EURASIP Journal on Image and Video Processing
a SpringerOpen Journal

RESEARCH

Open Access

Sparse Bayesian blind image deconvolution with parameter estimation

Bruno Amizic[1*], Rafael Molina[2] and Aggelos K Katsaggelos[1]

Abstract

In this article, we propose a novel blind image deconvolution method developed within the Bayesian framework. We concentrate on the restoration of blurred photographs taken by commercial cameras to show its effectiveness. The proposed method is based on a non-convex l_p quasi norm with $0 < p < 1$ that is used for the image, and a total variation (TV) based prior that is utilized for the blur. Bayesian inference is carried out by utilizing bounds for both the image and blur priors using a majorization-minimization principle. Maximum *a posteriori* estimates of the unknown image, blur and model parameters are calculated. Experimental results (i.e., restorations of more than 30 blurred photographs) are presented to demonstrate the advantage of the proposed method compared to existing ones.

1 Introduction

Blind image deconvolution (BID) refers to the process of estimating both the original image and the blur from the degraded noisy image observation by using partial information about the imaging system. Blind image deconvolution algorithms represent a valuable tool that can be used for improving image quality without requiring complicated calibrations of the real-time image acquisition and processing system (i.e., medical imaging, videoconferencing, space exploration, x-ray imaging, etc.).

The blind image deconvolution problem is encountered in many different technical areas, such as astronomical imaging, remote sensing, microscopy, medical imaging, optics, super-resolution applications, and motion tracking applications among others (see, for example, [1-9]). Astronomical imaging is one of the primary applications of blind image deconvolution algorithms [1,2]. Ground based imaging systems are subject to blurring due to the rapidly changing index of refractions of the atmosphere. Extraterrestrial observations of the Earth and the planets are degraded by motion blur as a result of slow camera shutter speeds relative to the rapid spacecraft motion. Blind image deconvolution is used for improving the quality of the Poisson distributed film grain noise present in the blurred X-rays, mammograms, and digital angiographic images. In such applications, many times, the blurring is unavoidable because the medical imaging systems limit the intensity (e.g., low X-ray intensity) of the incident radiation in order to protect the patient's health [10]. In optics, blind image deconvolution is used to restore the original image from the degradation introduced by a microscope or any other optical instrument [6,7]. The Hubble Space Telescope main mirror imperfections have provided an inordinate amount of images for the digital image processing community [1]. As a final example, in tracking applications the object being tracked might be blurred due to its speed or the motion of the camera. As a result, the track is lost with conventional tracking approaches and the application of blind restoration approaches can improve tracking results [9].

The standard formulation of the gray-scale image degradation model is given in matrix-vector form by

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

where the $N \times 1$ vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{n}$ represent, respectively, the original image, the available noisy and blurred image, and the observation noise, and $\mathbf{H}$ represents the blurring matrix created from the blur point spread function $\mathbf{h}$. The images are assumed to be of size $m \times n = N$, and they are lexicographically ordered into $N \times 1$ vectors. Given $\mathbf{y}$, the BID problem calls for finding estimates of $\mathbf{x}$ and $\mathbf{H}$ using prior knowledge on them.

A number of methods have been proposed to address BID (a recent literature review can be found in [11]). The most recent methods are based on a Bayesian framework,

*Correspondence: amizic@u.northwestern.edu
[1] Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, 60208-3118, USA
Full list of author information is available at the end of the article

and have addressed the removal of the camera motion from blurred photographs [12-19]. In [12] the unknown image and blur were estimated in a two step process. In the first step the blur is estimated from the blurred photograph by regularizing the image gradients with a mixture of Gaussian distributions and by regularizing the blur with a mixture of exponential distributions. In the second step the image is estimated from the blurred photograph and the estimated blur by utilizing the Richardson-Lucy (RL) algorithm. Finally, the restored image is obtained by performing the histogram equalization (to that of the observed image) on the output of RL. A multi-scale approach is utilized for the algorithm implementation. The multi-scale approach consists of down-sampling the blurred photograph to number of low resolution images, and by utilizing the proposed algorithm iteratively to obtain the blur estimate at each resolution.

Estimating camera motion has also been investigated in [13,14], where the unknown image and blur were estimated in a simultaneous fashion. Additionally, [14] concentrated on synthetic experiments where the performance of the algorithm was evaluated by the improvement in signal to noise ratio. Also, in [13,14] the regularization parameters are not automatically estimated but rather heuristically chosen by the user at each iteration in order to yield an unknown image estimate with good visual quality. The major disadvantage of the methods proposed in [13,14] compared to the method proposed in [12] is the lack of parameter estimation.

In this article, we extend our study in [20] by providing (1) a multi-scale based implementation of the algorithm which improves the quality of the obtained restored images, and (2) a complete comparison with many existing state of the art blind deconvolution methods. The proposed Bayesian algorithm for BID utilizes a variant of the non-convex l_p quasi norm based prior as the unknown image prior and the TV prior as the unknown blur prior. Furthermore, we utilize the Bayesian framework to provide the estimates for all model parameters. Finally, we evaluate the performance of the proposed algorithm and provide comparisons with [12-14,16,19] by restoring blurred photographs taken by commercial cameras.

This article is organized as follows. In Section 2 we provide the proposed Bayesian modeling of the BID problem. The Bayesian inference is presented in Section 3. In Section 4, we describe implementation details of the proposed algorithm. Experimental results are provided in Section 5 and conclusions are drawn in Section 6.

2 Bayesian modeling
As already discussed in the previous section, the observation noise is modeled as a zero mean white Gaussian ran-

dom vector. Therefore, the observation model is defined as

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

where β is the precision of the multivariate Gaussian distribution.

As the image prior we utilize a variant of the generalized Gaussian distribution, given by

$$p(\mathbf{x}|\boldsymbol{\alpha}) = \frac{1}{Z_{GG}(\boldsymbol{\alpha})} \exp\left[-\sum_{d \in D} \alpha_d \sum_i |\Delta_i^d(\mathbf{x})|^p\right], \quad (3)$$

where $Z_{GG}(\boldsymbol{\alpha})$ is the partition function, $0 < p < 1$, $\boldsymbol{\alpha}$ denotes the set $\{\alpha_d\}$ and $d \in D = \{h, v, hh, vv, hv\}$. $\Delta_i^h(\mathbf{x})$ and $\Delta_i^v(\mathbf{x})$ correspond to, respectively, the horizontal and vertical first order differences, at pixel i, that is, $\Delta_i^h(\mathbf{x}) = x_i - x_{l(i)}$ and $\Delta_i^v(\mathbf{x}) = x_i - x_{a(i)}$, where $l(i)$ and $a(i)$ denote the nearest neighbors of i, to the left and above, respectively. The operators $\Delta_i^{hh}(\mathbf{x})$, $\Delta_i^{vv}(\mathbf{x})$, $\Delta_i^{hv}(\mathbf{x})$ correspond to, respectively, horizontal, vertical and horizontal-vertical second order differences, at pixel i.

In this study, similarly to [15,21], we utilize a non-convex l_p quasi norm with $0 < p < 1$ since the derivatives of blurry photographs are expected to be sparse. The distributions of the image derivatives often have heavier tails that are better modeled with the non-convex l_p quasi norm prior with $0 < p < 1$ compared to the convex priors modeled with $p = 1, 2$.

For reducing the complexity of the problem we assume that $\alpha_h = \alpha_v = \alpha$ and $\alpha_{hh} = \alpha_{vv} = \alpha_{hv} = \alpha/2$. Additionally, similarly to [22], the partition function is approximated as $Z_{GG}(\alpha) \propto \alpha^{-\lambda_1 N/p}$, where λ_1 is a positive real number. We then simplify (3) accordingly to obtain the following image prior

$$p(\mathbf{x}|\boldsymbol{\alpha}) \propto \alpha^{\lambda_1 N/p} \exp\left[-\alpha \sum_{d \in D} 2^{1-o(d)} \sum_{i=1}^{N} |\Delta_i^d(\mathbf{x})|^p\right], \quad (4)$$

where $o(d) \in \{1, 2\}$ denotes the order of the difference operator $\Delta_i^d(\mathbf{x})$.

For the blur we utilize the total-variation prior given by (see [23] for more details)

$$p(\mathbf{h}|\gamma) \propto \gamma^{\lambda_2 N} \exp\left[-\gamma \text{TV}(\mathbf{h})\right], \quad (5)$$

where λ_2 is a positive real number and $\text{TV}(\mathbf{h})$ is defined as

$$\text{TV}(\mathbf{h}) = \sum_i \sqrt{(\Delta_i^h(\mathbf{h}))^2 + (\Delta_i^v(\mathbf{h}))^2}. \quad (6)$$

In this study, we use flat improper hyperpriors on α, β and γ, that is, we utilize

$$p(\alpha) \propto \text{const}, \qquad p(\beta) \propto \text{const}, \qquad p(\gamma) \propto \text{const}. \quad (7)$$

Note that with this choice of the hyperpriors, the observed image $\mathbf{y}$ is made solely responsible for the estimation of the image, blur and hyperparameters.

3 Bayesian inference

Bayesian inference on the unknown components of the blind image deconvolution problem is based on the estimation of the unknown posterior distribution $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y})$, given by

$$p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y}) = \frac{p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h}, \mathbf{y})}{p(\mathbf{y})}. \tag{8}$$

Assuming that $\mathbf{x}$ and $\mathbf{h}$ are independent, the joint distribution $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h}, \mathbf{y})$ can be factorized in terms of the observation model $p(\mathbf{y}|\beta, \mathbf{x}, \mathbf{h})$, the prior distributions $p(\mathbf{x}|\alpha)$ and $p(\mathbf{h}|\gamma)$, and the hyperparameter distributions $p(\alpha)$, $p(\beta)$ and $p(\gamma)$, that is,

$$p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h}, \mathbf{y}) = p(\mathbf{y}|\beta, \mathbf{x}, \mathbf{h})p(\mathbf{x}|\alpha)p(\mathbf{h}|\gamma)p(\alpha)p(\beta)p(\gamma). \tag{9}$$

In this study, we adopt the maximum *a posteriori* (MAP) approach to obtain a single point estimate, $\bar{\Theta} = (\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\mathbf{x}}, \bar{\mathbf{h}})$, that maximizes $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y})$ as follows,

$$
\begin{aligned}
\bar{\Theta} &= argmax_\Theta p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y}) \\
&= min_\Theta \left\{ \frac{\beta}{2} \|\mathbf{y} - \mathbf{Hx}\|^2 + \alpha \sum_{d \in D} 2^{1-o(d)} \sum_i |\Delta_i^d(\mathbf{x})|^p \right. \\
&\quad \left. + \gamma \operatorname{TV}(\mathbf{h}) - \frac{\lambda_1 N}{p} \log \alpha - \frac{N}{2} \log \beta - \lambda_2 N \log \gamma \right\}.
\end{aligned}
\tag{10}
$$

As can be seen from (10), obtaining the point estimate that maximizes the posterior distribution $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y})$ is not straightforward since it requires the minimization of a non-convex functional. Maximizing the posterior distribution $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y})$ by iteratively optimizing in one variable while fixing the others (the so called iterated conditional modes (ICM) method [24]) is equivalent to the variational Bayesian based maximization (see [25] for an example derivation) for the special case when all the posterior distributions are assumed to be degenerate.

In this article, we apply the majorization-minimization approach twice to bound the non-convex functional to be minimized. We start by bounding the non-convex image prior $p(\mathbf{x}|\alpha)$ by the functional $M_1(\alpha, \mathbf{x}, \mathbf{Z})$, that is

$$p(\mathbf{x}|\alpha) \geq \text{const} \cdot M_1(\alpha, \mathbf{x}, \mathbf{Z}). \tag{11}$$

The majorization-minimization approach has been utilized in several approaches for image restoration [25,26].

The functional $M_1(\alpha, \mathbf{x}, \mathbf{Z})$ is derived by considering the relationship between the weighted geometric and arithmetic means, which is given by

$$t^{p/2} z^{1-p/2} \leq \frac{p}{2} t + \left(1 - \frac{p}{2}\right) z, \tag{12}$$

where $t \geq 0, z > 0$, and $0 < p < 2$. We first rewrite (12) as

$$t^{p/2} \leq \frac{p}{2} \frac{t + \frac{2-p}{p} z}{z^{1-p/2}}. \tag{13}$$

Using (13) we obtain

$$|\Delta_i^d(\mathbf{x})|^p \leq \frac{p}{2} \frac{[\Delta_i^d(\mathbf{x})]^2 + \frac{2-p}{p} z_{d,i}}{z_{d,i}^{1-p/2}}. \tag{14}$$

Therefore, we have

$$
\begin{aligned}
p(\mathbf{x}|\alpha) &= \text{const} \cdot \alpha^{\lambda_1 N/p} \exp\left[-\alpha \sum_{d \in D} 2^{1-o(d)} \sum_i |\Delta_i^d(\mathbf{x})|^p\right] \\
&\geq \text{const} \cdot \alpha^{\lambda_1 N/p} \exp\left[-\frac{\alpha p}{2} \sum_{d \in D} 2^{1-o(d)} \right. \\
&\quad \left. \sum_i \frac{[\Delta_i^d(\mathbf{x})]^2 + \frac{2-p}{p} z_{d,i}}{z_{d,i}^{1-p/2}}\right].
\end{aligned}
\tag{15}
$$

Then (11) holds by setting

$$
\begin{aligned}
M_1(\alpha, \mathbf{x}, \mathbf{Z}) &= \alpha^{\lambda_1 N/p} \exp\left[-\frac{\alpha p}{2} \sum_{d \in D} 2^{1-o(d)} \right. \\
&\quad \left. \sum_i \frac{[\Delta_i^d(\mathbf{x})]^2 + \frac{2-p}{p} z_{d,i}}{z_{d,i}^{1-p/2}}\right],
\end{aligned}
\tag{16}
$$

where $\mathbf{Z}$ is a matrix with elements $z_{d,i}$, with $d \in \{h, v, hh, vv, hv\}$ and $i = 1, \ldots, N$.

Similarly, the majorization-minimization criterion is used to bound the blur prior $p(\mathbf{h}|\gamma)$ utilizing the functional $M_2(\gamma, \mathbf{h}, \mathbf{u})$. Let us define, for γ and any N-dimensional vector $\mathbf{u} \in (R^+)^N$, with components u_i, $i = 1, \ldots, N$, the following functional

$$M_2(\gamma, \mathbf{h}, \mathbf{u}) = \alpha^{\lambda_2 N} \exp\left[-\frac{\gamma}{2} \sum_i \frac{(\Delta_i^h(\mathbf{h}))^2 + (\Delta_i^v(\mathbf{h}))^2 + u_i}{\sqrt{u_i}}\right]. \tag{17}$$

Using the inequality in (13) with $p = 1$, for $t \geq 0$ and $z > 0$, that is,

$$\sqrt{t} \leq \sqrt{z} + \frac{1}{2\sqrt{z}}(t - z), \tag{18}$$

we obtain

$$p(\mathbf{h}|\gamma) \geq \text{const} \cdot M_2(\gamma, \mathbf{h}, \mathbf{u}). \tag{19}$$

The lower bounds of $p(\mathbf{x}|\alpha)$ and $p(\mathbf{h}|\gamma)$ defined above lead to the following lower bound of the distribution $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h}, \mathbf{y})$
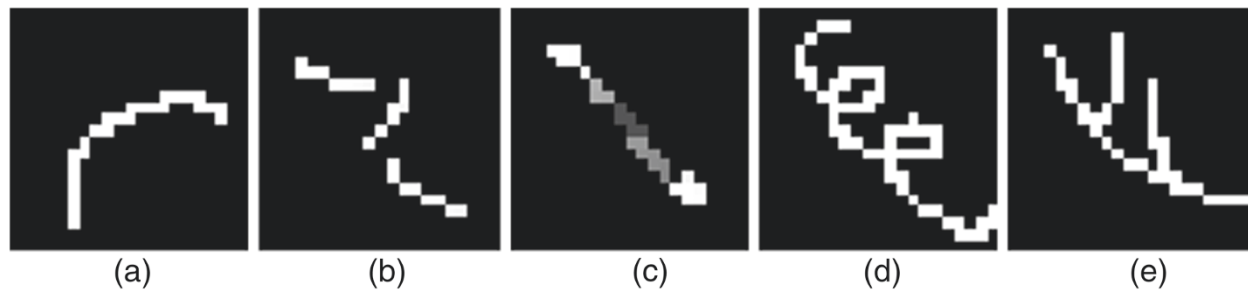
Figure 1 Five different synthetic non-parametric motion blurs included in the window of size 21×21: (a) Blur 1: the support is 14×12, **(b)** Blur 2: the support is 15×14, **(c)** Blur 3: the support is 14×14 **(d)** Blur 4: the support is 18×19, **(e)** Blur 5: the support is 18×15.
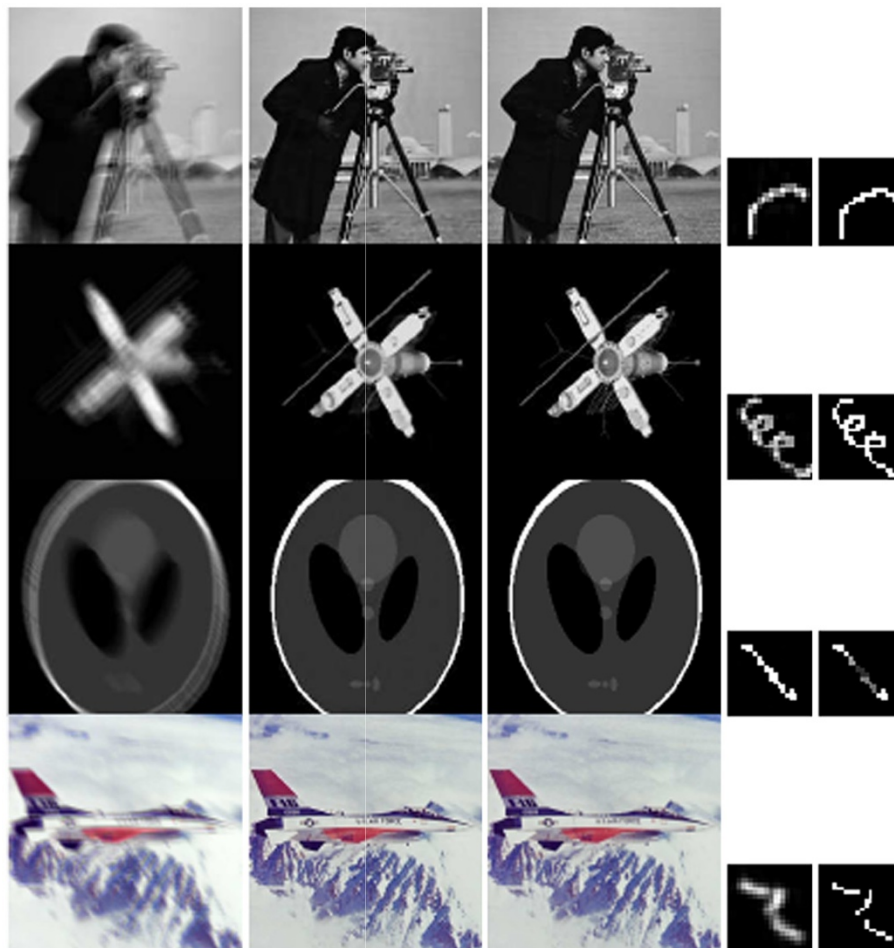

Figure 2 Example restorations from Table 1: 1st column represents four different blurred observations, 2nd column represents their respective restorations obtained by the proposed algorithm, 3rd column represents their respective original images, 4th column represents their respective blurs obtained by the proposed algorithm, 5th column represents their respective original blurs.

Table 1 ISNR$_{\hat{x}}$ and ISNR$_{\hat{h}}$ values, for the cameraman, satellite, shepp-logan, and airplane images degraded by five different motion blurs (BSNR = 40 dB)

Image	Blur 1		Blur 2		Blur 3		Blur 4		Blur 5	
	ISNR$_{\hat{x}}$	ISNR$_{\hat{h}}$	ISNR$_{\hat{x}}$	ISNR$_{\hat{h}}$	ISNR$_{\hat{x}}$	ISNR$_{\hat{h}}$	ISNR$_{\hat{x}}$	ISNR$_{\hat{h}}$	ISNR$_{\hat{x}}$	ISNR$_{\hat{h}}$
cameraman	11.41	20.80	7.07	16.74	6.64	18.41	5.96	21.94	6.92	21.22
satellite	18.50	39.44	17.85	35.49	17.98	20.56	8.76	25.76	9.02	26.17
shepp-logan	32.49	47.34	24.27	45.23	31.16	20.51	22.99	49.00	22.36	49.76
airplane-color	8.10	17.55	9.34	15.67	7.99	15.78	8.50	20.84	4.83	18.32

$$p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h}, \mathbf{y}) = p(\alpha)p(\beta)p(\gamma)p(\mathbf{x}|\alpha)p(\mathbf{h}|\gamma)p(\mathbf{y}|\beta, \mathbf{x}, \mathbf{h})$$
$$\geq \text{const} \cdot p(\alpha)p(\beta)p(\gamma)M_1(\alpha, \mathbf{x}, \mathbf{V})M_2$$
$$(\gamma, \mathbf{h}, \mathbf{u})p(\mathbf{y}|\beta, \mathbf{x}, \mathbf{h}).$$

Therefore, a single point estimate that maximizes the lower bound of the posterior distribution $p(\alpha, \beta, \gamma, \mathbf{x}, \mathbf{h} \mid \mathbf{y})$ is found as follows

$$\bar{\Theta} = min_\Theta \left\{ \frac{\beta}{2} \|\mathbf{y} - \mathbf{Hx}\|^2 + \frac{\alpha p}{2} \sum_{d \in D} 2^{1-o(d)} \right.$$
$$\sum_i \frac{[\Delta_i^d(\mathbf{x})]^2 + \frac{2-p}{p}z_{d,i}}{z_{d,i}^{1-p/2}} +$$
$$+ \frac{\gamma}{2} \sum_i \frac{(\Delta_i^h(\mathbf{h}))^2 + (\Delta_i^v(\mathbf{h}))^2 + u_i}{\sqrt{u_i}}$$
$$\left. - \frac{\lambda_1 N}{p} \log \alpha - \frac{N}{2} \log \beta - \lambda_2 N \log \gamma \right\}. \tag{20}$$

As shown in (20), we are effectively replacing the original non-convex minimization problem (10) by a series of convex ones by utilizing the majorization-minimization criteria and introducing the additional variational vectors $\mathbf{z}_d$ and $\mathbf{u}$. By iteratively solving this convex optimization problem in an alternating fashion with respect to all unknowns, we obtain a sequence of point estimates and derive the proposed algorithm as shown next.

3.1 Algorithm
Given $\alpha^1, \beta^1, \gamma^1, \mathbf{h}^1, u_i^1 = [\Delta_i^h(\mathbf{h}^1)]^2 + [\Delta_i^v(\mathbf{h}^1)]^2$, and $z_{d,i}^1$. for $k = 1, 2, \ldots$ until a stopping criterion is met:

1. Calculate
$$\mathbf{x}^k = \left[\beta^k(\mathbf{H}^k)^t(\mathbf{H}^k) + \alpha^k p \right.$$
$$\left. \sum_d 2^{1-o(d)}(\mathbf{\Delta}^d)^t \mathbf{W}_d^k(\mathbf{\Delta}^d) \right]^{-1} \beta^k(\mathbf{H}^k)^t\mathbf{y}, \tag{21}$$
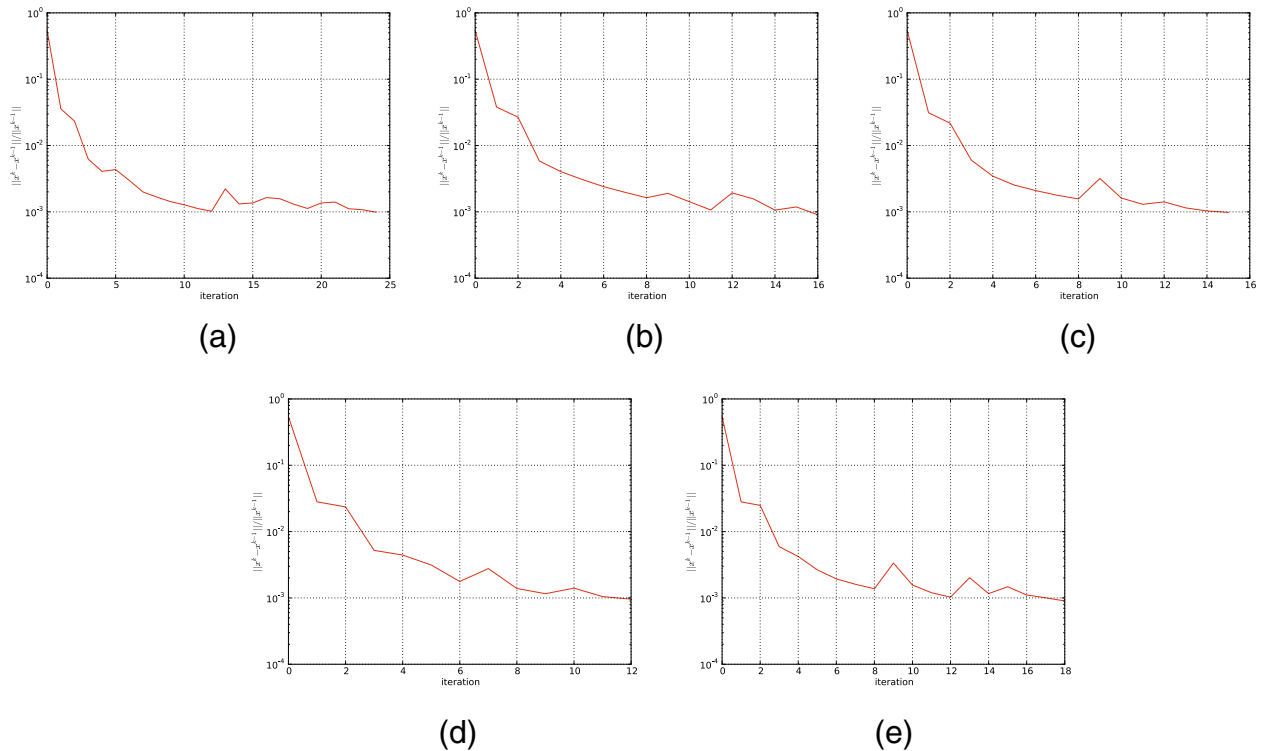

Figure 3 Example convergence curves for the finest scale of the proposed algorithm of the airplane-color restoration for five different synthetic non-parametric motion blurs: (a) Blur 1. (b) Blur 2. (c) Blur 3. (d) Blur 4 and (e) Blur 5.

Figure 4 Four different original images from [27]: (a) Image 1. (b) Image 2. (c) Image 3. (d) Image 4.

where $\mathbf{W}_d^k$ is a diagonal matrix with entries $\mathbf{W}_d^k(i,i) = (z_{d,i}^k)^{p/2-1}$.

2. Calculate

$$\mathbf{h}^{k+1} = \left[\beta^k (\mathbf{X}^k)^t (\mathbf{X}^k) + \gamma^k \right.$$

$$\left. \sum_{d \in \{h,v\}} (\mathbf{\Delta}^d)^t \mathbf{U}^k (\mathbf{\Delta}^d) \right]^{-1} \beta^k (\mathbf{X}^k)^t \mathbf{y}, \quad (22)$$

where $\mathbf{U}^k$ is a diagonal matrix with entries $\mathbf{U}^k(i,i) = (u_i^k)^{-1/2}$.

3. For each $d \in \{h, v, hh, vv, hv\}$ calculate

$$z_{d,i}^{k+1} = [\Delta_i^d(\mathbf{x}^k)]^2, \quad (23)$$

4. Calculate

$$u_i^{k+1} = [\Delta_i^h(\mathbf{h}^{k+1})]^2 + [\Delta_i^v(\mathbf{h}^{k+1})]^2, \quad (24)$$

5. Calculate

$$\alpha^{k+1} = \frac{\lambda_1 N/p}{\sum_{d \in D} 2^{1-o(d)} \sum_i |\Delta_i^d(\mathbf{x}^k)|^p}, \quad (25)$$

$$\beta^{k+1} = \frac{N}{\| \mathbf{y} - \mathbf{H}^{k+1}\mathbf{x}^k \|^2}, \quad (26)$$

$$\gamma^{k+1} = \frac{\lambda_2 N}{\mathrm{TV}(\mathbf{h}^{k+1})}, \quad (27)$$

In the line of study presented in [21] the parameter p is set to 4/5 (see [21] for a detailed discussion). Additionally, the parameters λ_1 and λ_2 are needed to approximate the partition functions for prior distributions p($\mathbf{x}|\alpha$) and p($\mathbf{h}|\gamma$), respectively. Unfortunately, the approximations of partition functions for the distributions p($\mathbf{x}|\alpha$) and p($\mathbf{h}|\gamma$) are necessary since its corresponding partition functions are analytically intractable. We follow the approaches proposed in [22,23], as already described in the previous section, and determine the values of the parameters λ_1 and λ_2 experimentally. The values of the parameters p, λ_1, and λ_2 are therefore set throughout all the experiments that follow. The robustness of the proposed method will be tested and evaluated under various blurring and noisy conditions.

Note that if the blur $\mathbf{h}$ and the hyperparameters α, β, and γ are assumed to be known, the proposed algorithm coincides with the iteratively re-weighted least squares (IRLS) algorithm presented in [21] (i.e., in this case for both algorithms the image estimate is calculated as shown in (21)). Note, that the l_p quasi norm based prior is also utilized in [15], and that this study simplifies the prior used in [21] by omitting the second order derivatives.

4 Multi-scale implementation

The restoration results presented in [12], and more recently in [27], showed the effectiveness of the multi-scale approach in implementing blind image deconvolution algorithms. Furthermore, it is shown in [27] that the multi-scale approach prevents the algorithm from converging to the unit impulse. Alternatively, the authors in [13,14] introduced heuristic re-weighting of the regularization parameters, at each iteration, to prevent the
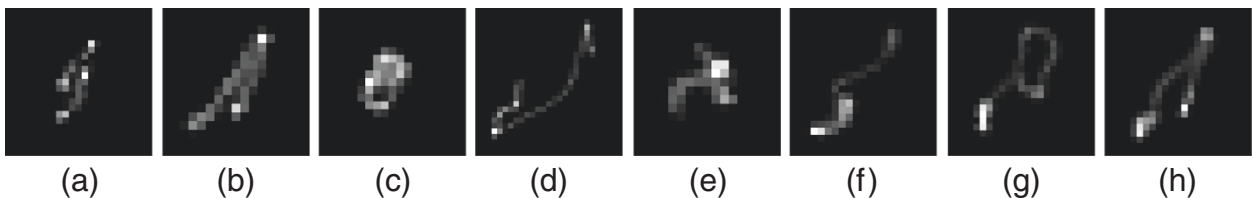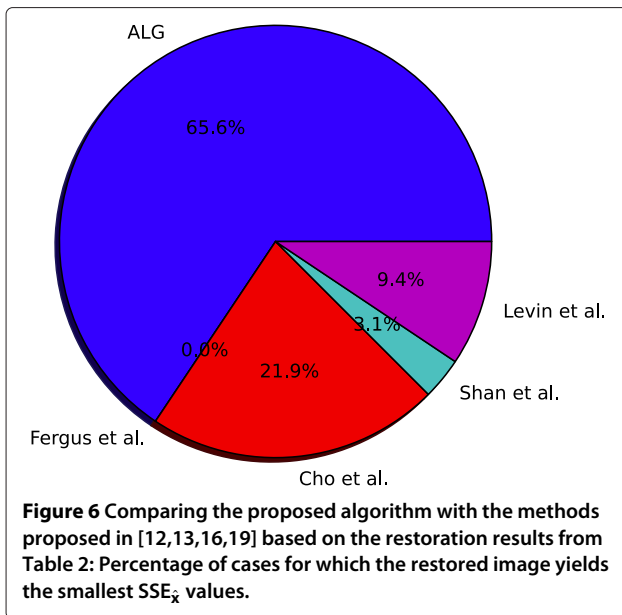

Figure 5 Eight different non-parametric motion blurs from [27]: (a) Blur 1. (b) Blur 2. (c) Blur 3. (d) Blur 4. (e) Blur 5. (f) Blur 6. (g) Blur 7 and (h) Blur 8.

Table 2 $SSE_{\hat{x}}$ and $ER_{\hat{x}}$ values, for the images and blurs defined in Figures 4 and 5, respectively

Blur	Method	Image 1		Image 2		Image 3		Image 4	
		$SSE_{\hat{x}}$	$ER_{\hat{x}}$	$SSE_{\hat{x}}$	$ER_{\hat{x}}$	$SSE_{\hat{x}}$	$ER_{\hat{x}}$	$SSE_{\hat{x}}$	$ER_{\hat{x}}$
Blur 1	ALG	**29.91**	**0.91**	**43.93**	**0.91**	**24.75**	**0.78**	43.73	1.46
	Fergus et al.	39.73	1.21	59.71	1.24	39.48	1.25	44.33	1.48
	Cho et al.	33.05	1.00	64.88	1.35	26.74	0.85	**31.59**	**1.05**
	Shan et al.	49.79	1.51	100.70	2.09	45.83	1.45	53.76	1.79
	Levin et al.	44.06	1.34	63.98	1.33	38.82	1.23	78.25	2.61
Blur 2	ALG	**33.47**	**0.89**	**50.00**	**0.97**	20.59	0.57	**42.18**	**0.96**
	Fergus et al.	40.70	1.08	66.12	1.29	41.55	1.15	93.14	2.11
	Cho et al.	34.13	0.91	53.42	1.04	30.18	0.83	102.24	2.32
	Shan et al.	38.87	1.03	313.48	6.11	29.74	0.82	146.73	3.33
	Levin et al.	48.50	1.29	74.05	1.44	46.73	1.29	128.82	2.92
Blur 3	ALG	**27.45**	1.06	**19.38**	**0.45**	**16.96**	0.92	17.78	1.16
	Fergus et al.	30.05	1.16	55.75	1.31	21.36	1.16	19.63	1.29
	Cho et al.	31.41	1.21	29.17	0.68	17.56	0.95	19.55	1.28
	Shan et al.	28.12	1.08	53.48	1.25	19.20	1.04	**16.33**	**1.07**
	Levin et al.	34.93	1.35	64.27	1.50	18.93	1.03	49.97	3.27
Blur 4	ALG	**51.59**	**1.08**	93.70	1.30	**28.69**	**0.74**	**44.15**	**1.11**
	Fergus et al.	125.80	2.64	112.92	1.56	81.28	2.09	11658.02	294.05
	Cho et al.	63.73	1.34	**80.16**	**1.11**	41.34	1.06	84.32	2.13
	Shan et al.	100.43	2.11	178.28	2.47	134.77	3.46	429.12	10.82
	Levin et al.	95.81	2.01	105.20	1.45	68.15	1.75	97.73	2.46
Blur 5	ALG	31.39	1.50	36.54	1.32	21.28	1.45	20.27	1.31
	Fergus et al.	27.32	1.31	39.50	1.42	21.58	1.47	20.61	1.34
	Cho et al.	38.59	1.85	**33.25**	**1.20**	23.30	1.59	**16.00**	**1.04**
	Shan et al.	30.76	1.47	51.94	1.87	17.71	1.21	20.85	1.35
	Levin et al.	**26.50**	**1.27**	35.98	1.29	**17.43**	**1.19**	34.01	2.20
Blur 6	ALG	**20.72**	**1.32**	**23.71**	**1.17**	**18.61**	**1.88**	**23.66**	**1.32**
	Fergus et al.	44.02	2.80	84.12	4.16	33.29	3.36	46.83	2.62
	Cho et al.	42.68	2.72	36.37	1.80	19.24	1.94	37.60	2.10
	Shan et al.	71.33	4.54	199.59	9.87	28.90	2.91	58.19	3.25
	Levin et al.	28.47	1.81	36.73	1.82	20.98	2.12	62.42	3.49
Blur 7	ALG	**38.51**	**1.90**	61.57	1.56	**20.74**	**1.59**	64.40	4.36
	Fergus et al.	206.70	10.22	152.18	3.86	137.37	10.53	501.31	33.90
	Cho et al.	43.46	2.15	**48.99**	**1.24**	26.81	2.06	**31.38**	**2.12**
	Shan et al.	252.56	12.49	250.20	6.34	230.76	17.69	300.06	20.29
	Levin et al.	45.91	2.27	64.07	1.62	27.05	2.07	97.97	6.63
Blur 8	ALG	**30.07**	**1.16**	**44.67**	**1.10**	33.23	1.43	77.69	3.37
	Fergus et al.	49.42	1.91	89.65	2.20	51.95	2.24	781.10	33.88
	Cho et al.	45.48	1.76	73.35	1.80	47.10	2.03	**41.63**	**1.81**
	Shan et al.	158.72	6.14	106.66	2.62	202.18	8.73	287.21	12.46
	Levin et al.	48.19	1.86	71.32	1.75	31.20	**1.35**	112.69	4.89

Entries of the table in bold denote the lowest values.

Figure 6 Comparing the proposed algorithm with the methods proposed in [12,13,16,19] based on the restoration results from Table 2: Percentage of cases for which the restored image yields the smallest $SSE_{\hat{x}}$ values.

algorithms from converging to unrealistic blur estimates. In our study, no heuristic adjustment of each parameter is performed but instead we estimate the parameters automatically within the Bayesian framework.

To avoid unrealistic blur estimates we adopt here a multi-scale scheme similar in spirit to the one proposed in [12]. Additionally, the proposed multi-scale approach

allows user to automatically initialize the proposed algorithm without visually inspecting the observed blurred image for determining the initial blur estimate. By analyzing the observed blurred image it is possible to come up with more informative initial blur estimates; however in this study our focus is to develop a completely automated algorithm once the blur support is provided.

The basic idea behind the multi-scale approach is to down-sample the observed blurred image to a number of low resolution images. At the lowest resolution the initial blur estimate (i.e., $\mathbf{h}^1$) is set to the uniform blur and the lowest resolution of the down-sampled observed image is utilized as the initial image estimate (i.e., $\mathbf{x}^1$). After convergence is achieved at each scale we up-sample the image and the blur estimates to the next higher resolution and re-run the proposed algorithm. This iterative process is repeated until the algorithm converges and image and blur estimates are obtained at their native resolutions. The detailed pseudocode used in the implementation of our multi-scale approach is shown in Appendix.

5 Experimental results

In this section, we present the experimental results obtained by the use of the proposed algorithm. As the performance metric, for the experiments in which the original image is known, we utilize the improvement in signal to noise ratio of the restored image (denoted as $\text{ISNR}_{\hat{x}}$), which is defined as $10 \log_{10} \left(\|\mathbf{x} - \mathbf{y}\|^2 / \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right)$,

Figure 7 Comparing the cumulative histogram of the proposed algorithm with the cumulative histogram of the methods proposed in [12,13,16,19] based on the restoration results from Table 2 in terms of the $ER_{\hat{x}}$ comparison metric.
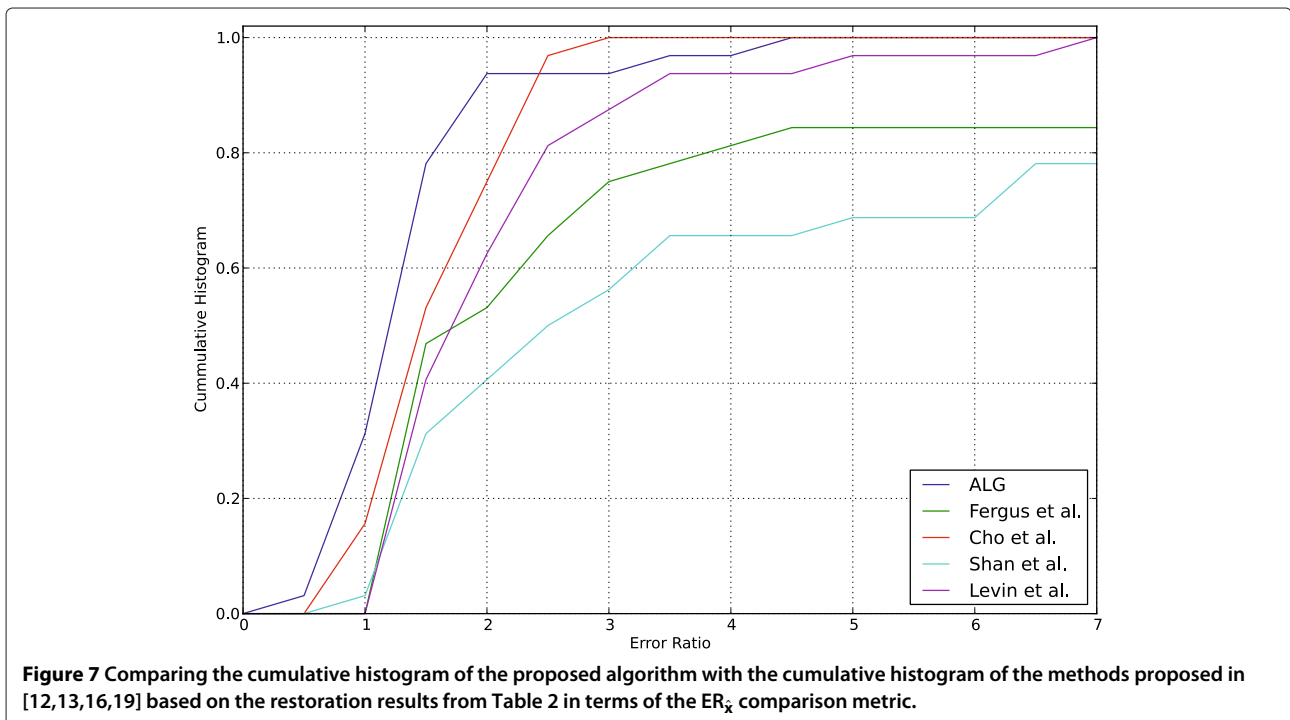
where $\mathbf{x}$, $\mathbf{y}$ and $\hat{\mathbf{x}}$ are the original, observed and estimated images, respectively. Analogously, when the blur is known we utilize the improvement in signal to noise ratio of the restored blur (denoted as $\text{ISNR}_{\hat{\mathbf{h}}}$), which is defined as $10 \log_{10} \left(\|\mathbf{h} - \mathbf{h}_\delta\|^2 / \|\mathbf{h} - \hat{\mathbf{h}}\|^2 \right)$, where $\mathbf{h}$, $\mathbf{h}_\delta$ and $\hat{\mathbf{h}}$ are

the original blur, the unit impulse, and the estimated blur, respectively.

In addition, after the blur support, greater than the original one, is specified by the user, all unknown parameters and the estimates of the unknown image and blur

Figure 8 Example restorations from Table 2: 1st column represents eight different blurred observations, 2nd column represents their respective original undistorted versions, 3rd column represents their respective restorations obtained by the proposed algorithm, 4th column represents their respective restorations obtained by the method proposed in [12], 5th column represents their respective original blurs, 6th column represents their respective restored blurs obtained by the proposed algorithm, 7th column represents their respective restored blurs obtained by the method proposed in [12].

are estimated automatically as described in Sections 3 and 4. Specifying the blur support for the unknown blur is common with the state-of-the-art approaches (see [12-14,16,19]).

Similarly to these approaches, the proposed algorithm is very robust when the support of the blur is largely overestimated by the user, as can be seen in all the experiments that follow. Also, for the experiments in which
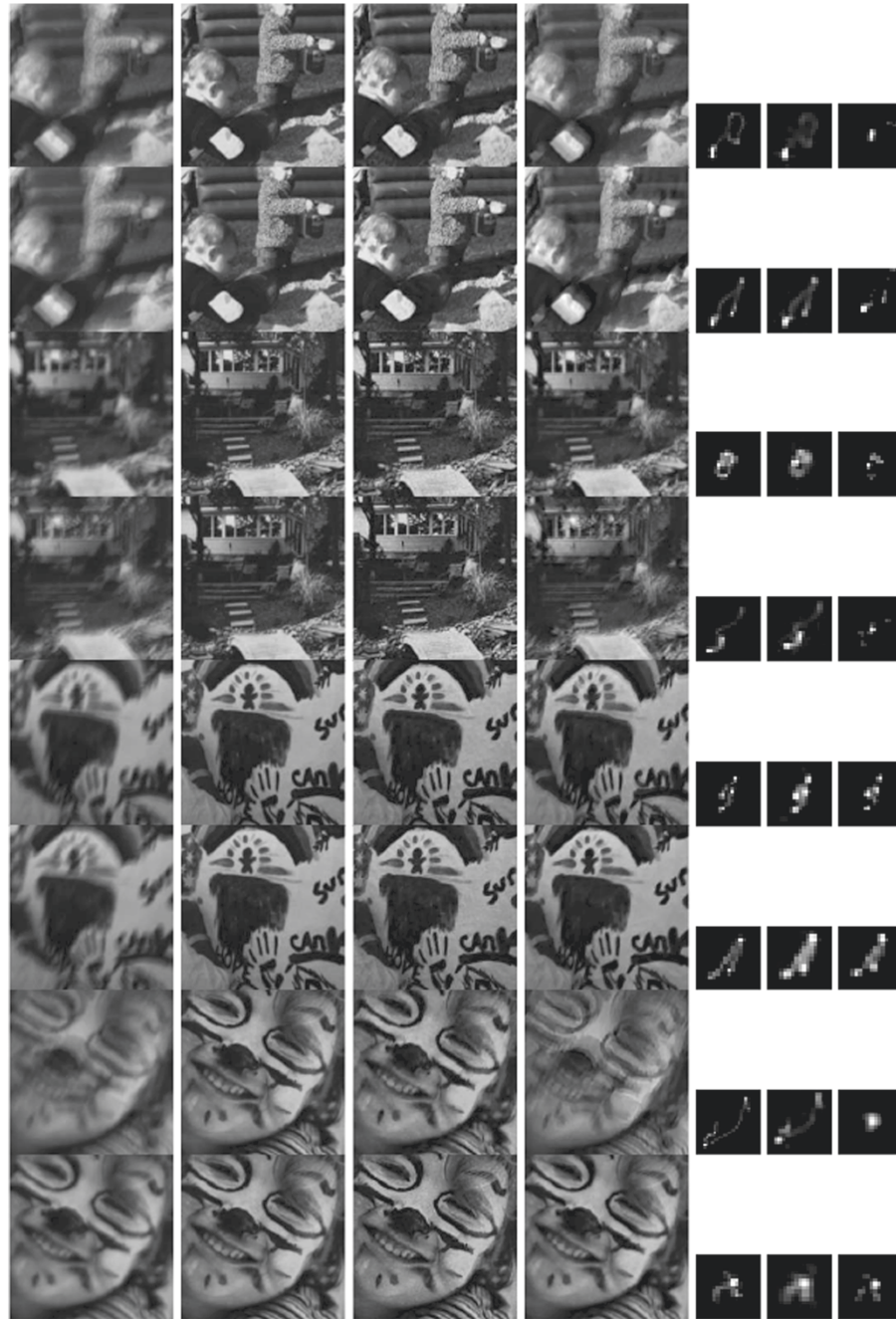

Figure 9 Example restorations from Table 2: 1st column represents eight different blurred observations, 2nd column represents their respective original undistorted versions, 3rd column represents their respective restorations obtained by the proposed algorithm, 4th column represents their respective restorations obtained by the method proposed in [16], 5th column represents their respective original blurs, 6th column represents their respective restored blurs obtained by the proposed algorithm, 7th column represents their respective restored blurs obtained by the method proposed in [16].

blurred colored images are considered, only the luminance component of the observed image is restored while the observed chroma components are used to obtain the restored colored image once the original luminance is estimated. Finally, the proposed algorithm is terminated when

the criterion $\|\mathbf{x}^k - \mathbf{x}^{k-1}\|/\|\mathbf{x}^{k-1}\| < 10^{-3}$ is achieved or the number of iterations reaches 100. After each iteration, we enforce the following constraints on the blur estimates: the positivity (blur elements less than zero are set to zero), the support constraint (blur elements
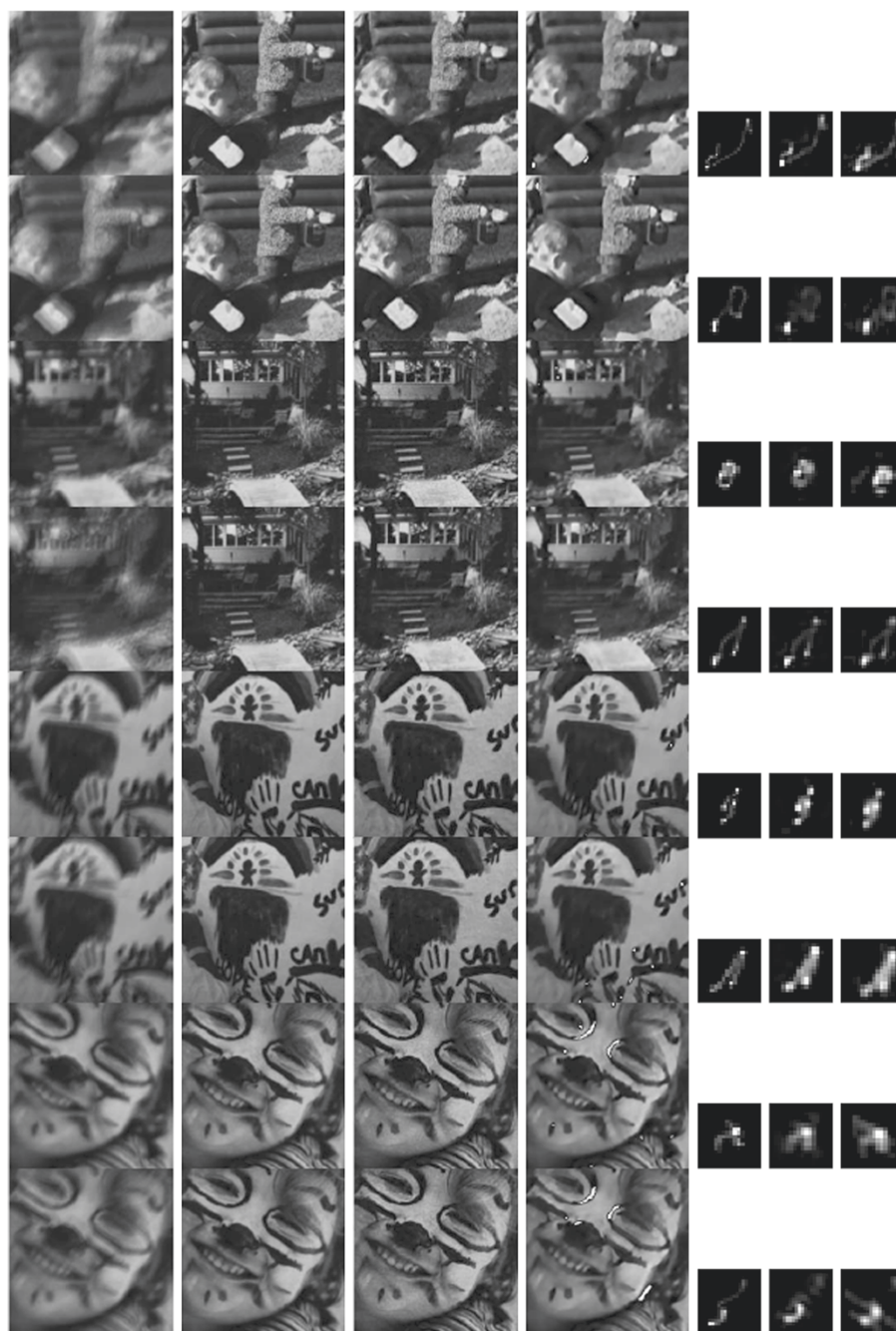

Figure 10 Example restorations from Table 2: 1st column represents eight different blurred observations, 2nd column represents their respective original undistorted versions, 3rd column represents their respective restorations obtained by the proposed algorithm, 4th column represents their respective restorations obtained by the method proposed in [13], 5th column represents their respective original blurs, 6th column represents their respective restored blurs obtained by the proposed algorithm, 7th column represents their respective restored blurs obtained by the method proposed in [13].

outside of the blur support estimate are set to zero), and the energy conservation (sum of the blur elements equals one).

In the first set of experiments, we evaluate the performance of the proposed method on four standard images (cameraman, satellite, shepp-logan phantom and airplane) which are widely used in image restoration experiments. The original images are then blurred with five different motion blurs, which are shown in Figure 1. Realizations of white Gaussian noise are added to the respective blurred

Figure 11 Example restorations from Table 2: 1st column represents eight different blurred observations, 2nd column represents their respective original undistorted versions, 3rd column represents their respective restorations obtained by the proposed algorithm, 4th column represents their respective restorations obtained by the method proposed in [19], 5th column represents their respective original blurs, 6th column represents their respective restored blurs obtained by the proposed algorithm, 7th column represents their respective restored blurs obtained by the method proposed in [19].
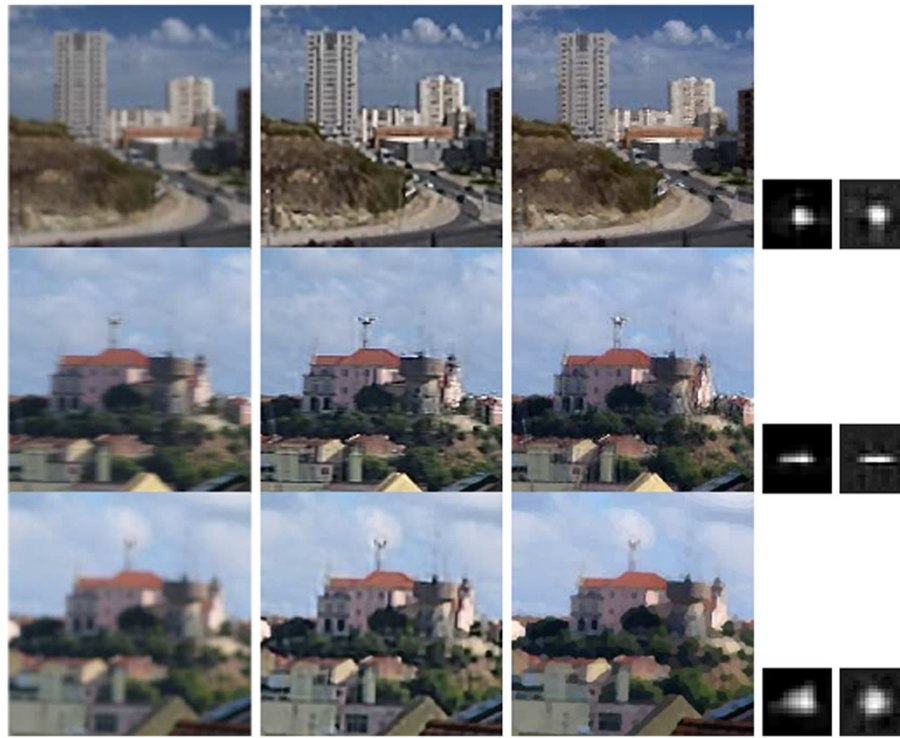
Figure 12 Comparing the proposed method with the method proposed in [14]: 1st column represents three different blurred observations, 2nd column represents their respective restorations obtained by the proposed algorithm, 3rd column represents their respective restorations obtained by the method proposed in [14], 4th column represents their respective blurs obtained by the proposed algorithm, 5th column represents their respective blurs obtained by the method proposed in [14].

images in order to obtain degraded images with the blurred signal to noise (BSNR) ratio of 40 dB. The blurred signal to noise ratio is defined as follows

$$BSNR = 10 \log_{10} \frac{Var(\mathbf{Hx})}{Var(\mathbf{n})}, \qquad (28)$$

where $Var(\cdot)$ denotes the variance of the random sequence. Example restorations obtained by the proposed algorithm are shown in Figure 2. The restoration results in terms of the previously defined $ISNR_{\hat{\mathbf{x}}}$ and $ISNR_{\hat{\mathbf{h}}}$, metrics are shown in Table 1. It can be observed from Table 1 that the proposed algorithm is very robust and it is capable of restoring the blurred images very successfully under various non-parametric motion blurs. Example convergence curves obtained by the proposed algorithm are shown in Figure 3.

In the second set of experiments, we evaluate the performance of the proposed method on a set of 32 blurred test images taken by a commercial camera. The test images are obtained from [27] and they are available online (www.wisdom.weizmann.ac.il/~levina/papers/LevinEtalCVPR09Data.zip). The set of 32 blurred

images was obtained by taking the original images shown in Figure 4 and by putting them side by side in order to form a calibration image. Once the calibration image was formed, a commercial camera was mounted on a tripod and eight photos of the calibration image were obtained. Note that during the acquisition process, the Z-axis rotation handle was locked in while the X-axis and Y-axis handles were loosened up in order to simulate in-plane camera shake (see [27] for details; resulting blurs are shown in Figure 5). In this study, we consider a comparison with the following methods [12,13,16,19]. For convenience, from now on, the methods proposed in [12,13,16], and the best method from [19] will be denoted, respectively, as Fergus et al., Shan et al., Cho et al., and Levin at al., while the proposed algorithm will be denoted as ALG.

The restoration results, for the second set of experiments where the original image and blur are both known, in terms of Sum of Squared Errors (i.e., $SSE_{\hat{\mathbf{x}}} = \|\mathbf{x} - \hat{\mathbf{x}}\|^2$) and the SSE ratio test (i.e., $ER_{\hat{\mathbf{x}}} = SSE_{\hat{\mathbf{x}}}/SSE_{\tilde{\mathbf{x}}}$) defined in [27] are shown in Table 2. Note that $\tilde{\mathbf{x}}$ is an image estimation obtained from the non-blind method from [27]. It can be observed from Table 2 that the proposed

algorithm is very robust and that it is capable of restoring blurred images taken by a commercial camera very successfully under various non-parametric motion blurs. In addition, the proposed algorithm is very competitive with the state-of-the-art methods. As can be noted in Table 2, the $SSE_{\hat{x}}$ metric for the proposed algorithm is respectively, on the average, 427, 89, 6, and 21 smaller than the $SSE_{\hat{x}}$ metric for the Fergus at al., Shan at al., Cho at al., and Levin et al. methods. Note that in Fergus' method, the blur estimation is performed separately from the image estimation. In order to understand the differences in the restoration results we provide some additional information. In Figure 6 it can be seen that for 66 % of tested cases the proposed algorithm yields the smallest $SSE_{\hat{x}}$ values. In addition, Figure 7 shows that for a number of test cases the proposed algorithm is capable of achieving very small $ER_{\hat{x}}$ values. For example, there are 78 % of the cases for which the proposed algorithm has $ER_{\hat{x}}$ smaller than 1.5 while at the same time (as the second best) there are 53 % of the test cases for which Cho et al. method achieves such condition. Example restorations and comparison with Fergus at al., Cho at al., Shan at al, and Levin at al. are shown, respectively, in Figures 8, 9, 10, and 11.

In the third set of experiments, we compare the performance of the proposed method with the method proposed in [14] by using the same set of blurred photographs as presented in [14]. Note that in the method proposed in [14] the parameters are not estimated but rather they are manually tuned which results in the sequence of numbers for each parameter. As can be seen in Figure 12 the restoration results obtained by the proposed algorithm are very competitive with the method proposed in [14]. It is clear from Figure 12 that the proposed algorithm produces much sharper restoration results with higher visual quality. Since we lack the true knowledge of the scene, the comparison metrics $SSE_{\hat{x}}$ and $ER_{\hat{x}}$ are undefined for this experiment.

6 Conclusions

In this article, a novel blind image deconvolution algorithm is presented. The proposed algorithm was developed within a Bayesian framework utilizing a non-convex l_p quasi norm based sparse prior on the image, and a total-variation prior on the unknown blur. The proposed algorithm is completely automated once the blur support is provided. Experimental results demonstrate that using sparse priors and the proposed parameter estimation, both the unknown image and blur can be estimated with very high accuracy. Furthermore, numerous restorations of photographs taken by commercial cameras are provided to demonstrate the robustness and effectiveness of the proposed approach. Finally, it was shown that the performance of the proposed algorithm is competitive

to existing state-of-the-art blind image deconvolution algorithms.

Appendix
Algorithm 1 is shown below.

Pseudocode for the multi-scale implementation.
Input: The observed $M \times N$ image (i.e., $\mathbf{y}$) and the initial $K \times K$ uniform blur (denoted as $\mathbf{h}_{init}$).
Output: The image (i.e., $\hat{\mathbf{x}}$) and blur (i.e., $\hat{\mathbf{h}}$) estimates.
begin

$S \longleftarrow \lceil \log_2(\min(M,N)/K) \rceil$
   – represents the total number of scales (see [12]);
$p \longleftarrow 4/5; \quad \lambda_1 \longleftarrow 1/2; \quad \lambda_2 \longleftarrow 1/16;$
$\mathbf{y}_S \longleftarrow \mathbf{y};$
for $s = 1, 2, \ldots, S$ **do**
   $\mathbf{y} \longleftarrow$ imresize $\left(\mathbf{y}_S, (3/2)^{1/2(s-S)}, \text{'bilinear'}\right)$ –
   MATLAB function;
   $N = \text{length}(\mathbf{y});$
   if $s = 1$ **then**
      $\mathbf{x}^1 \longleftarrow \mathbf{y};$
      $\mathbf{h}^1 \longleftarrow$ imresize$(\mathbf{h}_{init}, (3/2)^{1/2(s-S)},$
       'bilinear'); $\quad \mathbf{h}^1 \longleftarrow \mathbf{h}^1/\sum(\mathbf{h}^1);$
      $z^1_{d,i} \longleftarrow 1;$
      $\alpha^1 \longleftarrow \frac{\lambda_1 N/p}{\sum_{d \in D} 2^{1-o(d)} \sum_i |\Delta_i^d(\mathbf{x}^1)|^p};$
      $\beta^1 \longleftarrow \frac{N}{\|\mathbf{y}-\mathbf{H}^1\mathbf{x}^1\|^2};$
      $\gamma^1 \longleftarrow \frac{\lambda_2 N}{TV(\mathbf{h}^1)};$
   else
      $\mathbf{x}^1 \longleftarrow$ imresize$(\mathbf{x}, (3/2)^{1/2}, \text{'bilinear'});$
      $\mathbf{h}^1 \longleftarrow$ imresize$(\mathbf{h}, (3/2)^{1/2}, \text{'bilinear'});$
      $\mathbf{h}^1 \longleftarrow \mathbf{h}^1/\sum(\mathbf{h}^1);$
      $z^1_{d,i} \longleftarrow [\Delta_i^d(\mathbf{x}^1)]^2;$
      $\alpha^1 \longleftarrow \alpha; \quad \beta^1 \longleftarrow \beta; \quad \gamma^1 \longleftarrow \gamma$
      if $s = S$ **then**
         $\lambda_1 \longleftarrow 1;$
      end
   end
   $u_i^1 \longleftarrow [\Delta_i^h(\mathbf{h}^1)]^2 + [\Delta_i^v(\mathbf{h}^1)]^2;$
   $[\mathbf{x}, \mathbf{h}, \alpha, \beta, \gamma, z_{d,i}, u_i] \longleftarrow$
   Inference$(\mathbf{y}, \mathbf{h}^1, \alpha^1, \beta^1, \gamma^1, z^1_{d,i}, u_i^1)$
      – as outlined in Section 3;
end
$\hat{\mathbf{x}} \longleftarrow \mathbf{x};$
$\hat{\mathbf{h}} \longleftarrow \mathbf{h};$

end

Author details
[1] Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, 60208-3118, USA. [2] Departamento de Ciencias de la Computación e I.A., Universidad de Granada, 18071 Granada, Spain.

References
1. J Krist, in *Astronomical Data Analysis Software and Systems IV*, ed. by byRA Shaw, HE Payne, and JJE Hayes. Simulation of HST PSFs Using Tiny Tim (Astronomical Society of the Pacific, San Francisco, USA, 1995), pp. 349–353
2. TJ Schultz, Multiframe blind deconvolution of astronomical images. J. Opt. Soc. Am. A. **10**, 1064–1073 (1993)
3. T Bretschneider, P Bones, S McNeill, D Pairman, in *Proceedings of the American Society for Photogrammetry & Remote Sensing*. Image-based quality assessment of SPOT data, (2001). [Unpaginated CD-ROM]
4. FS Gibson, F Lanni, Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. J. Opt. Soc. Am. A. **8**, 1601–1613 (1991)
5. O Michailovich, D Adam, A novel approach to the 2-D blind deconvolution problem in medical ultrasound. IEEE Trans. Med. Imaging. **24**, 86–104 (2005)
6. M Roggemann, Limited degree-of-freedom adaptive optics and image reconstruction. Appl. Opt. **30**, 4227–4233 (1991)
7. P Nisenson, R Barakat, Partial atmospheric correction with adaptive optics. J. Opt. Soc. Am. A. **4**, 2249–2253 (1991)
8. AK RMCA Segall, Katsaggelos, High-resolution images from low-resolution compressed video. IEEE Signal Process. Mag. **20**(3), 37–48 (2003)
9. S Dai, M Yang, Y Wu, AK Katsaggelos, in *Proc. IEEE Int Image Processing Conf*. Tracking motion-blurred targets in video, (2006), pp. 2389–2392
10. CJKK Faulkner, M Louka, in *Third Int. Conf. on Image Proc. and Its Applications*. Veiling glare deconvolution of images produced by x-ray image intensifiers, (1989), pp. 669–673
11. TE Bishop, SD Babacan, B Amizic, AK Katsaggelos, T Chan, R Molina, *Blind Image Deconvolution: Problem Formulation and Existing Approaches*. (CRC Press , 2007)
12. R Fergus, B Singh, A Hertzmann, ST Roweis, WT Freeman, Removing camera shake from a single photograph. ACM Trans Graph. **25**(3), 787–794 (2006)
13. Q Shan, J Jia, A Agarwala, in *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*. High-quality motion deblurring from a single image (ACM, New York, 2008), pp. 1–10
14. M Almeida, L Almeida, Blind and semi-blind deblurring of natural images. IEEE Trans. Image Process. **19**, 36–52 (2010)
15. D Krishnan, R Fergus, in *Advances in Neural Information Processing Systems 22*, ed. by Y Bengio, D Schuurmans, J Lafferty, CKI Williams, and A Culotta. Fast image deconvolution using hyper-Laplacian priors, (2009), pp. 1033–1041
16. S Cho, S Lee, Fast motion deblurring. ACM Trans Graph. (SIGGRAPH ASIA 2009). **28**(5). (Article No. 145, 2009)
17. TS Cho, N Joshi, CL Zitnick, SB Kang, R Szeliski, WT Freeman, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. A content-aware image prior, (2010), pp. 169–176
18. T Hou, S Wang, H Qin, Image deconvolution with multi-stage convex relaxation and its perceptual evaluation. IEEE Trans. Image Process. **20**(12), 3383–3392 (2011)
19. A Levin, Y Weiss, F Durand, W Freeman, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Efficient marginal likelihood optimization in blind deconvolution, (2011), pp. 2657–2664
20. B Amizic, SD Babacan, R Molina, AK Katsaggelos, in *European Signal Processing Conference Eusipco*. Sparse Bayesian Blind Image Deconvolution with Parameter Estimation (Aalborg, Denmark, 2010), pp. 626–630
21. A Levin, R Fergus, F Durand, WT Freeman, in *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*. Image and depth from a conventional camera with a coded aperture (ACM, New York, 2007), p. 70
22. A Mohammad-Djafari, in *Maximum Entropy and Bayesian Methods*. A full bayesian approach for inverse problems (Kluwer Academic Publishers, 1996), pp. 135–143
23. J Bioucas-Dias, M Figueiredo, J Oliveira, in *Proceedings of EUSIPCO'2006*. Adaptive total-variation image deconvolution: a majorization-minimization approach (Florence, Italy, p. 2006
24. J Besag, On the statistical analysis of dirty pictures. J. Royal Stat. Soc. Series B (Methodological). **48**(3), 259–302 (1986)
25. S Babacan, R Molina, A Katsaggelos, Parameter estimation in TV image restoration using variational distribution approximation. IEEE Trans. Image Process. **17**(3), 326–339 (2008)
26. J Bioucas-Dias, M Figueiredo, J Oliveira, in *ICASSP*, vol. 2. Total-variation image deconvolution: a majorization-minimization approach, (2006), p. II
27. A Levin, Y Weiss, F Durand, WT Freeman, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition CVPR*. Understanding and evaluating blind deconvolution algorithms, (2009), pp. 1964–1971