BAYESIAN BLIND DECONVOLUTION FROM DIFFERENTLY EXPOSED IMAGE PAIRS

S. Derin Babacan[1], Jingnan Wang[1], Rafael Molina[2], Aggelos K. Katsaggelos[1]

[1]EECS Department
Northwestern University, Evanston, IL USA
sdb@northwestern.edu, jwa180@eecs.northwestern.edu,
aggk@eecs.northwestern.edu

[2]Departamento de Ciencias
de la Computación e I.A.
Universidad de Granada, Granada, Spain
rms@decsai.ugr.es

ABSTRACT

Photographs acquired under low-light conditions require long exposure times and therefore exhibit significant blurring due to the shaking of the camera. Using shorter exposure times results in sharper images but with a very high level of noise. In this paper we address this problem and present a novel blind deconvolution algorithm for a pair of differently exposed images. We formulate the problem in a hierarchical Bayesian framework by utilizing prior knowledge on the unknown image and blur, and also on the dependency between two observed images. By incorporating a fully Bayesian analysis, the developed algorithm estimates all necessary algorithm parameters along with the unknowns, such that no user-intervention is needed. Moreover, we employ a variational Bayesian inference procedure, which allows for the statistical compensation of errors occurring at different stages of the restoration, and also provides uncertainties of the estimates. Experimental results demonstrate the high restoration performance of the proposed algorithm.

Index Terms— Blind deconvolution, Bayesian methods, variational distribution approximations, image stabilization, parameter estimation.

1. INTRODUCTION

Taking high-quality photographs under low-light conditions is a major challenge. A longer exposure time than usual is required to obtain an image with low-noise, but any motion of the camera during exposure causes blur in the recorded image. On the other hand, a short exposure time will result in an image with a very high level of noise. Possible solutions include increasing the light sensitivity (ISO) of the camera sensor, which increases the noise level; increasing the aperture, which results in a smaller depth of field in the acquired image; and using a tripod to stabilize the camera which is not practical in many cases. Digital image stabilization methods, applied at a post-processing stage, provide a powerful means to obtain high-quality images using the low-quality observations.

This paper addresses the problem of blind deconvolution from a short- and long-exposed image pair. A number of methods are proposed for blind deconvolution of a single observation (see, for example, [1] for a recent review), but due to the challenging nature of the problem, obtaining a high-quality restoration result is very hard in most cases. Multi-frame blind deconvolution [1–3] methods attempt to decrease the ill-posedness of the problem by combining information from a set of images, and therefore the restoration performance can be significantly improved. Another possible approach is to discard the blurred image and apply denoising algorithms to the

sharp short-exposed image. Although many very advanced denoising methods are available, the noise level is too high so that features of the underlying image are concealed, and the denoising algorithms cannot easily separate image and noise.

The specific case of blind deconvolution from a pair of short- and long-exposed images has been considered in [4–6]. In [4], the blur point spread function (PSF) is identified separately from the image by Tikhonov regularization and hysteresis thresholding, and then classical image restoration methods are utilized in order to restore the original image. On the other hand, a joint identification method is proposed in [5], where the unknown image and the PSF are estimated simultaneously. The image is modeled using a total-variation (TV) based prior, and the blurs are estimated by imposing the constraint that the blur on the short exposure image is very small. No explicit blur model is utilized in this work, and the blur is denoised by thresholding in an ad hoc manner. Finally, sparsity priors along with continuity constraints on the blurs are utilized in [6], and the image is modeled using a mixture-of-Gaussians prior on the image derivatives. However, the model is derived in a somewhat ad hoc manner, and the resulting algorithm has many parameters to tune, which makes it hard to apply to a wide range of images.

In this paper we systematically model the unknowns within a novel hierarchical Bayesian formulation and develop a blind deconvolution algorithm which jointly estimates the unknown image and blur. We utilize a TV-prior on the image and a sparsity prior on the blur which also imposes positivity. By incorporating a fully-Bayesian approach, all required parameters are estimated along with the unknowns, so that the proposed algorithm does not require user-intervention. Finally, our inference procedure is based on variational distribution approximation, which provides estimates of the distributions of the unknowns. These distributions are implicitly used to incorporate the uncertainties of estimates and to compensate for the estimation errors.

The rest of this paper is organized as follows. In Sec. 2 we formulate the image acquisition processes mathematically. The unknown variables in our model are cast into a hierarchical Bayesian framework as presented in Sec. 3. The variational inference to estimate the unknowns and the proposed algorithm are presented in Sec. 4. Experimental results are presented in Sec. 5 and conclusions are drawn in Sec. 6.

2. PROBLEM FORMULATION

We assume a linear and space invariant degradation model, so that the observation processes can mathematically be expressed as follows:

$$\mathbf{y}_1 = \mathbf{H}\mathbf{x} + \mathbf{n}_1 \tag{1}$$

$$\mathbf{y}_2 = \mathbf{x} + \mathbf{n}_2, \tag{2}$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are the observed images, $\mathbf{x}$ the unknown original image, $\mathbf{n}_1$ and $\mathbf{n}_2$ the noise components. We use matrix-vector notation throughout the paper, so that the images $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{x}$, $\mathbf{n}_1$ and $\mathbf{n}_2$ are $N \times 1$ vectors, where N is the number of pixels in each image. For the sake of simplicity, photometric and geometric calibration between the images $\mathbf{y}_1$ and $\mathbf{y}_2$ is not taken into account in (1) and (2). Generally, the average luminance levels of these images are significantly different due to different exposure times, and the images have to be geometrically registered because of the camera movement between the two acquisitions. In this work, we register the images both photometrically and geometrically as a pre-processing step using standard techniques (see, for example, [4]).

Using (1) and (2), the restoration problem is then to find an estimate of $\mathbf{x}$ from $\mathbf{y}_1$ and $\mathbf{y}_2$ using prior knowledge about $\mathbf{H}$, $\mathbf{n}_1$, $\mathbf{n}_2$ and $\mathbf{x}$.

3. HIERARCHICAL BAYESIAN MODEL

The proposed hierarchical model is composed of two stages. In the first stage, prior distributions are utilized for the observation, the unknown image and the blur. The parameters of these distributions are called *hyperparameters*, which are modeled by hyperprior distributions in the second stage. In the following subsections we present these prior models which form the proposed hierarchical Bayesian model.

3.1. Observation models

We assume that the observation noise in both images follows independent Gaussian distributions, that is, from (1) and (2),

$$p(\mathbf{y}_1|\mathbf{x},\mathbf{h},\beta_1) \propto \beta_1^{N/2} \exp\left[-\frac{\beta_1}{2} \parallel \mathbf{y}_1 - \mathbf{Hx} \parallel^2\right], \qquad (3)$$

and

$$p(\mathbf{y}_2|\mathbf{x},\beta_2) \propto \beta_2^{N/2} \exp\left[-\frac{\beta_2}{2} \parallel \mathbf{y}_2 - \mathbf{x} \parallel^2\right], \qquad (4)$$

where β_1 and β_2 are precisions (inverse variances) of the noises, with $\beta_1 \gg \beta_2$.

Note that the dependency between the observations $\mathbf{y}_1$ and $\mathbf{y}_2$ is very high, as they are images of the same scene. To exploit this dependency, we incorporate the coprimeness condition employed in some multichannel blind deconvolution methods (see, for example, [7]) by combining (1) and (2) to obtain that given $\mathbf{h}$

$$\mathbf{y}_1 - \mathbf{Y}_2\mathbf{h} = \mathcal{N}(\underline{0}, \beta_1 I + \beta_2 \mathbf{H}\mathbf{H}^T) \qquad (5)$$

We note that the model in (5) is obtained from (3) and (4) and so from the statistical point of view it does not provide more information on the unknown blur. However we have experimentally observed that defining the observation model

$$\begin{aligned}p(\mathbf{y}_1,\mathbf{y}_2|\mathbf{x},\mathbf{h},\beta_1,\beta_2,\beta_{12}) \quad &\propto \quad p(\mathbf{y}_1|\mathbf{x},\mathbf{h},\beta_1)p(\mathbf{y}_2|\mathbf{x},\beta_2) \\ &\times \beta_{12}^{N/2}\exp\left[-\frac{\beta_{12}}{2} \parallel \mathbf{y}_1 - \mathbf{Y}_2\mathbf{h} \parallel^2\right],\end{aligned}(6)$$

with $\beta_{12} > 0$ produces both a better restored image and a better estimated blur. Notice that in $p(\mathbf{y}_1,\mathbf{y}_2|\mathbf{x},\mathbf{h},\beta_1,\beta_2,\beta_12)$ we have introduced a third independent observation model which is modelled independently from the ones in (3) and (4). Note also that the noise of this additional observation is modelled as independent white noise with variance β_{12}^{-1}.

3.2. Prior model on the blur

Since the blur is mainly caused by the shaking of the camera during the long exposure time, it exhibits the characteristics of the nonuniform motion blur. Hence, it is expected to be very sparse, i.e., most

of the PSF coefficients being zero or very small. In order to exploit this information, we utilize a mixture prior of D exponential distributions on each PSF coefficient, that is,

$$p(\mathbf{h}|\{\tau_{jd}\}, \{\sigma_{jd}\}) = \prod_j \sum_{d=1}^{D} \tau_{jd} \, \mathrm{Expon}\left(h_j \mid \sigma_{jd}\right) \qquad (7)$$

with τ_{jd} the mixture coefficients for each pixel j and

$$\mathrm{Expon}\left(h_j \mid \sigma_{jd}\right) = \begin{cases} \sigma_{jd}\exp\left(-\sigma_{jd}h_j\right) & \text{if } h_j \geq 0, \\ 0 & \text{if } h_j < 0. \end{cases} \qquad (8)$$

Note that this prior enforces sparsity to a great extent, and the degree of sparsity is increased by increasing the number of mixture coefficients. In addition to imposing sparsity, note that (8) also imposes positivity on the blur coefficients h_j [8, 9]. This property makes the prior especially useful, since unlike most previous work the positivity constraint is considered during the optimization process, and not imposed artificially after the optimization.

3.3. Prior model on the image

As the prior model for the image $\mathbf{x}$ we utilize the quadratic approximation of the TV prior [10]

$$p(\mathbf{x}|\alpha_{\mathrm{im}}) \propto \alpha_{\mathrm{im}}^{N/2} \exp\left[-\frac{1}{2}\alpha_{\mathrm{im}}\mathrm{TV}(\mathbf{x})\right], \qquad (9)$$

where

$$\mathrm{TV}(\mathbf{x}) = \sum_j \sqrt{(\Delta_j^h(\mathbf{x}))^2 + (\Delta_j^v(\mathbf{x}))^2}. \qquad (10)$$

The operators $\Delta_j^h(\mathbf{x})$ and $\Delta_j^v(\mathbf{x})$ correspond to, respectively, horizontal and vertical first order differences, at pixel j, that is, $\Delta_j^h(\mathbf{x}) = x_j - x_{l(j)}$ and $\Delta_j^v(\mathbf{x}) = x_j - x_{a(j)}$, where $l(j)$ and $a(j)$ denote the nearest neighbors of j, to the left and above, respectively. The TV function has the advantage of preserving the edge structure while imposing smoothness on the solutions.

3.4. Hyperpriors on the hyperparameters

In the second stage of the hierarchical model, we employ Gamma priors on α_{im}, β_1, β_2, β_{12}, σ_{jd} and Dirichlet priors on mixture coefficients $\{\tau_{jd}\}_{d=1}^{D}$ in order to obtain a tractable Bayesian analysis, as these are the conjugate distributions to their respective priors. The hyperpriors can be expressed as

$$p(\alpha_{\mathrm{im}}) = \mathrm{Gamma}\left(\alpha_{\mathrm{im}}|a^{(\alpha_{\mathrm{im}})}, b^{(\alpha_{\mathrm{im}})}\right) \qquad (11)$$

$$p(\beta_1) = \mathrm{Gamma}\left(\beta_1|a^{(\beta)}, b^{(\beta)}\right) \qquad (12)$$

$$p(\beta_2) = \mathrm{Gamma}\left(\beta_2|a^{(\beta)}, b^{(\beta)}\right) \qquad (13)$$

$$p(\beta_{12}) = \mathrm{Gamma}\left(\beta_{12}|a^{(\beta)}, b^{(\beta)}\right) \qquad (14)$$

$$p(\sigma_{jd}) = \mathrm{Gamma}\left(\sigma_{jd}|a^{(\sigma_{jd})}, b^{(\sigma_{jd})}\right) \qquad (15)$$

$$p(\{\tau_{jd}\}_{d=1}^{D}) = \mathrm{Dirichlet}\left(\{\tau_{jd}\}_{d=1}^{D}|c^{(\tau_j)}\right) \qquad (16)$$

The shape and scale parameters a and b of the Gamma distributions are set to a small common value (e.g., 0.01) to make the estimation process rely more on the observations than prior knowledge. However, we found out in our experiments that the proposed algorithm is very robust to the selection of these parameters.

Finally, combining the first and second stage of the model we obtain the following global distribution

$$p(\mathbf{x}, \mathbf{h}, \mathbf{y}_1, \mathbf{y}_2, \alpha_{\mathrm{im}}, \beta_1, \beta_2, \beta_{12}, \{\tau_{jd}\}, \{\sigma_{jd}\})$$
$$= p(\mathbf{x}|\alpha_{\mathrm{im}})\, p(\mathbf{h}|\{\tau_{jd}\}, \{\sigma_{jd}\}) p(\mathbf{y}_1, \mathbf{y}_2 | \mathbf{x}, \mathbf{h}, \beta_1, \beta_2, \beta_{12})$$
$$\times\; p(\alpha_{\mathrm{im}}) p(\beta_1)\, p(\beta_2) p(\beta_{12}) \prod_{j=1}^{N} \left[p(\{\tau_{jd}\}_{d=1}^{D}) \prod_{d=1}^{D} p(\sigma_{jd}) \right].$$

4. VARIATIONAL INFERENCE

In Bayesian formulations, the inference is based on the posterior distribution, which in our case is intractable. Therefore, in this work we utilize variational distribution approximations. Let us denote by Θ the set of unknowns, i.e., $\Theta = \{\mathbf{x}, \mathbf{h}, \alpha_{\mathrm{im}}, \beta_1, \beta_2, \beta_{12}, \{\sigma_{jd}\}, \{\tau_{jd}\}\}$. The goal is to approximate the posterior distribution $p(\Theta|\mathbf{y}_1, \mathbf{y}_2)$ by another distribution $q(\Theta)$ which allows a tractable analysis. The assumption made during this approximation is that the approximate posterior distribution is separable, that is,

$$q(\Theta) = q(\mathbf{x}) q(\mathbf{h}) q(\alpha_{\mathrm{im}}) q(\beta_1) q(\beta_2) q(\beta_{12}) \prod_{d}^{D} q(\sigma_{jd}) q(\tau_{jd}).$$

Unfortunately the general results of the variational Bayesian analysis cannot be directly utilized due to the TV and mixture priors in our model. The problems caused by the TV prior can be avoided by utilizing a majorization-minimization approach, whose details are given in [10] and are omitted here. In order to utilize the mixture priors, the Jensen's inequality is utilized as follows [9]

$$\log \prod_j \sum_d \tau_{jd} \,\mathrm{Expon}\,(h_j \mid \sigma_{jd})$$
$$\leq \sum_j \sum_d \mu_{jd} \log \left(\frac{\tau_{jd}}{\mu_{jd}} \mathrm{Expon}\,(h_j \mid \sigma_{jd}) \right) \quad (17)$$

with $\sum_{d=1}^{D} \mu_{jd} = 1$, $j = 1, \ldots, N$. Using the right hand side of (17) and the quadratic upper bound of the TV prior [10], we obtain an upper bound for the true posterior distribution $\mathbf{F}(\Theta)$. The distribution approximation $q(\theta)$ of each unknown $\theta \in \Theta$ can then be found by minimizing the Kullback-Leibner divergence between this upper bound $\mathbf{F}(\Theta)$ and $q(\theta)$, which results in the following general solution [11]

$$q(\theta) = \mathrm{const} \times \exp\left(E_{q(\Theta_\theta)}\left[\log \mathbf{F}(\Theta) \right] \right), \quad (18)$$

where Θ_θ denotes the set Θ with θ removed from the set, and $E_{q(\Theta_\theta)}[\cdot]$ denotes the expectation with respect to the distribution $q(\Theta_\theta)$. The posterior approximations $q(\theta)$ can be computed by holding $q(\Theta_\theta)$ constant and solving (18) with respect to $q(\theta)$.

In this work we additionally assume that $q(\mathbf{x})$ and $q(\mathbf{h})$ are degenerate distributions and denote by $\hat{\mathbf{x}}$ and $\hat{\mathbf{h}}$, respectively, the image and blur where these distributions are degenerate. Using this approximation, we obtain

$$q(\alpha_{\mathrm{im}}) = \mathrm{Gamma}\left(\alpha_{\mathrm{im}} | \bar{a}^{(\alpha_{\mathrm{im}})}, \bar{b}^{(\alpha_{\mathrm{im}})} \right) \quad (19)$$

$$q(\beta_1) = \mathrm{Gamma}\left(\beta_1 | \bar{a}^{(\beta_1)}, \bar{b}^{(\beta_1)} \right) \quad (20)$$

$$q(\beta_2) = \mathrm{Gamma}\left(\beta_2 | \bar{a}^{(\beta_2)}, \bar{b}^{(\beta_2)} \right) \quad (21)$$

$$q(\beta_{12}) = \mathrm{Gamma}\left(\beta_{12} | \bar{a}^{(\beta_{12})}, \bar{b}^{(\beta_{12})} \right) \quad (22)$$

$$q(\sigma_{jd}) = \mathrm{Gamma}\left(\sigma_{jd} | \bar{a}^{(\sigma_{jd})}, \bar{b}^{(\sigma_{jd})} \right) \quad (23)$$

$$q(\{\tau_{jd}\}_{d=1}^{D}) = \mathrm{Dirichlet}\left(\{\tau_{jd}\}_{d=1}^{D} | \{\bar{c}^{(\tau_{jd})}\}_{d=1}^{D} \right) \quad (24)$$

whose means are given by

$$E(\alpha_{\mathrm{im}}) = \frac{\bar{b}^{(\alpha_{\mathrm{im}})}}{\bar{a}^{(\alpha_{\mathrm{im}})}} = \frac{b^{(\alpha_{\mathrm{im}})} + \frac{N}{2}}{a^{(\alpha_{\mathrm{im}})} + \sum_j \sqrt{w_j}} \quad (25)$$

$$E(\beta_1) = \frac{\bar{b}^{(\beta_1)}}{\bar{a}^{(\beta_1)}} = \frac{b^{(\beta_1)} + \frac{N}{2}}{a^{(\beta_1)} + \frac{1}{2}\| \mathbf{y}_1 - \hat{\mathbf{H}}\hat{\mathbf{x}} \|^2} \quad (26)$$

$$E(\beta_2) = \frac{\bar{b}^{(\beta_2)}}{\bar{a}^{(\beta_2)}} = \frac{b^{(\beta_2)} + \frac{N}{2}}{a^{(\beta_2)} + \frac{1}{2}\| \mathbf{y}_2 - \hat{\mathbf{x}} \|^2} \quad (27)$$

$$E(\beta_{12}) = \frac{\bar{b}^{(\beta_{12})}}{\bar{a}^{(\beta_{12})}} = \frac{b^{(\beta_{12})} + \frac{N}{2}}{a^{(\beta_{12})} + \frac{1}{2}\| \mathbf{y}_1 - \mathbf{Y}_2\hat{\mathbf{h}} \|^2} \quad (28)$$

$$E(\sigma_{jd}) = \frac{\bar{b}^{(\sigma_{jd})}}{\bar{a}^{(\sigma_{jd})}} = \frac{b^{(\sigma_{jd})} + \mu_{jd}}{a^{(\sigma_{jd})} + \mu_{jd}\, h_j} \quad (29)$$

$$\mu_{jd} \propto \tau_{jd}\,\mathrm{Expon}\left(\hat{h}_j|\sigma_{jd} \right), \; \sum_{d=1}^{D} \mu_{jd} = 1 \quad (30)$$

$$\bar{c}^{(\tau_{jd})} = c^{(\tau_{jd})} + \mu_{jd} \quad (31)$$

where $E(\cdot)$ denotes the mean of the distribution, and

$$w_j = (\Delta_j^h(\hat{\mathbf{x}}))^2 + (\Delta_j^v(\hat{\mathbf{x}}))^2 \; j = 1, \ldots, N, \quad (32)$$

$$\mathbf{W} = \mathrm{diag}\left(\frac{1}{\sqrt{w_j}} \right), \; j = 1, \ldots, N. \quad (33)$$

The matrix $\mathbf{W}$ in (33) is the spatial adaptivity matrix which controls the amount of smoothing at each pixel location depending on the intensity variation at that pixel, as expressed by the vector $\mathbf{w}$ representing the total variation of the estimated image.

The restored image $\hat{\mathbf{x}}$ satisfies

$$\hat{\mathbf{x}} = \Sigma_{\mathbf{x}}\left(E(\beta_1)\hat{\mathbf{H}}^T\mathbf{y}_1 + E(\beta_2)\mathbf{y}_2 \right) \quad (34)$$

$$\Sigma_{\mathbf{x}}^{-1} = E(\alpha_{\mathrm{im}})(\Delta^h)^T\mathbf{W}(\Delta^h) + E(\alpha_{\mathrm{im}})(\Delta^v)^T\mathbf{W}(\Delta^v)$$
$$+ E(\beta_1)\hat{\mathbf{H}}^T\hat{\mathbf{H}} + E(\beta_2)\mathbf{I} \quad (35)$$

The blur estimate $\hat{\mathbf{h}}$ is found by maximizing the posterior distribution, that is,

$$\hat{\mathbf{h}} = \underset{\mathbf{h}}{\mathrm{argmax}} \; p(\hat{\mathbf{x}}, \mathbf{h}, \mathbf{y}_1, \mathbf{y}_2, \alpha_{\mathrm{im}}, \beta_1, \beta_2, \beta_{12}, \{\tau_{jd}\}, \{\sigma_{jd}\}),$$

where parameters are fixed to the means of their corresponding $q(\cdot)$ distributions. Therefore we obtain

$$\tilde{\mathbf{h}} = \beta_1\mathbf{X}^T\mathbf{X} + \beta_{12}\mathbf{Y}_2^T\mathbf{Y}_2 \quad (36)$$

$$\hat{\mathbf{h}} = (\tilde{\mathbf{h}})^{-1}\left(-\mathrm{diag}\left(\sum_{d=1}^{D} \sigma_{jd}^{-1}\mu_{jd} \right) + \beta_1\mathbf{X}^T\mathbf{y}_1 + \beta_{12}\mathbf{Y}_2^T\mathbf{y}_1 \right). \quad (37)$$

However, $\hat{\mathbf{h}}$ can have negative values. To solve this problem we consider $\tilde{h}_j$ and $\hat{h}_j$ to be the parameters of a rectified Gaussian distribution $\mathcal{N}^R\left(h_j|\hat{h}_j, \tilde{h}_j \right)$ of h_j and use the means of these distributions given by [9]

$$E(h_j) = \hat{h}_j + \sqrt{\frac{2}{\pi\tilde{h}_j}}\frac{1}{\mathrm{erfcx}(-\hat{h}_j\sqrt{\frac{\tilde{h}_j}{2}})}, \quad (38)$$

where $\mathrm{erfcx}(\cdot)$ is the scaled complementary error function as the components of the blur estimate instead of $\hat{\mathbf{h}}$.

Finally, the algorithm iterates among estimating the image using (34), estimating the blur using (38) and then estimating the hyperparameters using (25)-(31).
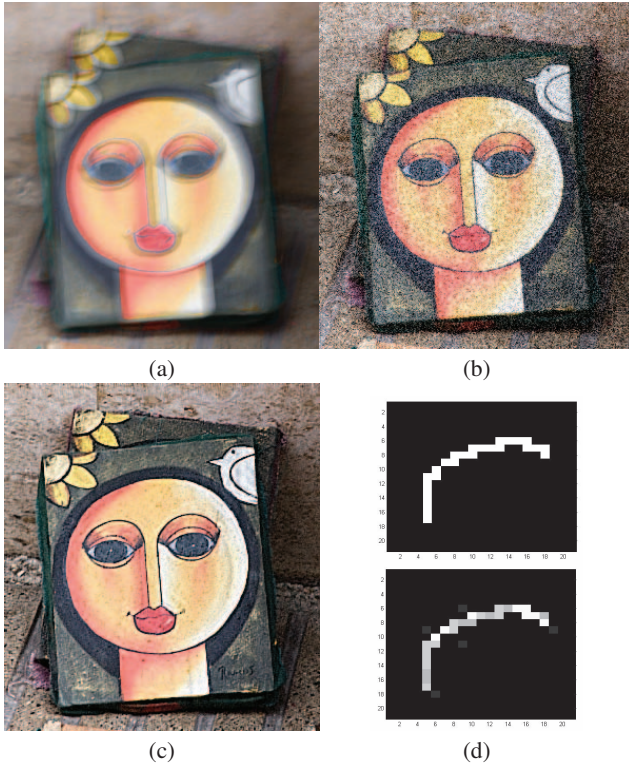
Fig. 1. (a) Observed blurred image simulating a long-exposure acquisition, (b) observed noisy image simulating a short-exposure acquisition, (c) restored image using the proposed algorithm, (d) (top) original blur PSF, (bottom) blur PSF estimated by the proposed method.

5. EXPERIMENTAL RESULTS

The proposed algorithm has been evaluated extensively with both synthetic and real images. A synthetically generated image pair is shown in Fig. 1, where the image in Fig. 1(a) is the long-exposure observation $\mathbf{y}_1$ blurred with the PSF shown in Fig. 1(d) (top), which simulates the shake of the camera during a long-exposure. The short-exposure noisy image $\mathbf{y}_2$ is shown in Fig. 1(b). White Gaussian noise with variances 1.6 and 1000 are added to obtain degraded observations $\mathbf{y}_1$ and $\mathbf{y}_2$, corresponding to signal-to-noise ratios (SNR) of 42dB and 5dB, respectively.

The observed image $\mathbf{y}_1$ is used as the initial estimate of $\mathbf{x}$, and the initial estimate of $\mathbf{h}$ is calculated by the division of the observations $\mathbf{y}_1$ and $\mathbf{y}_2$ in the frequency domain using their Fourier transforms. Note that although the initial PSF calculated in this fashion is a very crude estimate to the original PSF, it provides a very fast initialization of the algorithm. The number of mixture distributions is set to $D = 3$, but other values gave similar results. All other parameters are calculated using (25)-(31). As the convergence criterion we use the ratio of the norm of the difference of the image estimates between two consecutive iterations and the norm of the last image estimate, and the algorithm is stopped when this ratio is smaller than 10^{-5}. The convergence is generally achieved within 20 iterations, where each iteration takes approximately 10 seconds in Matlab running on a Pentium Core2 CPU at 2.66 GHz.

The restored image is shown in Fig. 1(c), and the recovered PSF is shown in Fig. 1(d) (bottom). Note that the support of the PSF used in this experiment is 21×21, and the original image is of size 300×300, so that the degradation caused by the blur

is severe. Moreover, the noise level in the second observed image is very high, which makes the estimation of the blur PSF difficult. Even with these severely degraded observations, the proposed algorithm provides high-quality restoration results shown in Fig. 1(c)-(d). The mean-squared error between the restored and the original image is 48.5, and the running time of the algorithm for this image set was approximately four minutes. The proposed method also produces high quality restorations with real images and compares favorably with other algorithms, which are not presented in this paper due to space limitations (some additional results can be found in http://ivpl.eecs.northwestern.edu/research/topics/image-and-video-recovery).

6. CONCLUSIONS

In this paper we presented a novel Bayesian formulation for blind deconvolution from differently exposed image pairs. The unknown image, blur and model parameters, including the noise variance, is estimated solely from the observations without prior knowledge or user intervention. The developed algorithm simultaneously estimates the distributions of the unknowns which allows for the computation of the estimation uncertainties and also incorporates these uncertainties within the restoration procedure. We have shown that although the algorithm is fully-automated and no ad hoc methods (such as blur thresholding or denoising) is performed, it provides very high quality restored images even with high degradations.

7. REFERENCES

[1] T. E. Bishop, S. D. Babacan, B. Amizic, A. K. Katsaggelos, T. Chan, and R. Molina, "Blind image deconvolution: problem formulation and existing approaches," in *Blind image deconvolution: theory and applications*, P. Campisi and K. Egiazarian, Eds., chapter 1. CRC press, 2007.

[2] A. Rav-Acha and S. Peleg, "Two motion-blurred images are better than one," *Pattern Recogn. Lett.*, vol. 26, no. 3, pp. 311–317, 2005.

[3] F. Šroubek and J. Flusser, "Multichannel blind iterative image restoration," *IEEE Trans. Image Processing*, vol. 12, no. 9, pp. 1094–1106, 2003.

[4] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," in *SIGGRAPH '07*, 2007, p. 1.

[5] M. Tico and M. Vehvilainen, "Image stabilization based on fusing the visual information in differently exposed images," *ICIP 2007*, vol. 1, pp. 117–120, 16 2007-Oct. 19 2007.

[6] J. Chen, L. Yuan, C.-K. Tang, and L. Quan, "Robust dual motion deblurring," *CVPR 2008*, pp. 1–8, June 2008.

[7] F. Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Processing*, vol. 7, pp. 45–53, July 2005.

[8] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph," *SIGGRAPH 2006*, vol. 25, pp. 787–794, 2006.

[9] J. Miskin, *Ensemble Learning for Independent Component Analysis*, Ph.D. thesis, Astrophysics Group, University of Cambridge, 2000.

[10] S. D. Babacan, R. Molina, and A.K. Katsaggelos, "Parameter estimation in TV image restoration using variational distribution approximation," *IEEE Trans. Image Processing*, , no. 3, pp. 326–339, 2008.

[11] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, 2006.