Bayesian Deblurring with Integrated Noise Estimation

Uwe Schmidt* Kevin Schelten* Stefan Roth
Department of Computer Science, TU Darmstadt

Abstract

Conventional non-blind image deblurring algorithms involve natural image priors and maximum a-posteriori (MAP) estimation. As a consequence of MAP estimation, separate pre-processing steps such as noise estimation and training of the regularization parameter are necessary to avoid user interaction. Moreover, MAP estimates involving standard natural image priors have been found lacking in terms of restoration performance. To address these issues we introduce an integrated Bayesian framework that unifies non-blind deblurring and noise estimation, thus freeing the user of tediously pre-determining a noise level. A sampling-based technique allows to integrate out the unknown noise level and to perform deblurring using the Bayesian minimum mean squared error estimate (MMSE), which requires no regularization parameter and yields higher performance than MAP estimates when combined with a learned high-order image prior. A quantitative evaluation demonstrates state-of-the-art results for both non-blind deblurring and noise estimation.

1. Introduction

Although image blur is sometimes used for artistic purposes, it frequently corrupts valuable image information and produces visually disturbing images. Deblurring techniques thus attempt to restore a sharp explanation from a blurred input image. This paper is concerned with *non-blind* deblurring, where the blur is assumed to be known, which is an important component of the more general blind deblurring problem [14]. Even when the image blur is known, for example from inertial sensors [8], deblurring is a difficult problem, partly rooted in the loss of high spatial frequencies due to the blur. This is further exacerbated by image noise, which arises from the image capturing process. The simplest conceivable approach of deblurring by inverting the blur matrix cannot effectively overcome these difficulties.

Instead, most techniques rely on natural image priors based on Markov random fields (MRFs) to regularize the problem and to avoid restoration artifacts. The restored image is coupled to the blurred input through a likelihood model, which requires the image noise level to be known. To compute the deblurred image, most approaches rely on MAP estimation [11, 13]. However, such MAP approaches have various shortcomings: *(1)* The noise level must be given by the user or estimated in a separate process prior to deblurring. This is problematic, since the performance can be sensitive to the choice of noise level (see Fig. 1). Moreover, conventional noise estimation techniques are not necessarily designed for blurred images. *(2)* For best results, the influence of prior and likelihood needs to be calibrated with a regularization parameter. This parameter depends on the noise magnitude; therefore, a set of suitable regularization parameters must be determined in an off-line training step for a necessarily limited set of noise levels. *(3)* Finally, MAP estimates have been found to either exhibit oversmoothing or residual high-frequency artifacts [3, 22].

In this paper, we propose a non-blind deblurring algorithm that extends the conventional Bayesian approach with *integrated noise estimation*. In particular, we treat the noise level as a nuisance parameter that is integrated out using a sampling-based algorithm. As a consequence of unifying noise estimation and image restoration, manual noise selection or a separate pre-processing step are no longer needed. Moreover, we replace MAP with MMSE estimation based on Gibbs sampling [22], which reduces oversmoothing and allows to eliminate the regularization parameter when combined with an appropriate image prior. Finally, we employ a learned high-order MRF prior [20] for regularization, which as far as we are aware is the first time learned priors have been used for image deblurring.

We perform a number of quantitative experiments to evaluate both the noise estimation component itself, as well as the integrated deblurring approach; in both cases we find favorable performance. In particular, our non-blind deblurring approach with integrated noise estimation outperforms state-of-the-art deblurring methods for which the noise level is estimated with a separate procedure [30]. In addition, our method outperforms previous ones even when those are supplied with the (usually unknown) ground truth noise level. Finally, we report results for the special case of blind image denoising, for which we obtain results very close to the non-blind case.

*The first two authors contributed equally to this work.

In Proc. of the *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2625–2632, Colorado Springs, CO, 2011. doi:10.1109/CVPR.2011.5995653

2. Related Work

The practical significance of non-blind deblurring not only includes applications with known blur kernel [*e.g.* 13], but also stems from the design of state-of-the-art blind deblurring algorithms. Levin *et al.* [14] theoretically and practically demonstrate the stability of *first* estimating the kernel from a marginalized density, and *then* inferring a sharp image by non-blind deblurring with the kernel estimate. Although some algorithms differ from this and alternate between kernel and latent image estimation [*e.g.* 23], non-blind deblurring forms an important element of not only state-of-the-art image-based blind methods [2, 6, 9, 23, 27], but also of highly competitive hardware-based techniques [1, 8, 24].

A classical approach to non-blind deblurring is the Lucy-Richardson method [16, 19], which is known to cause ringing artifacts, but frequently is used as a baseline [11, 23]. Ringing can be reduced if a pair of blurred and noisy images is available [27]. More recent deblurring approaches involve natural image priors and MAP estimation. Krishnan and Fergus [11] propose a fast MAP algorithm with a first-order prior by exploiting a half-quadratic scheme [7]. Shan *et al.* [23] use a likelihood model based on derivatives of up to second order and a two-level prior. Levin *et al.* [13] propose a MAP algorithm with a high-order prior involving second-order derivative filters. This method has found widespread use due to its leading performance [2, 8, 14], or is often compared to experimentally [11, 23].

As discussed above, MAP approaches to deblurring have a number of disadvantages. On one hand, MAP estimates have been found to yield either oversmoothed results lacking textural detail, or results with high-frequency artifacts. Cho *et al.* [3] address this issue with a content-aware prior, which adapts the image model locally to the properties of the respective image region. Schmidt *et al.* [22] replace MAP with MMSE estimation, and show improved results for image denoising. We here follow the approach of [22] and generalize it to non-blind deblurring. Additionally, the sampling framework allows to integrate noise estimation in a rather natural way.

On the other hand, MAP approaches require the noise level and/or regularization parameter to be known or estimated separately. This has been addressed using variational Bayesian techniques that approximate the posterior by a simpler, analytically tractable density. One can thus compute marginal expectations of the hidden variables, including the noise level, under the approximative distribution. Miskin and MacKay [17] propose a variational Bayesian framework for blind deconvolution with integrated noise estimation, but assume that pixels are i. i. d., which leads to sub-par deblurring results. Fergus *et al.* [6] incorporate a similar automatic noise estimate into kernel estimation, but do so for noise on the image gradients instead of image noise. Moreover, [6] use a simple gradient prior. In contrast, our non-blind deblurring algorithm based on sampling naturally extends to high-order priors, and allows to formulate and estimate sensor noise in the spatial domain.

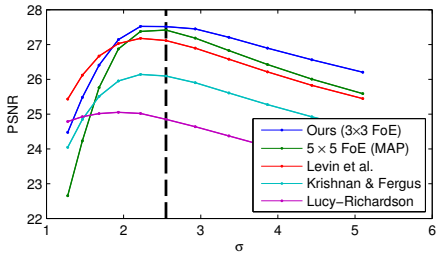


Figure 1. **Dependence of deblurring results on the noise level.** Average results for 8 images and various assumed noise levels (ground truth $\sigma_{\text{GT}} = 2.55$). All methods suffer from an incorrect choice of σ. Especially for MAP-based approaches the best performance is achieved for $\sigma < \sigma_{\text{GT}}$. The trained regularization parameter is thus not fully representative of the test set.

The issue of estimating regularization parameters extends well beyond deblurring. In stereo, Zhang and Seitz [28] address this by performing joint MAP estimation of the disparity and the MRF parameters. In optical flow, Krajsek and Mester [10] marginalize over the flow field based on a Laplace approximation in order to obtain a maximum marginal likelihood estimate for the model parameters. Our approach, in contrast, does not require approximations of the posterior.

Since many image restoration approaches involve noise-dependent tuning parameters, much effort has gone into automatic noise estimation. For a single color image, Liu *et al.* [15] infer the noise level in RGB channels using a piecewise smooth image model. For gray-level images, Zoran and Weiss [30] estimate the noise standard deviation by modeling a link between kurtosis values and image noise. The wavelet-based approach of De Stefano *et al.* [4] follows similar ideas. The widely used MAD framework [5, 29] infers a noise estimate from the wavelet coefficients of the highest-frequency subband. We note that most noise estimation procedures do not explicitly consider the special case of noise inference on blurred images. One exception is [30], which at least reports experimental results for this case. The advantage of our integrated noise estimation approach is that it is directly applicable to deblurring.

3. Deblurring with High-order Priors

As is usual in the deblurring literature, we assume that the unknown sharp image $\mathbf{x} \in \mathbb{R}^m$ is blurred with a blur matrix $\mathbf{K} \in \mathbb{R}^{n \times m}$ and corrupted with additive white Gaussian noise $\mathbf{n}$:

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \mathbf{n}, \qquad \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}). \tag{1}$$

Here, $\mathbf{y} \in \mathbb{R}^n$ is the observed, blurred image. We note that all subsequent derivations also hold for the case of spatially varying, non-uniform blur. However, since there is little ground-truth data available for spatially varying blur, our experiments are limited to spatially uniform blur. In other terms, we usually assume that $\mathbf{K}\mathbf{x} \equiv \mathbf{k} \otimes \mathbf{x}$ corresponds to a convolution of the desired image $\mathbf{x}$ with the blur kernel $\mathbf{k}$.

In this paper, we consider the problem of non-blind image deblurring with a known blur matrix $\mathbf{K}$. Taking a Bayesian approach thus amounts to formulating the posterior

$$p(\mathbf{x}|\mathbf{y}, \mathbf{K}, \sigma) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{K}, \sigma) \cdot p(\mathbf{x}). \tag{2}$$

The assumption of additive white Gaussian noise gives rise to the likelihood

$$p(\mathbf{y}|\mathbf{x}, \mathbf{K}, \sigma) = \mathcal{N}(\mathbf{y}; \mathbf{K}\mathbf{x}, \sigma^2\mathbf{I}). \tag{3}$$

The term $p(\mathbf{x})$ denotes a natural image prior. This prior is necessary, since simply inverting the blur matrix $\mathbf{K}$ to recover the original image $\mathbf{x}$ is infeasible due to the presence of noise or a misspecified blur kernel. Hence, recent deblurring techniques employ sparse image priors to regularize the solution to this ill-posed problem [*e.g.* 11, 13].

High-order MRF prior. Instead of hand-defined, frequently gradient-based priors prevalent in deblurring, we here rely on a learned high-order prior, specifically Fields of Experts (FoEs) [20]. These high-order MRFs are based on spatially extended cliques, and capture the properties of a natural image $\mathbf{x}$ in terms of responses to a bank of learned, linear filters:

$$p(\mathbf{x}) = \frac{1}{Z} e^{-\epsilon\|\mathbf{x}\|^2/2} \prod_{c \in \mathcal{C}} \prod_{i=1}^{N} \phi\big(\mathbf{J}_i^{\mathrm{T}} \mathbf{x}_{(c)}; \boldsymbol{\alpha}_i\big) \tag{4}$$

Here, the response to the learned, linear filters $\mathbf{J}_i$ is modeled as the product of expert functions ϕ with parameters $\boldsymbol{\alpha}_i$, which are learned as well. Moreover, the MRF cliques are denoted as $c \in \mathcal{C}$, and Z denotes the partition function. The broad Gaussian factor $e^{-\epsilon\|\mathbf{x}\|^2/2}$ ensures that the model is normalizable.

Instead of the originally proposed Student-t experts [20], we follow [22] and model the expert functions using Gaussian scale mixtures (GSMs) [25] as

$$\phi(\mathbf{J}_i^{\mathrm{T}} \mathbf{x}_{(c)}; \boldsymbol{\alpha}_i) = \sum\nolimits_{j=1}^{J} \alpha_{ij} \mathcal{N}(\mathbf{J}_i^{\mathrm{T}} \mathbf{x}_{(c)}; 0, \eta_{ij}^2). \tag{5}$$

Apart from being a more faithful model of natural images, this formulation admits efficient Gibbs sampling-based inference [12], which we later exploit for our integrated noise estimation approach. To that end, the FoE density from Eq. (4) can be augmented with discrete latent variables $\mathbf{z}$ to yield the joint distribution

$$p(\mathbf{x}, \mathbf{z}) \propto e^{-\epsilon\|\mathbf{x}\|^2/2} \prod_{c \in \mathcal{C}} \prod_{i=1}^{N} \alpha_{iz_{ic}} \mathcal{N}(\mathbf{J}_i^{\mathrm{T}} \mathbf{x}_{(c)}; 0, \eta_{iz_{ic}}^2), \tag{6}$$

from which the FoE model in Eq. (4) arises by marginalizing over the latent variables $\mathbf{z}$.

The advantage of the augmented distribution $p(\mathbf{x}, \mathbf{z})$ is that the conditional distributions are tractable. Specifically, $p(\mathbf{x}|\mathbf{z})$ is a multivariate Gaussian and $p(\mathbf{z}|\mathbf{x})$ is a discrete distribution [*cf*. 22].

4. Bayesian Deblurring using Sampling

To infer the deblurred image from the posterior in Eq. (2), we extend the approach of [22] and in contrast to previous deblurring methods compute the posterior mean, or Bayesian minimum mean squared error estimate (MMSE):

$$\hat{\mathbf{x}} = \arg\min_{\tilde{\mathbf{x}}} \int \|\tilde{\mathbf{x}} - \mathbf{x}\|^2 p(\mathbf{x}|\mathbf{y}, \mathbf{K}, \sigma)\, \mathrm{d}\mathbf{x} = E[\mathbf{x}|\mathbf{y}, \mathbf{K}, \sigma]. \tag{7}$$

The advantage over the more common MAP approach, at least for image denoising [22], is that that it leads to superior results, both for smooth and textured image regions. Secondly, MMSE estimates yield a higher correlation between the image restoration performance and the generative quality of the model. This on one hand lets us take advantage of powerful learned priors, and on the other hand allows us to work without any regularization parameter that balances the prior and the likelihood, which is very desirable especially when the noise level is not known.

We perform MMSE estimation by extending the sampling approach of [22] to image deblurring. To that end, we note that the posterior can be augmented with discrete latent variables $\mathbf{z}$:

$$p(\mathbf{x}, \mathbf{z}|\mathbf{y}, \mathbf{K}, \sigma) \propto p(\mathbf{y}|\mathbf{x}, \mathbf{K}, \sigma) \cdot p(\mathbf{x}, \mathbf{z}). \tag{8}$$

Due to the Gaussian form of the likelihood from Eq. (3) we obtain the conditional distributions

$$p(\mathbf{z}|\mathbf{x}, \mathbf{y}, \mathbf{K}, \sigma) \propto \prod_{c \in \mathcal{C}} \prod_{i=1}^{N} \alpha_{iz_{ic}} \mathcal{N}(\mathbf{J}_i^{\mathrm{T}} \mathbf{x}_{(c)}; 0, \eta_{iz_{ic}}^2) \tag{9}$$

$$p(\mathbf{x}|\mathbf{z}, \mathbf{y}, \mathbf{K}, \sigma) \propto \mathcal{N}(\mathbf{y}; \mathbf{K}\mathbf{x}, \sigma^2\mathbf{I}) \cdot \mathcal{N}(\mathbf{x}; \mathbf{0}, \mathbf{P}_{\mathbf{z}}^{-1}). \tag{10}$$

Note that the conditional $p(\mathbf{z}|\mathbf{x}, \mathbf{y}, \mathbf{K}, \sigma) = p(\mathbf{z}|\mathbf{x})$ is not affected by the likelihood term and hence is the same as in the denoising case [22]. Sampling from Eq. (9) is easy, because the distribution decomposes and each z_{ic} can be sampled independently from a univariate discrete distribution.

The conditional distribution $p(\mathbf{x}|\mathbf{z}, \mathbf{y}, \mathbf{K}, \sigma)$, on the other hand, is a generalization of the denoising case (where $\mathbf{K} = \mathbf{I}$). Eq. (10) is a Gaussian in $\mathbf{x}$ with precision (inverse covariance) matrix $\mathbf{Q}_{\mathbf{z}} = \mathbf{P}_{\mathbf{z}} + \frac{1}{\sigma^2}\mathbf{K}^{\mathrm{T}}\mathbf{K}$ and mean $\mathbf{Q}_{\mathbf{z}}^{-1}\mathbf{K}^{\mathrm{T}}\frac{\mathbf{y}}{\sigma^2}$, and can be sampled by solving two large sparse systems of linear equations [12]. The matrix $\mathbf{P}_{\mathbf{z}}$ is the same as in denoising and depends on the linear filters $\mathbf{J}_i$ and the current value of the auxiliary variables $\mathbf{z}$ [see 22].

Sampling from the posterior (Eq. (2)) thus proceeds using a Gibbs sampler that alternates between sampling from Eqs. (9) and (10) to obtain a sequence of samples $\{\{\mathbf{z}^{(1)}, \mathbf{x}^{(1)}\}, \ldots, \{\mathbf{z}^{(T)}, \mathbf{x}^{(T)}\}\}$. The MMSE estimate of $\mathbf{x}$ is approximated by averaging samples $\mathbf{x}^{(t)}$ from the posterior after B burn-in iterations ($B < t \leq T$); the samples of $\mathbf{z}$ are simply discarded at the end. Alternatively, we can use a Rao-Blackwellized MMSE estimator of $\mathbf{x}$ by averaging the conditional expectations from Eq. (10) [*cf*. 18]:

$$\hat{\mathbf{x}}_{\text{RB}} \approx \frac{1}{T-B} \sum\nolimits_{t=B+1}^{T} \mathbf{Q}_{\mathbf{z}^{(t)}}^{-1} \mathbf{K}^{\mathrm{T}} \frac{\mathbf{y}}{\sigma^2}. \tag{11}$$

We find that essentially the same performance can be achieved using either estimator, although many fewer iterations of the Gibbs sampler are necessary to satisfy our conservative convergence criteria (similar to [22]) when using Rao-Blackwellization. Hence, all results in the remainder of the paper were obtained using this approach.

5. Integrated Noise Estimation

Most MAP-based approaches in low-level vision rely on the choice of a regularization parameter λ that calibrates the influence of prior and likelihood on the posterior. This parameter is dependent on the noise level and must in practice be determined in an off-line training step. Nonetheless, even after training the regularization parameter, the MAP framework still requires the user to provide a noise level estimate, which can significantly affect the application performance when selected incorrectly. Fig. 1 shows how the image deblurring performance depends on the chosen noise level for a selection of deblurring methods, and illustrates that a reliable noise estimate is crucial for optimum deblurring performance. One of the properties of the MMSE-based deblurring approach proposed in Sec. 4 is that it does not require an off-line training of a regularization parameter. In contrast, all other approaches shown in Fig. 1 require such a procedure.

In the following, we further extend the framework of Sec. 4 by integrating noise estimation. Specifically, we adopt a Bayesian approach and treat σ as a nuisance parameter, which we (approximately) integrate out:

$$p(\mathbf{x}|\mathbf{y}, \mathbf{K}) = \int p(\mathbf{x}, \sigma|\mathbf{y}, \mathbf{K}) \, \mathrm{d}\sigma. \tag{12}$$

To that end we incorporate the noise σ as a new variable in an extended joint distribution $p(\mathbf{x}, \mathbf{z}, \sigma|\mathbf{y}, \mathbf{K})$. Since the input image and the blur kernel provide sufficient constraints in practice, we assume a uniform prior on σ, *i.e.* $p(\sigma) = \text{const}$. To estimate the integral from Eq. (12), we extend the Gibbs sampler from Sec. 4 by another step for sampling the conditional distribution $p(\sigma|\mathbf{x}, \mathbf{z}, \mathbf{y}, \mathbf{K})$, which is a Gamma distribution $\mathcal{G}(x; a, b) = \frac{x^{a-1} e^{-x/b}}{b^a \, \Gamma(a)}$ on the inverse noise variance [6]:

$$p(\sigma|\mathbf{x}, \mathbf{z}, \mathbf{y}, \mathbf{K}) \propto \mathcal{N}(\mathbf{y}; \mathbf{K}\mathbf{x}, \sigma^2 \mathbf{I}) \tag{13}$$

$$\propto \sigma^{-n} \exp\left(-\frac{\|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2}{2\sigma^2}\right) \tag{14}$$

$$\propto \mathcal{G}\left(\frac{1}{\sigma^2}; \frac{n}{2} + 1, \frac{2}{\|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2}\right). \tag{15}$$

Gibbs sampling proceeds by sampling σ, $\mathbf{z}$ and $\mathbf{x}$ alternatingly. Hence, we can obtain an MMSE estimate of the deblurred image $\mathbf{x}$ without knowledge of the noise level as

$$\hat{\mathbf{x}} = \arg\min_{\tilde{\mathbf{x}}} \iint \|\tilde{\mathbf{x}} - \mathbf{x}\|^2 p(\mathbf{x}, \sigma|\mathbf{y}, \mathbf{K}) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\sigma = E[\mathbf{x}|\mathbf{y}, \mathbf{K}]. \tag{16}$$

In practice, we employ Eq. (11) for this purpose, but replace σ with the current sample $\sigma^{(t)}$.

If we are not interested in the deblurred image, but rather in an MMSE estimate of the noise level σ itself, we can compute

$$\hat{\sigma} = \arg\min_{\tilde{\sigma}} \iint \|\tilde{\sigma} - \sigma\|^2 p(\mathbf{x}, \sigma|\mathbf{y}, \mathbf{K}) \, \mathrm{d}\mathbf{x} \, \mathrm{d}\sigma = E[\sigma|\mathbf{y}, \mathbf{K}] \tag{17}$$

by averaging the samples of σ. Alternatively, we can also apply Rao-Blackwellization here; based on the conditional expectations of Eq. (15), we obtain the Rao-Blackwellized MMSE estimate of σ as

$$\hat{\sigma}_{\text{RB}} \approx \frac{1}{T-B} \sum\nolimits_{t=B+1}^{T} \left(\frac{n+2}{\|\mathbf{y} - \mathbf{K}\mathbf{x}^{(t-1)}\|^2}\right)^{-1/2}. \tag{18}$$

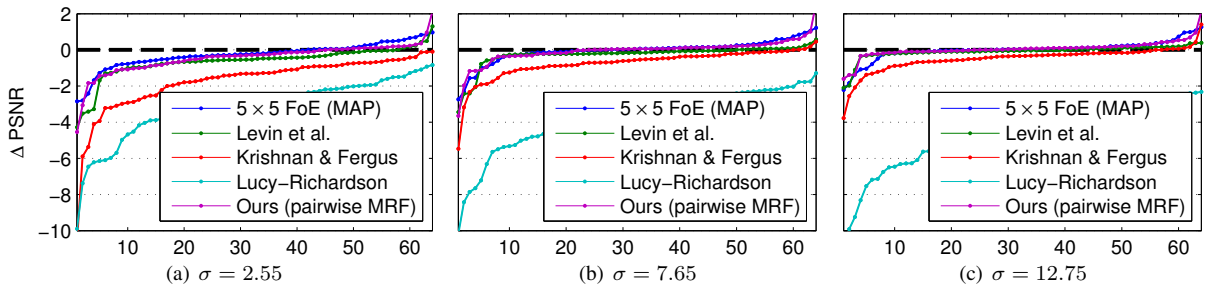
In contrast to the estimation framework of Fergus *et al.* [6], our sampling-based method allows to naturally incorporate a high-order image prior and model sensor noise in the spatial domain. Moreover, the fundamental difference to standard MAP approaches is that our noise estimation process is a fully automatic, built-in procedure of the deblurring algorithm, which arises naturally by treating the noise standard deviation as a variable of the posterior. This has the advantage that the noise estimation procedure is specialized to the image restoration problem at hand, here deblurring.

We would also like to point out the degenerate case of identity blur $\mathbf{K} = \mathbf{I}$, in which case samples drawn from $p(\mathbf{x}, \mathbf{z}, \sigma|\mathbf{y}, \mathbf{K})$ can effectively be used for blind denoising with simultaneous noise estimation. We evaluated this in Sec. 6 and found that despite unknown noise level, the results differ negligibly on average from the non-blind case.

6. Experiments

We evaluate our approach in the context of four different tasks: non-blind deblurring (with integrated and without noise estimation), blind image denoising, and noise estimation in denoising as well as in deblurring applications[1]. To

[1] MATLAB code is available on the authors' webpages.



(a) $\sigma = 2.55$ (b) $\sigma = 7.65$ (c) $\sigma = 12.75$

Figure 2. **Sorted PSNR differences** for the 64 deblurred images between our method (3×3 FoE) and all others. Noise estimates (using the method of Zoran and Weiss [30]) instead of the ground truth σ were used for the other approaches. Everything below the black dashed line means that our method (3×3 FoE) was better than the competing one. *Best viewed in color.*

	PSNR (dB)						SSIM					
	$\sigma = 2.55$		$\sigma = 7.65$		$\sigma = 12.75$		$\sigma = 2.55$		$\sigma = 7.65$		$\sigma = 12.75$	
	GT	NE	GT	NE	GT	NE	GT	NE	GT	NE	GT	NE
Lucy-Richardson [16, 19]	25.38	25.34	21.85	21.88	19.83	19.86	0.703	0.697	0.423	0.425	0.244	0.245
Krishnan & Fergus [11]	26.97	26.86	24.91	24.88	23.93	23.94	0.800	0.793	0.671	0.669	0.608	0.605
Levin *et al.* [13]	28.03	27.96	25.36	25.36	24.29	24.34	0.823	0.817	0.689	0.686	0.625	0.624
5×5 FoE (MAP) [20]	28.44	28.33	25.66	25.59	24.48	24.43	0.842	0.835	**0.711**	**0.708**	**0.646**	**0.642**
Ours (pairwise MRF)	28.24	28.17	25.63	25.58	**24.51**	**24.48**	0.833	0.830	0.700	0.696	0.633	0.629
Ours (3×3 FoE)	**28.66**	**28.61**	**25.68**	**25.64**	24.46	24.43	**0.850**	**0.846**	**0.711**	0.707	0.640	0.637

Table 2. **Average deblurring results for 64 test images.** *GT* denotes that the ground truth noise parameter σ was used, *NE* indicates that σ was assumed to be unknown and an estimate was used instead. For all methods except ours, the approach of Zoran and Weiss [30] was used for noise estimation prior to deblurring. We handle the unknown noise parameter as described in Sec. 5.

	Estimate $\hat{\sigma}$		PSNR	SSIM
	avg.	$\langle\epsilon\rangle$	avg. dB	avg.
5×5 FoE (MAP) [20]	GT	—	27.44	0.746
5×5 FoE (MAP) [21]	GT	—	27.86	0.776
Pairwise MRF (MMSE) [22]	GT	—	27.54	0.758
3×3 FoE (MMSE) [22]	GT	—	**27.95**	**0.788**
Zoran & Weiss [30]	23.16	8.8%	—	—
Ours (pairwise MRF)	22.81	10.1%	27.16	0.733
Ours (3×3 FoE)	24.20	**5.8%**	27.88	0.783

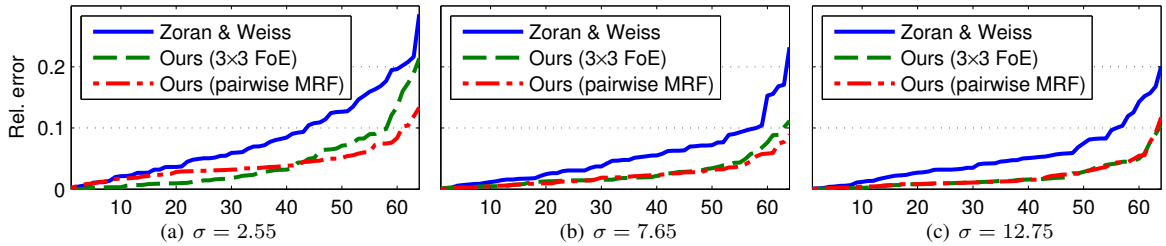
Table 1. **Average denoising results and noise estimates** $\hat{\sigma}$ for 68 test images and $\sigma = 25$ (partly reproduced from [22]). The average relative estimation error is denoted by $\langle\epsilon\rangle = \sum_{m=1}^{M} |\hat{\sigma}_m - \sigma| / (M\sigma)$; GT denotes that the ground truth value for σ was used.

separate the contribution of our general approach from the effects of the learned, high-order prior, we also report results for a learned pairwise MRF. If unspecified, the experimental discussion below refers to using the 3×3 FoE. To facilitate comparisons, we use the parameters of the learned pairwise MRF and 3×3 FoE models from [22].

Blind denoising with noise estimation. We begin by highlighting the special case of no blur, *i.e.* the blur matrix $\mathbf{K}$ equals the identity matrix – this corresponds to image denoising. This case is interesting, since our method enables us to perform blind denoising (unknown σ) with integrated noise estimation. Based on the approach described in Secs. 4 and 5 (with $\mathbf{K} = \mathbf{I}$), we performed a series of experiments on 68 images used in [20]. The results are summarized in Tab. 1 using the common peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [26]. Most importantly, we find that despite unknown noise level our average results are only slightly worse than the non-blind denoising results with a 3×3 FoE reported in [22] (27.88dB vs. 27.95dB). Nonetheless, the performance on individual images can significantly differ. Moreover, if we use our approach to perform MMSE-based noise estimation, we obtain results that are superior to those of Zoran and Weiss [30]. This is an interesting result due to the conceptual simplicity of our approach: It is solely guided by a noise model and a natural image prior.

The performance in case of a pairwise MRF drops behind the non-blind setting more significantly (~ 0.4dB worse). Moreover, noise estimates are also slightly inferior to [30] in this case. It is interesting to note that we observe neither effect in case of deblurring (see below). Finally, Tab. 1 also shows that our MMSE-based approach with integrated noise estimation outperforms the MAP-based approaches with 5×5 FoEs of [20] and [21] despite the fact that they rely on knowledge of the noise parameter σ.

Non-blind deblurring. A general problem with evaluating non-blind deblurring algorithms is the scarcity of ground-truth data or even realistic blur kernels. We thus use the eight publicly available blur kernels from [14] on eight different images of size 128×128 pixels each, to synthetically blur 64 images overall; the images were randomly cropped from the same test set as used for blind denoising. After blurring the images, we added white Gaussian noise of



(a) $\sigma = 2.55$ (b) $\sigma = 7.65$ (c) $\sigma = 12.75$

Figure 3. **Relative estimation errors** $|\hat{\sigma} - \sigma|/\sigma$ for the 64 blurred test images and three noise levels, comparing our method with Zoran and Weiss [30]. Each curve is sorted *separately* and does not indicate a performance comparison between the methods for a given image.

	$\sigma = 2.55$			$\sigma = 7.65$			$\sigma = 12.75$		
	avg.	std. dev.	$\langle\epsilon\rangle$	avg.	std. dev.	$\langle\epsilon\rangle$	avg.	std. dev.	$\langle\epsilon\rangle$
Zoran & Weiss [30]	2.52	0.27	8.21%	7.44	0.50	5.34%	12.41	0.80	5.18%
Ours (pairwise MRF)	2.55	0.12	**4.02%**	7.68	0.23	**2.26%**	12.82	0.38	2.01%
Ours (3×3 FoE)	2.64	0.14	4.28%	7.79	0.23	2.56%	12.87	0.35	**1.95%**

Table 3. **Noise estimation results** for the 64 blurred test images and three noise levels, comparing our method with Zoran and Weiss [30]. The average relative estimation error is denoted by $\langle\epsilon\rangle = \sum_{m=1}^{M} |\hat{\sigma}_m - \sigma|/(M\sigma)$.

varying strength to obtain three test sets for noise levels 1%, 3%, and 5% ($\sigma = 2.55, 7.65, 12.75$). To simulate realistic conditions, all pixel values were subsequently rounded to one of 256 discrete intensities. Additionally, we deblurred with a slightly perturbed version of the true blur kernel by adding white Gaussian noise with variance 10^{-6} to it. The motivation is to mimic a more realistic scenario where the estimated blur kernel always contains some error.

We compare our method against the recent MAP-based approaches of Levin *et al*. [13] and Krishnan and Fergus [11], as well as the classical Lucy-Richardson algorithm [16, 19]. Both [11, 13] use image priors with hyper-Laplacian potential functions $e^{-\rho(x)}$ with $\rho(x) = |x|^{\alpha}$. We used the original implementation of [13] unchanged with $\alpha = 4/5$, and chose $\alpha = 2/3$ for the original implementation of [11] to achieve best performance. In addition, we also applied the 5×5 FoE prior from [20] to the non-blind deblurring case, and used an EM algorithm [see 12] to perform MAP estimation.

For all competing methods, the regularization parameters were determined (per noise level) on a separate training set of 16 images, using each of the eight blur kernels from [14] twice. In order to better compare the results with our method, which does not rely on knowing the noise level, we also obtained noise estimates with the approach of [30] and used those estimates instead of the ground truth noise value for the competing methods. For completeness, we also report results for our method with known noise level, which show that integrated noise estimation performs almost identically to deblurring with the (usually unknown) ground truth noise level.

The quantitative deblurring results for the 64 images and three noise levels are summarized in Tab. 2. Fig. 2 visually illustrates the performance difference of our approach to various baselines. The results show that Lucy-Richardson [16, 19] and the very fast method of Krishnan and Fergus [11] are far behind the other methods for all noise levels, particularly behind ours (at least 0.5dB worse on average). We find that our method particularly outperforms competing approaches for the smallest – and for many applications likely the most realistic – noise level. For example, the widely used method of Levin *et al*. [13] is outperformed by about 0.6dB. An interesting observation is also that deblurring with the 5×5 FoE prior from [20] outperforms [13], which as far as we are aware has not been observed before. This shows that the learned, high-order MRF prior is at least partly responsible for the observed performance difference to [13]. Nonetheless, our Bayesian approach with integrated noise estimation still performs about 0.25dB better than using a 5×5 FoE with MAP estimation, despite using smaller 3×3 cliques. Interestingly, even when combined with a simple pairwise MRF our method comes close in performance to the 5×5 FoE. This demonstrates that the proposed MMSE approach substantially improves performance, while offering an integrated noise estimation procedure and not requiring a regularization parameter.

A qualitative comparison on two larger images that were not part of the test set can be seen in Figs. 4 and 5. We observe in both images that our method is especially good at preserving textural detail, while at the same time allowing for very smooth regions, which visually appear too crisp or show artifacts in case of the competing methods (*e.g*. the sky in Fig. 4 or the background near the top of Fig. 5).

Noise estimation for deblurring. We finally evaluated the noise estimation performance itself by comparing with the approach of Zoran and Weiss [30], which is one of the most competitive techniques in this area. We report results for 64 blurred images and three noise levels in Tab. 3; a visual comparison is given in Fig. 3. We find that our estimates are significantly better than those of [30] in terms of the

average relative estimation error. This holds for the pairwise MRF and high-order FoE model as well as all noise levels, particularly the large ones, and demonstrates the advantage of having a noise estimation procedure that is specifically adapted to the problem of image deblurring.

Computational considerations. Unsurprisingly, the Bayesian approach taken here is computationally more expensive than MAP estimation (energy minimization). The reported results are intended to show the performance limits of our method and consequently use conservative thresholds for assessing sampler convergence. For practical purposes it is easily possible to relax the convergence criteria with little to no influence on the deblurring performance. For the 64 test images in Tab. 2 and $\sigma = 2.55/7.65/12.75$, a simple MATLAB implementation achieves an average runtime of $4.6/3.9/3.9$ minutes, while staying within 0.03dB of the quoted results. A similar MAP implementation used for the 5×5 FoE [20] is only about 4 times faster. The methods of [11] and [13] are optimized for speed and need at most a few seconds per image. While much faster, their deblurring performance is significantly worse (1.7dB and 0.6dB). Furthermore, they require additional effort for noise estimation and determining suitable regularization parameters.

7. Conclusions and Future Work

Based on posterior sampling, we presented an integrated Bayesian framework for unified non-blind deblurring and noise estimation. By relying on MMSE estimation and marginalizing out the noise parameter, we not only free the user from tedious parameter tuning, but also achieve improved application performance by exploiting learned, high-order priors. In an extensive experimental evaluation, our framework was demonstrated to outperform state-of-the-art MAP deblurring and noise estimation algorithms using standard quantitative measures.

To cope with the computational cost of Bayesian inference, future work should consider more efficient sampling algorithms. Moreover, since our integrated noise estimation approach in principle extends to other problems, such as super-resolution, future work may consider exploring further applications of the proposed framework. In general, it may be possible to estimate additional parameters of the problem at hand; for deblurring in particular, estimating the blur kernel in an integrated fashion would extend our framework to the problem of blind deblurring (assuming spatially uniform blur). Finally, future work should also be devoted to gathering ground-truth data for spatially varying blur and evaluating our and other approaches on this task.

Acknowledgements. We thank George Papandreou for suggesting Rao-Blackwellized MMSE estimation.

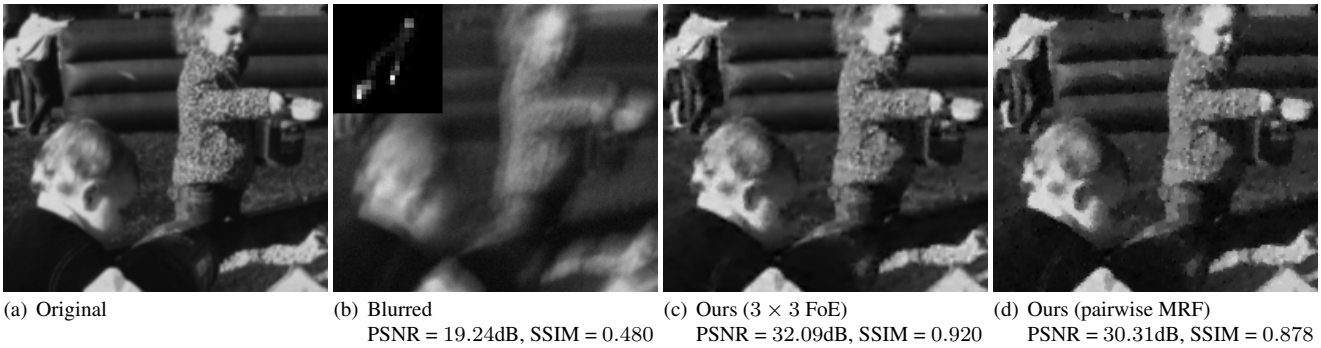
References

[1] M. Ben-Ezra and S. Nayar. Motion-based motion deblurring. *IEEE TPAMI*, 26(6):689–698, Apr. 2004.
[2] S. Cho and S. Lee. Fast motion deblurring. *ACM T. Graphics*, 28, 2009.
[3] T. Cho, N. Joshi, C. Zitnick, S. Kang, R. Szeliski, and W. Freeman. A content-aware image prior. In *CVPR 2010*.
[4] A. De Stefano, P. White, and W. Collis. Training methods for image noise level estimation on wavelet components. *EURASIP J. App. Signal Process.*, 2004(1):2400–2407, Jan. 2004.
[5] D. Donoho and J. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425, 1994.
[6] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM T. Graphics*, 3(25), July 2006.
[7] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE TIP*, 4(7):932–946, July 1995.
[8] N. Joshi, S. Kang, C. Zitnick, and R. Szeliski. Image deblurring using inertial measurement sensors. *ACM T. Graphics*, 29(4), July 2010.
[9] N. Joshi, R. Szeliski, and D. Kriegman. PSF estimation using sharp edge prediction. In *CVPR 2008*.
[10] K. Krajsek and R. Mester. A maximum likelihood estimator for choosing the regularization parameters in global optical flow methods. In *ICIP 2006*.
[11] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-Laplacian priors. In *NIPS*2009*.
[12] E. Levi. Using natural image priors – Maximizing or sampling? Master's thesis, The Hebrew University of Jerusalem, 2009.
[13] A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM T. Graphics*, 26(3), July 2007.
[14] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR 2009*.
[15] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. Noise estimation from a single image. In *CVPR 2006*, volume 1, pages 901–908.
[16] L. B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745, 1974.
[17] J. Miskin and D. J. C. MacKay. Ensemble learning for blind image separation and deconvolution. *Adv. in Ind. Comp. Analysis*, 2000.
[18] G. Papandreou and A. Yuille. Gaussian sampling by local perturbations. In *NIPS*2010*.
[19] W. Richardson. Bayesian-based iterative method of image restoration. *J. Opt. Soc. America*, 62(1):55–59, 1972.
[20] S. Roth and M. J. Black. Fields of experts. *IJCV*, 82(2):205–229, Apr. 2009.
[21] K. G. G. Samuel and M. F. Tappen. Learning optimized MAP estimates in continuously-valued MRF models. In *CVPR 2009*.
[22] U. Schmidt, Q. Gao, and S. Roth. A generative perspective on MRFs in low-level vision. In *CVPR 2010*.
[23] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM T. Graphics*, 27(3):73–73, Aug. 2008.
[24] Y. Tai, H. Du, M. Brown, and S. Lin. Image/video deblurring using a hybrid camera. In *CVPR 2008*.
[25] M. J. Wainwright and E. P. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. In *NIPS*1999*, pages 855–861.
[26] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, Apr. 2004.
[27] L. Yuan, J. Sun, L. Quan, and H. Shum. Image deblurring with blurred/noisy image pairs. *ACM T. Graphics*, 26(3), July 2007.
[28] L. Zhang and S. M. Seitz. Estimating optimal parameters for MRF stereo from a single image pair. *IEEE TPAMI*, 29(2):331–342, Feb. 2007.
[29] V. Zlokolica, A. Pizurica, and W. Philips. Noise estimation for video processing based on spatio-temporal gradients. *IEEE Signal Process. Letters*, 13(6):337–340, June 2006.
[30] D. Zoran and Y. Weiss. Scale invariance and noise in natural images. In *ICCV 2009*.

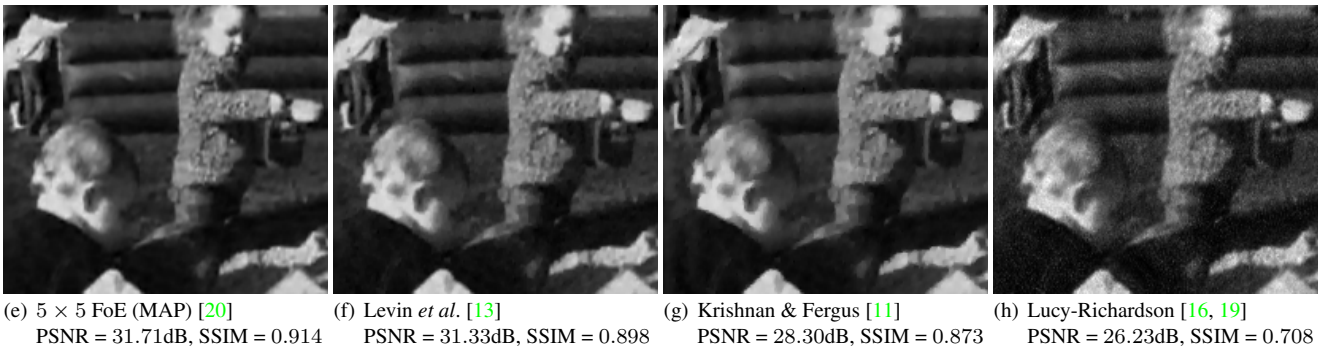
(a) Original
(b) Blurred
PSNR = 23.89dB, SSIM = 0.588
(c) Ours (3×3 FoE)
PSNR = 29.05dB, SSIM = 0.864
(d) Ours (pairwise MRF)
PSNR = 28.84dB, SSIM = 0.842

(e) 5×5 FoE (MAP) [20]
PSNR = 28.81dB, SSIM = 0.844
(f) Levin *et al*. [13]
PSNR = 28.54dB, SSIM = 0.826
(g) Krishnan & Fergus [11]
PSNR = 28.36dB, SSIM = 0.825
(h) Lucy-Richardson [16, 19]
PSNR = 27.01dB, SSIM = 0.693

Figure 4. **Deblurring example (cropped).** Visual comparison of non-blind methods, where the ground truth noise level was used for all approaches except ours. Our method simultaneously preserves rock texture and smooth sky. The 15×15 blur kernel shown in *(b)* is spatially resized and its entries are scaled for better visualization. *Best viewed on screen.*



(a) Original
(b) Blurred
PSNR = 19.24dB, SSIM = 0.480
(c) Ours (3×3 FoE)
PSNR = 32.09dB, SSIM = 0.920
(d) Ours (pairwise MRF)
PSNR = 30.31dB, SSIM = 0.878



(e) 5×5 FoE (MAP) [20]
PSNR = 31.71dB, SSIM = 0.914
(f) Levin *et al*. [13]
PSNR = 31.33dB, SSIM = 0.898
(g) Krishnan & Fergus [11]
PSNR = 28.30dB, SSIM = 0.873
(h) Lucy-Richardson [16, 19]
PSNR = 26.23dB, SSIM = 0.708

Figure 5. **Deblurring example (cropped).** Visual comparison of non-blind methods, where the ground truth noise level was used used for all approaches except ours. The blur exhibits sparseness typical to camera shake. Our method simultaneously preserves clothing texture and smooth background regions. The 23×23 blur kernel shown in *(b)* is spatially resized and its entries are scaled for better visualization. *Best viewed on screen.*