown parameters. Finally, application to other fields such as computer vision [2] and computational photography [44], which benefit from high-quality image restoration with automatic parameter tuning enabled by VB, is also of interest.

References

[1] Gonzalez, R. C. and Woods, R. E.: *Digital Image Processing*, Prentice Hall, 3rd edition (2008).

[2] Szeliski, R.: *Computer vision: algorithms and applications*, Springer (2011).

[3] Levin, A., Weiss, Y., Durand, F. and Freeman, W. T.: Understanding and evaluating blind deconvolution algorithms, *Proc. CVPR*, pp. 1964–1971 (2009).

[4] Bishop, C. M.: *Pattern recognition and machine learning*, Springer (2006).

[5] Boyd, S. and Vandenberghe, L.: *Convex Optimization*, Cambridge University Press (2004).

[6] Babacan, S. D., Molina, R. and Katsaggelos, A. K.: Parameter estimation in TV image restoration using variational distribution approximation, *IEEE Trans. Image Process.*, Vol. 17, No. 3, pp. 326–339 (2008).

[7] Babacan, S. D., Molina, R., Do, M. N. and Katsaggelos, A. K.: Bayesian blind deconvolution with general sparse image priors, *Proc. ECCV*, pp. 341–355 (2012).

[8] Ruiz, P., Zhou, X., Mateos, J., Molina, R. and Katsaggelos, A. K.: Variational Bayesian blind image deconvolution: a review, *Digit. Signal Process.*, Vol. 47, pp. 116–127 (2015).

[9] Chambolle, A.: An algorithm for total variation minimization and applications, *J. Math. Imaging Vis.*, Vol. 20, No. 1, pp. 89–97 (2004).

[10] Dong, Y., Hintermüller, M. and Rincon-Camacho, M. M.: Automated Regularization Parameter Selection in multi-scale Total Variation Models for Image Restoration, *J. Math. Imaging Vis.*, Vol. 40, No. 1, pp. 82–104 (2011).

[11] Chen, K., Piccolomini, E. L. and Zama, F.: An automatic regularization parameter selection algorithm in the total variation model for image deblurring, *Numerical Algorithms*, Vol. 67, No. 1, pp. 73–92 (2014).

[12] Telea, A., Preusser, T., Garbe, C., Droske, M. and Rumpf, M.: A variational approach to joint denoising, edge detection and motion estimation, *Pattern Recognition*, pp. 525–535 (2006).

[13] Bar, L., Berkels, B., Rumpf, M. and Sapiro, G.: A variational framework for simultaneous motion estimation and restoration of motion-

blurred video, *Proc. ICCV*, pp. 1–8 (2007).

[14] Sun, D. and Liu, C.: Non-causal temporal prior for video deblocking, *Proc. ECCV*, pp. 510–523 (2012).

[15] Zhang, H. and Carin, L.: Multi-shot imaging: joint alignment, deblurring and resolution-enhancement, *Proc. CVPR*, pp. 2925–2932 (2014).

[16] Liu, C. and Sun, D.: On Bayesian adaptive video super resolution, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 36, No. 2, pp. 346–360 (2014).

[17] Chantas, G., Galatsanos, N., Likas, A. and Saunders, M.: Variational Bayesian image restoration based on a product of t-distributions image prior, *IEEE Trans. Image Process.*, Vol. 17, No. 10, pp. 1795–1805 (2008).

[18] Chantas, G., Galatsanos, N., Molina, R. and Katsaggelos, A. K.: Variational Bayesian image restoration with a product of spatially weighted total variation image priors, *IEEE Trans. Image Process.*, Vol. 19, No. 2, pp. 351–362 (2010).

[19] Tico, M. and Pulli, K.: Low-light imaging solutions for mobile devices, *Proc. ACSSC*, pp. 851–855 (2009).

[20] Horn, B. K. P. and Schunck, B. G.: Determining optical flow, *Artifical Intelligence*, Vol. 17, pp. 185–203 (1981).

[21] Black, M. J. and Anandan, P.: The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, *Computer Vision and Image Understanding*, Vol. 63, No. 1, pp. 75–104 (1996).

[22] Brox, T., Bruhn, A., Papenberg, N. and Weickert, J.: High accuracy optical flow estimation based on a theory for warping, *Proc. ECCV*, pp. 25–36 (2004).

[23] Rudin, L. I., Osher, S. and Fatemi, E.: Nonlinear total variation based noise removal algorithms, *Physica D*, Vol. 60, No. 1, pp. 259–268 (1992).

[24] Hore, A. and Ziou, D.: Image quality metrics: PSNR vs. SSIM, *Proc. ICPR*, pp. 2366–2369 (2010).

[25] Jung, M., Marquina, A. and Vese, L. A.: Variational multiframe restoration of images degraded by noisy (stochastic) blur kernels, *J. Comput. Appl. Math.*, Vol. 240, pp. 123–134 (2013).

[26] Fan, Y., Wei, X. and Qin, S.: Fast and robust deblurring method with multi-frame images based on PSF estimation and total variation optimization, *Optik*, Vol. 124, No. 16, pp. 2285–2291 (2013).

[27] Zhou, X., Mateos, J., Zhou, F., Molina, R. and Katsaggelos, A. K.: Variational Dirichlet blur kernel estimation, *IEEE Trans. Image Process.*, Vol. 24, No. 12, pp. 5127–5139 (2015).

[28] Attias, H.: Inferring parameters and structure of latent variable models by variational Bayes, *Proc. UAI*, pp. 21–30 (1999).

[29] Babacan, S. D., Molina, R. and Katsaggelos, A. K.: Variational Bayesian blind deconvolution using a total variation prior, *IEEE Trans. Image Process.*, Vol. 18, No. 1, pp. 12–26 (2009).

[30] Wipf, D. and Zhang, H.: Revisiting Bayesian blind deconvolution, *J. Mach. Learn. Res.*, Vol. 15, No. 8, pp. 3595–3634 (2014).

[31] Hirsch, M., Schuler, C. J., Harmeling, S. and Schölkopf, B.: Fast removal of non-uniform camera shake, *Proc. ICCV*, pp. 463–470 (2011).

[32] Whyte, O., Sivic, J., Zisserman, A. and Ponce, J.: Non-uniform deblurring for shaken images, *Proc. CVPR* (2010).

[33] Goldman, D. B.: Vignette and exposure calibration and compensation, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 32, No. 12, pp. 2276–2288 (2010).

[34] Cho, H., Lee, H. and Lee, S.: Radial bright channel prior for single image vignetting correction, *Proc. ECCV*, pp. 189–202 (2014).

[35] Zheng, Y., Lin, S., Kambhamettu, C., Yu, J. and Kang, S. B.: Single-image vignetting correction, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 31, No. 12, pp. 2243–2256 (2009).

[36] Zheng, Y., Lin, S., Kang, S. B., Xiao, R., Gee, J. C. and Kambhamettu, C.: Single-image vignetting correction from gradient distribution symmetries, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 35, No. 6, pp. 1480–1494 (2013).

[37] Lopez-Fuentes, L., Oliver, G. and Massanet, S.: Revisiting image vignetting correction by constrained minimization of log-intensity entropy, *Proc. IWANN*, pp. 450–463 (2015).

[38] Lopez-Fuentes, L., Massanet, S. and González-Hidalgo, M.: Image vignetting reduction via a maximization of fuzzy entropy, *Proc. FUZZ-IEEE*, pp. 1–6 (2017).

[39] Petersen, K. B. and Pedersen, M. S.: The Matrix Cookbook, Technical report, Technical University of Denmark (2012).

[40] Kang, S. B. and Weiss, R.: Can we calibrate a camera using an image of a flat, textureless Lambertian surface?, *Proc. ECCV*, pp. 640–653 (2000).

[41] Martin, D., Fowlkes, C., Tal, D. and Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, *Proc. ICCV*, pp. 416–423 (2001).

[42] Zhang, K., Zuo, W., Gu, S. and Zhang, L.: Learning deep CNN denoiser prior for image restoration, *arXiv preprint arXiv:1704.03264* (2017).

[43] Salimans, T., Kingma, D. and Welling, M.: Markov chain Monte Carlo and variational inference: bridging the gap, *Proc. ICML*, pp. 1218–1226 (2015).

[44] Raskar, R.: Computational photography: epsilon to coded photography, *Emerging trends in visual computing*, pp. 238–253 (2009).